Performance Modeling of Feature-Based Classification in SAR Imagery

Michael Boshra and Bir Bhanu

College of Engineering, University of California
Riverside, California 92521

ABSTRACT

We present a novel method for modeling the performance of a vote-based approach for target classification in SAR imagery. In this approach, the geometric locations of the scattering centers are used to represent 2-D model views of a 3-D target for a specific sensor under a given viewing condition (azimuth, depression and squint angles). Performance of such an approach is modeled in the presence of data uncertainty, occlusion, and clutter. The proposed method captures the structural similarity between model views, which plays an important role in determining the classification performance. In particular, performance would improve if the model views are dissimilar and vice versa. The method consists of the following steps. In the first step, given a bound on data uncertainty, model similarity is determined by finding feature correspondence in the space of relative translations between each pair of model views. In the second step, statistical analysis is carried out in the vote, occlusion and clutter space, in order to determine the probability of misclassifying each model view. In the third step, the misclassification probability is averaged for all model views to estimate the probability-of-correct-identification (PCI) plot as a function of occlusion and clutter rates. Validity of the method is demonstrated by comparing predicted PCI plots with ones that are obtained experimentally. Results are presented using both XPATCH and MSTAR SAR data.

Keywords: Performance prediction, Expected bounds on performance, Classification under occlusion, clutter, and data uncertainty, Model-based ATR, Model similarity

1. INTRODUCTION

Model-based automatic target recognition (ATR) has received considerable attention during the last two decades.[1,2,6] It can be defined as follows: Given a set of model targets and sensory data possibly belonging to one of these targets, the objective is to determine identity and/or pose of the scene target, if it exists, by comparing features extracted from the scene data with those of the model targets. The performance of the recognition process depends on properties of both the scene data (e.g., the extent of measurement error, missing and spurious features, etc.) and the model targets (e.g., articulation of model parts, similarity of model views, etc.). Systematic modeling of all these properties in a single approach has been a challenge for predicting the recognition performance.

In this paper, we present a method for predicting the probabilities of correct target identification and target misclassification in SAR imagery. The geometric locations of the scattering centers are used to represent the 2-D model views of a 3-D target under a given viewing condition (azimuth, depression and squint angles). A vote-based decision criterion is used for selecting the best target hypothesis; i.e., we choose the hypothesis with the maximum number of model features (votes) that are consistent with scene data features in an appropriate configuration. The following scene data and model properties are considered in our approach for performance modeling:

- *Distortion of Scene Data:* Three distortion types are considered: data uncertainty, missing scene features (occlusion), and spurious scene features (clutter).
- *Similarity of Model Views:* Model similarity means the degree of structural, or geometric, overlap between two target views. A model view which is considerably similar to other views can be misclassified as any of them under moderate levels of data distortion, whereas a distinct view can be correctly classified even under very high distortion levels.

Further author information -
Email: {michael, bhanu}@vislab.ucr.edu
URL: http://www.vislab.ucr.edu

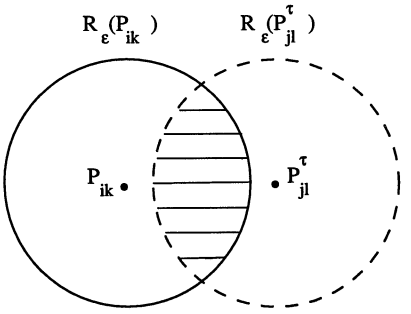


Figure 1. The similarity between P_{ik} and P^{τ}_{jl} is proportional to the area of the intersection (shaded) region.

Since, as mentioned in Section 3.2, data uncertainty is assumed to be uniform, it can be easily shown that this probability is proportional to the area corresponding to the intersection of the uncertainty regions associated with P_{ik}, and P^{τ}_{jl} (see Fig. 1):

$$FFS_{\epsilon}(P_{ik}, P^{\tau}_{jl}) = \frac{\mathrm{A}(R_{\epsilon}(P_{ik}) \cap R_{\epsilon}(P^{\tau}_{jl}))}{\mathrm{A}(R_{\epsilon}(P_{ik}))}$$

where $A(X)$ is the area of region X. It is clear that $FFS_{\epsilon}(P_{ik}, P^{\tau}_{jl}) > 0$, only if $\| P_{ik} - P^{\tau}_{jl} \| < 2\epsilon$. In such a case, features P_{ik} and P^{τ}_{jl} are said to be *similar*.

Next, we introduce a similarity measure between point P^{τ}_{jl} and view $\mathcal{M}_i$. We define this similarity to be the probability that an uncertain measurement of *any* feature in $\mathcal{M}_i$ can be interpreted as an uncertain measurement of P^{τ}_{jl}. Formally, *feature/view similarity* is defined as follows:

$$FVS_{\epsilon}(\mathcal{M}_i, P^{\tau}_{jl}) = 1 - \prod_{k}(1 - \frac{\mathrm{A}(R_{\epsilon}(P_{ik}) \cap R_{\epsilon}(P^{\tau}_{jl}))}{\mathrm{A}(R_{\epsilon}(P_{ik}))}). \tag{2}$$

Now, we turn to our original problem of measuring the similarity between $\mathcal{M}_i$ and $\mathcal{M}^{\tau}_j$. A plausible way to quantify such a similarity is to estimate the number of votes that $\mathcal{M}^{\tau}_j$ would get if the scene data view is an uncertain version of $\mathcal{M}_i$. That is, the similarity between $\mathcal{M}_i$ and $\mathcal{M}^{\tau}_j$ is the number of features in $\mathcal{M}^{\tau}_j$ that are consistent with uncertain measurements of the features in $\mathcal{M}_i$. This number of consistent features, denoted by $VVS_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j)$, is a random variable whose probability distribution can be defined as follows:

$$\Pr(VVS_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j) = x) \approx \sum_{\forall \mathcal{N}, |\mathcal{N}|=x} \prod_{l \in \mathcal{N}} FVS_{\epsilon}(\mathcal{M}_i, P^{\tau}_{jl}) \prod_{l \notin \mathcal{N}} (1 - FVS_{\epsilon}(\mathcal{M}_i, P^{\tau}_{jl}))$$

where $\mathcal{N}$ is a subset of feature indices in $\mathcal{M}^{\tau}_j$. The above equation is approximate since it assumes that the correspondence between similar features in $\mathcal{M}_i$ and $\mathcal{M}^{\tau}_j$ is many-to-one, thus making the terms $FVS_{\epsilon}(\mathcal{M}_i, P^{\tau}_{jl})$, $\forall l$, statistically independent.

It can be easily shown that $VVS_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j)$ is bounded by the number of *coincident* features, those having exactly the same location, in both views (lower bound), and the number of features in $\mathcal{M}^{\tau}_j$ that are similar to any feature in $\mathcal{M}_i$ (upper bound). In order to simplify the analysis, we choose our *view/view similarity* measure, $S_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j)$, to be the mean of $VVS_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j)$. That is,

$$\begin{aligned} S_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j) &= E(VVS_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j)) \\ &= \sum_{l} FVS_{\epsilon}(\mathcal{M}_i, P^{\tau}_{jl}). \end{aligned} \tag{3}$$

Notice that, in general, view/view similarity is not symmetric, i.e., $S_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j) \neq S_{\epsilon}(\mathcal{M}^{\tau}_j, \mathcal{M}_i)$. However, when there is a one-to-one correspondence between similar features in $\mathcal{M}_i$ and $\mathcal{M}^{\tau}_j$, $S_{\epsilon}(\mathcal{M}_i, \mathcal{M}^{\tau}_j)$ becomes symmetric.

3.4. Model Similarity: Variable-Location Case

In this section, we develop a method for measuring the similarity between views $\mathcal{M}_i$ and $\mathcal{M}_j$ as a function of the relative location between them. As will be seen later, this similarity function is fundamental for determining the misclassification probability. The obvious problem in this case is the infinite number of relative locations between $\mathcal{M}_i$ and $\mathcal{M}_j$. One possible approach is discretizing the transformation space and computing the similarity between the two views, equation (3), at each sample relative location. The problems associated with this approach are difficulty of "filling in the gaps" between samples, and complexity of determining the optimal sampling frequency. We alleviate the exhaustive sampling approach by considering that, for our purposes, we are interested in only those locations that result in a *minimum* degree of similarity between the two views, i.e., a minimum number of similar features.

Locations which result in a minimum degree of similarity between $\mathcal{M}_i$ and $\mathcal{M}_j$ can be obtained as follows. Features of $\mathcal{M}_i$ are aligned with those of $\mathcal{M}_j$ to generate $\mid \mathcal{M}_i \mid\mid \mathcal{M}_j \mid$ locations of $\mathcal{M}_j$. Each one of these locations corresponds to an instance of $\mathcal{M}_j$ which has at least a single feature that is coincident with one in $\mathcal{M}_i$. The estimated locations in this alignment-based approach become uncertain, when we consider the uncertainty associated with measured features of $\mathcal{M}_i$. For example, matching $P_{ik} \in \mathcal{M}_i$ with $P_{jl} \in \mathcal{M}_j$ generates a location $P_{ik} - P_{jl}$, with associated uncertainty region $R_\epsilon(P_{ik} - P_{jl})$. The set of generated instances, along with the associated uncertainty, covers *all* possible instances of $\mathcal{M}_j$ that share at least a single similar feature with $\mathcal{M}_i$, which is the minimum degree of similarity that we are interested in. We consider only these *uncertain instances* when determining the similarity between $\mathcal{M}_i$ and $\mathcal{M}_j$ (after eliminating redundant ones). Thus, our approach reduces the infinite number of possible instances that need to be considered to only a quadratic number of uncertain instances.

The next step is to compute the similarity between $\mathcal{M}_i$ and each uncertain instance of $\mathcal{M}_j$ defined by $\tau_\epsilon \subset \mathcal{T}$, which is generated by a scene/model feature match. Attempting to compute the similarity measure defined in (3) as a function of $\tau \in \tau_\epsilon$ is very complex, and appears to be solvable only through sampling of τ_ϵ. Since τ_ϵ is expected to be very small, we resort to a simpler solution by treating the infinite number of instances defined by τ_ϵ, as a *single* instance, $\mathcal{M}_j^{\tau_\epsilon}$, with *uncertainty* in its point features. Each $P_{jl}^{\tau_\epsilon} \in \mathcal{M}_j^{\tau_\epsilon}$ is a region that accommodates the variation in the location of feature P_{jl}^{τ}. It can be easily shown that $P_{jl}^{\tau_\epsilon} = R_\epsilon(P_{jl})$. The similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$, $S_\epsilon(\mathcal{M}_i, \mathcal{M}_j^{\tau_\epsilon})$, is measured in a manner similar to the fixed-location case (equation (3)). Specifically,

$$S_\epsilon(\mathcal{M}_i, \mathcal{M}_j^{\tau_\epsilon}) = \sum_l FVS_\epsilon(\mathcal{M}_i, P_{jl}^{\tau_\epsilon}) \tag{4}$$

where $FVS_\epsilon(\mathcal{M}_i, P_{jl}^{\tau_\epsilon})$ is a feature/view similarity measure between $P_{jl}^{\tau_\epsilon}$ and $\mathcal{M}_i$. This measure is defined as follows (see equation (2)):

$$FVS_\epsilon(\mathcal{M}_i, P_{jl}^{\tau_\epsilon}) = 1 - \prod_k \left(1 - \frac{\mathrm{A}(R_\epsilon(P_{ik}) \cap R_\epsilon(P_{jl}^{\tau_\epsilon}))}{\mathrm{A}(R_\epsilon(P_{ik}))}\right). \tag{5}$$

Region $R_\epsilon(P_{jl}^{\tau_\epsilon})$, the dilation of $P_{jl}^{\tau_\epsilon}$ by ϵ, is referred to the *uncertainty region* of $P_{jl}^{\tau_\epsilon}$. Notice that $FVS_\epsilon(\mathcal{M}_i, P_{jl}^{\tau_\epsilon}) > 0$, only if the center of $P_{jl}^{\tau_\epsilon}$ lies at a distance less than 3ϵ from any $P_{ik} \in \mathcal{M}_i$.

3.5. Effects of Scene Data and Model Properties on Classification Performance

In this section, we outline the effects of data uncertainty, occlusion, clutter, and model similarity on misclassification probability. In order to simplify both the presentation in this section and the probabilistic analysis in the next one, we assume that the correspondence between similar features in $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$ is one-to-one, i.e., each feature in $\mathcal{M}_i$ has at most one similar feature in $\mathcal{M}_j^{\tau_\epsilon}$, and vice versa. Further, we place a "threshold" on the degree of similarity between features in $\mathcal{M}_i$ and the corresponding ones in $\mathcal{M}_j^{\tau_\epsilon}$. This is done by considering the top most $\lfloor S_\epsilon(\mathcal{M}_i, \mathcal{M}_j^{\tau_\epsilon}) + 0.5 \rfloor$ similar features to be completely similar (i.e., $FFS_\epsilon(\cdot,\cdot) = 1$), and the rest to be completely dissimilar (i.e., $FFS_\epsilon(\cdot,\cdot) = 0$). Fig. 2 shows a conceptual view of $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$ after the thresholding process. Notice that, for a pair of similar features P_{ik} and $P_{jl}^{\tau_\epsilon}$, the uncertainty region of P_{ik}, $R_\epsilon(P_{ik})$, completely lies inside that of $P_{jl}^{\tau_\epsilon}$, $R_\epsilon(P_{jl}^{\tau_\epsilon})$.

The probability of misclassifying $\mathcal{M}_i$ as $\mathcal{M}_j^{\tau_\epsilon}$ depends on the sizes of similar and dissimilar feature sets. A schematic diagram of similar and dissimilar feature sets in both views is shown in Fig. 3. In this figure, the intersection area denotes the similar features in $\mathcal{M}_i$ ($\mathcal{M}_{ij}$) or $\mathcal{M}_j^{\tau_\epsilon}$ ($\mathcal{M}_{ji}^{\tau_\epsilon}$), while the other ones denote dissimilar features in $\mathcal{M}_i$ ($\mathcal{M}_{i/j} = \mathcal{M}_i - \mathcal{M}_{ij}$), and $\mathcal{M}_j^{\tau_\epsilon}$ ($\mathcal{M}_{j/i}^{\tau_\epsilon} = \mathcal{M}_j^{\tau_\epsilon} - \mathcal{M}_{ji}^{\tau_\epsilon}$). Note that, according to our assumption of

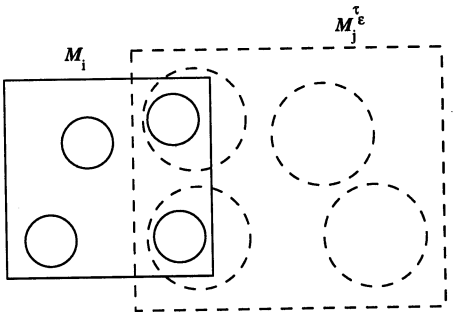


Figure 2. A conceptual view of $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$ after thresholding.

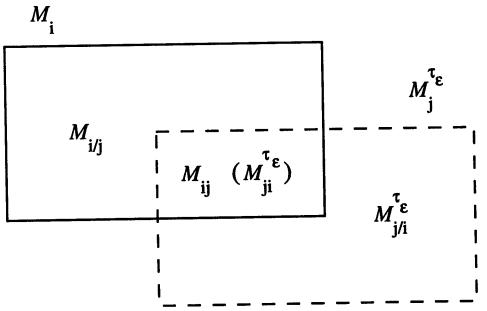


Figure 3. A schematic diagram of views $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$.

one-to-one correspondence between similar features, $\mathcal{M}_{ij}$ is of the same size as $\mathcal{M}_{ji}^{\tau_\epsilon}$. In our vote-based scheme, $\mathcal{M}_i$ gets misclassified as $\mathcal{M}_j^{\tau_\epsilon}$, if $\mathcal{M}_j^{\tau_\epsilon}$ receives more votes (i.e., has more consistent features) than $\mathcal{M}_i$.

The effects of scene and model factors on the vote count for both views can be outlined as follows:

- Initially, without any kind of distortion, $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$ get $\mid \mathcal{M}_i \mid$ and $S_0(\mathcal{M}_i, \mathcal{M}_j^{\tau_0})$ votes, respectively. Notice that $S_0(\mathcal{M}_i, \mathcal{M}_j^{\tau_0})$ is simply the number of coincident features in $\mathcal{M}_i$, and $\mathcal{M}_j^{\tau_0}$.

- As features in $\mathcal{M}_i$ get "blurred" due to data uncertainty, the similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$ increases (refer to equations (4) and (5)), thus increasing the expected number of votes for $\mathcal{M}_j^{\tau_\epsilon}$, while, obviously, keeping the same number of votes for $\mathcal{M}_i$. Notice that the expected number of votes for $\mathcal{M}_j^{\tau_\epsilon}$ becomes $\mid \mathcal{M}_{ij} \mid = S_\epsilon(\mathcal{M}_i, \mathcal{M}_j^{\tau_\epsilon})$ when the uncertainty bound reaches ϵ.

- The effect of occlusion depends on the number of occluded features in each of $\mathcal{M}_{i/j}$ and $\mathcal{M}_{ij}$. Occlusion of features in $\mathcal{M}_{i/j}$, which distinguish $\mathcal{M}_i$ from $\mathcal{M}_j^{\tau_\epsilon}$, will reduce the vote difference between the two views. On the other hand, occlusion of features in $\mathcal{M}_{ij}$ will reduce the number of votes for both of them, and so, using our vote-based selection criterion, it does not have an effect on the misclassification probability.

- Clutter features can accidently increase the number of votes for either of the two views. The probability that each view gets accidental votes is proportional to both the number of unmatched features, and the area of the uncertainty region associated with each feature. For $\mathcal{M}_i$, the unmatched features are the occluded ones, whereas for $\mathcal{M}_j^{\tau_\epsilon}$, the unmatched ones are $\mathcal{M}_{j/i}^{\tau_\epsilon}$ in addition to those features in $\mathcal{M}_{ji}^{\tau_\epsilon}$ that correspond to the occluded ones in $\mathcal{M}_{ij}$. The uncertainty regions corresponding to $\mathcal{M}_j^{\tau_\epsilon}$ are four-times larger than those of $\mathcal{M}_i$, and so $\mathcal{M}_j^{\tau_\epsilon}$ is likely to get more accidental votes than $\mathcal{M}_i$, even if it has less unmatched features. Notice that the area of the uncertainty region is $\pi\epsilon^2$ and $4\pi\epsilon^2$ for features of $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$, respectively. Another important point is concerning occluded features in $\mathcal{M}_{ij}$. A clutter feature that is consistent with one in $\mathcal{M}_{ij}$ will always be similar to the corresponding feature in $\mathcal{M}_{ji}^{\tau_\epsilon}$, but not vice versa. This is because, as mentioned, for similar pairs of features P_{ik} and $P_{jl}^{\tau_\epsilon}$, the uncertainty region of P_{ik} lies inside that of $P_{jl}^{\tau_\epsilon}$, which has a larger area (refer to Fig. 2).

The above trends are quantified in the next section in order to determine the misclassification probability.

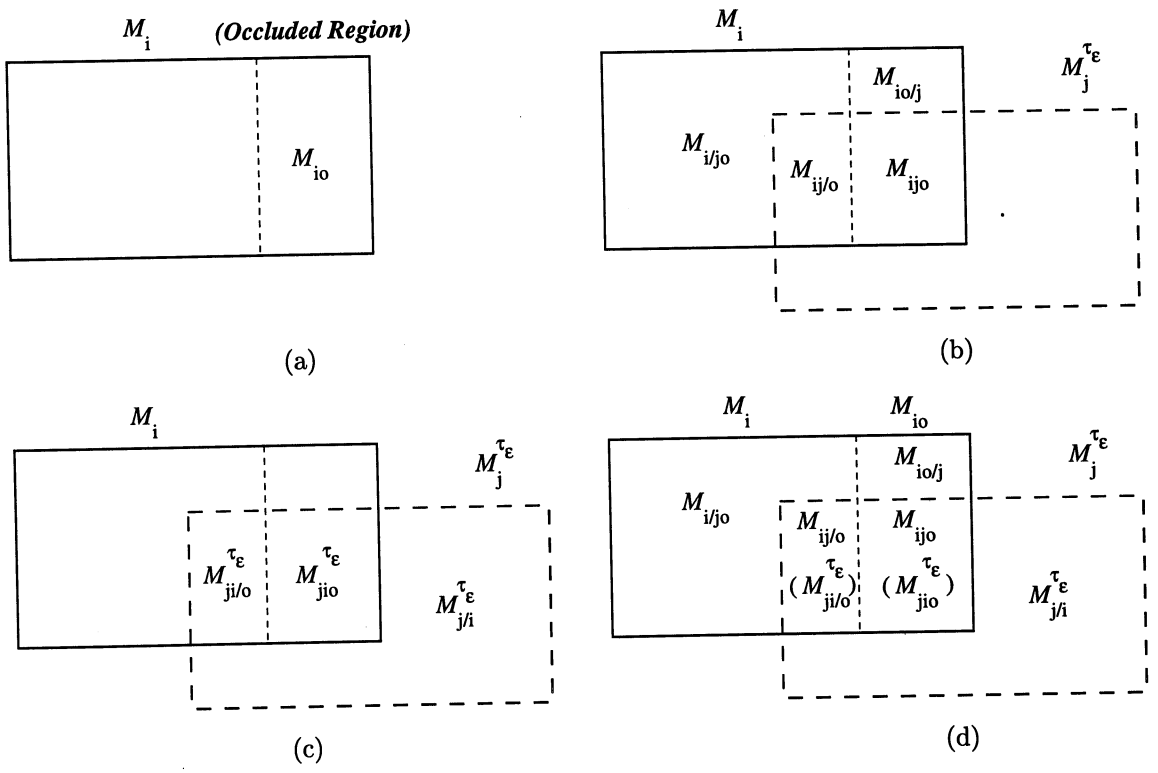


Figure 4. Schematic diagrams showing various regions: (a) occluded region in $\mathcal{M}_i$, (b) regions in $\mathcal{M}_i$, (c) regions in $\mathcal{M}_j^{\tau_\epsilon}$ and (d) all regions.

4. MISCLASSIFICATION PROBABILITY

In this section, we first derive the probability of misclassifying $\mathcal{M}_i$ considering only a single similar instance, $\mathcal{M}_j^{\tau_\epsilon}$, and then we determine the misclassification probability in the general case.

4.1. Single-Instance Case

Given scene view $\mathcal{S}_i(\epsilon, o, c)$, a distorted version of $\mathcal{M}_i$ (see Section 3.2), we would like to determine $\Pr(\mathcal{M}_j^{\tau_\epsilon}|\mathcal{M}_i, o, c)$, the probability of misclassifying $\mathcal{M}_i$ as $\mathcal{M}_j^{\tau_\epsilon}$ assuming that there are o and c occluded and clutter features, respectively. Let $V(\mathcal{M}, o, c)$ be the number of consistent features in $\mathcal{M}$ and $\mathcal{S}_i(\epsilon, o, c)$, i.e.,

$$V(\mathcal{M}, o, c) = \begin{cases} |\{X : X \in \mathcal{M}, \exists\, Y \in \mathcal{S}_i(\epsilon, o, c) \text{ s.t. } X \in Y\}| & \mathcal{M} = \mathcal{M}_i \\ |\{X : X \in \mathcal{M}, \exists\, Y \in \mathcal{S}_i(\epsilon, o, c) \text{ s.t. } X \cap Y \neq \Phi\}| & \mathcal{M} = \mathcal{M}_j^{\tau_\epsilon} \end{cases}.$$

Recall that views $\mathcal{S}_i(\epsilon, o, c)$ and $\mathcal{M}_j^{\tau_\epsilon}$ consist of uncertainty regions, while $\mathcal{M}_i$ consists of point features. Since our selection criterion is based on the number of votes, the misclassification probability can be simply written as follows:

$$\Pr(\mathcal{M}_j^{\tau_\epsilon}|\mathcal{M}_i, o, c) = \Pr(V(\mathcal{M}_j^{\tau_\epsilon}, o, c) > V(\mathcal{M}_i, o, c)) + 0.5 \times \Pr(V(\mathcal{M}_j^{\tau_\epsilon}, o, c) = V(\mathcal{M}_i, o, c)). \quad (6)$$

Notice that, in the above equation, ties are broken randomly, and accordingly, the probability of misclassification in case of a tie is 0.5.

Now, we derive an expression for the right hand side of (6). Using the notation $\mathcal{M}_{xy} = \mathcal{M}_x \cap \mathcal{M}_y$ and $\mathcal{M}_{x/y} = \mathcal{M}_x - \mathcal{M}_y$, we define the following sets:

- $\mathcal{M}_{io}$: set of occluded features in $\mathcal{M}_i$ (Fig. 4(a)),
- $\mathcal{M}_{i/jo}$ and $\mathcal{M}_{ij/o}$: subsets of unoccluded features in $\mathcal{M}_{i/j}$ and $\mathcal{M}_{ij}$, respectively (Fig. 4(b)),
- $\mathcal{M}_{io/j}$ and $\mathcal{M}_{ijo}$: subsets of occluded features in $\mathcal{M}_{i/j}$ and $\mathcal{M}_{ij}$, respectively (Fig. 4(b)), and

- $\mathcal{M}_{jio}^{\tau_\epsilon}$ and $\mathcal{M}_{ji/o}^{\tau_\epsilon}$: subsets of features in $\mathcal{M}_{ji}^{\tau_\epsilon}$ that are similar to those in $\mathcal{M}_{ijo}$, and $\mathcal{M}_{ij/o}$, respectively (Fig. 4(c)).

The above regions are shown together in Fig. 4(d). From the discussion in Section 3.5, it can be shown that the misclassification probability depends on the following random variables:

1. The number of unoccluded features in $\mathcal{M}_{i/j}$ ($U = \mid \mathcal{M}_{i/jo} \mid$), and
2. The number of clutter features that are consistent with those in $\mathcal{M}_{io/j}$ (V), $\mathcal{M}_{ijo}$ (W_1), $\mathcal{M}_{j/i}^{\tau_\epsilon}$ (W_2), and $\mathcal{M}_{jio}^{\tau_\epsilon}$ but *not* $\mathcal{M}_{ijo}$ (W_3).

Notice that the last case (corresponding to W_3) is possible since the uncertainty region of each feature in $\mathcal{M}_{ijo}$ is assumed to lie inside its similar one in $\mathcal{M}_{jio}^{\tau_\epsilon}$ (refer to Fig. 2). Ignoring the possibility that a feature in $\mathcal{M}_{jio}^{\tau_\epsilon}$ gets more than one clutter feature, which has a very small probability, we can write the number of votes for each view as follows:

$$V(\mathcal{M}_i, o, c) = U + V + \mid \mathcal{M}_{ij/o} \mid + W_1 \tag{7}$$

$$V(\mathcal{M}_j^{\tau_\epsilon}, o, c) \approx \mid \mathcal{M}_{ij/o} \mid + W_1 + W_2 + W_3. \tag{8}$$

Accordingly, the misclassification probability is

$$\Pr(\mathcal{M}_j^{\tau_\epsilon} | \mathcal{M}_i, o, c) \approx \Pr(W > U + V) + 0.5 \times \Pr(W = U + V)$$

where $W = W_2 + W_3$. The above equation can be rewritten as follows:

$$\Pr(\mathcal{M}_j^{\tau_\epsilon} | \mathcal{M}_i, o, c) \approx \sum_u \Pr(U = u|o) \sum_v \sum_{w \geq u+v} (1 - \frac{\delta(w - v - u)}{2}) \Pr(V = v, W = w | u, o, c) \tag{9}$$

where $\delta(x)$ is an impulse function that returns 1 if $x = 0$, and 0 otherwise. Note that

$$u \in [\max(0, \mid \mathcal{M}_{i/j} \mid - o), \min(\mid \mathcal{M}_{i/j} \mid - o + \mid \mathcal{M}_{ij} \mid, \mid \mathcal{M}_{i/j} \mid)].$$

We begin by determining the *first component* of the right hand side in (9), $\Pr(U = u|o)$. Since we are assuming that all feature subsets of $\mathcal{M}_i$ of size o are equally likely to be occluded, the probability that u features in $\mathcal{M}_{i/j}$ will remain unoccluded follows a hypergeometric distribution. In particular,

$$\Pr(U = u|o) = H_U(u; \mid \mathcal{M}_i \mid - o, \mid \mathcal{M}_{i/j} \mid, \mid \mathcal{M}_{ij} \mid) \tag{10}$$

where

$$H_X(x; n, a, b) = \Pr(X = x | n, a, b) = \frac{C(a, x) C(b, n - x)}{C(a + b, n)}$$

and $C(x, y) = \frac{x!}{(x-y)!\ y!}$.

The *second component* of the right hand side in (9), $\Pr(V = v, W = w | u, o, c)$, is determined as follows. We are interested in two non-connected sub-regions of the clutter area I (which is assumed to cover both $\mathcal{M}_i$ and $\mathcal{M}_j^{\tau_\epsilon}$):

1. I_V: union of $R_\epsilon(P)$ for all the ($\mid \mathcal{M}_{i/j} \mid - u$) occluded features $P \in \mathcal{M}_{i/j}$, and
2. I_W: union of $R_\epsilon(P^{\tau_\epsilon})$ for all $P^{\tau_\epsilon} \in \mathcal{M}_{j/i}^{\tau_\epsilon} \cup \mathcal{M}_{jio}^{\tau_\epsilon}$ *minus* union of $R_\epsilon(P)$ for all occluded features $P \in \mathcal{M}_{ijo}$.

Let $K_c(R)$ be the number of clutter features that fall in region R out of c features. Under the assumption of uniform clutter and small data uncertainty ϵ (refer to Section 3.2), we ignore the probability that more than one clutter feature is consistent with the same model feature, which is very small. Accordingly, we make the following approximation:

$$\Pr(V = v, W = w | u, o, c) \approx \Pr(K_c(I_V) = v, K_c(I_W) = w). \tag{11}$$

It is easy to show that $\Pr(K_c(I_V) = v, K_c(I_W) = w)$ follows a multinomial distribution. That is,

$$\Pr(K_c(I_V) = v, K_c(I_W) = w) = M_{V,W}(v, w; c, p_V, p_W) \tag{12}$$

where $p_V = \frac{A(I_V)}{A(I)}$, $p_W = \frac{A(I_W)}{A(I)}$, and

$$M_{X,Y}(x, y; n, p_x, p_y) = \frac{n!}{x!\ y!\ (n-x-y)!} p_x^x p_y^y (1 - p_x - p_y)^{n-x-y}.$$

The areas of regions I_V and I_W are

$$\begin{aligned} A(I_V) &\approx (\mid \mathcal{M}_{i/j} \mid -u)\pi\epsilon^2, \text{ and} \\ A(I_W) &\approx (\mid \mathcal{M}^{\tau_\epsilon}_{j/i} \mid + \mid \mathcal{M}^{\tau_\epsilon}_{jio} \mid) * 4\pi\epsilon^2 - \mid \mathcal{M}_{ijo} \mid * \pi\epsilon^2. \end{aligned}$$

The above estimates are approximate because they assume that the uncertainty regions are non-overlapping.

Substituting from (10), (11), and (12) into (9), we get the misclassification probability:

$$\begin{aligned} &\Pr(\mathcal{M}_j^{\tau_\epsilon} | \mathcal{M}_i, o, c) \approx \\ &\sum_u H_U(u; \mid \mathcal{M}_i \mid -o, \mid \mathcal{M}_{i/j} \mid, \mid \mathcal{M}_{ij} \mid) \sum_v \sum_{w \geq v+u} (1 - \frac{\delta(w-v-u)}{2}) M_{V,W}(v, w; c, p_V, p_W). \end{aligned} \tag{13}$$

4.2. General Case

In this section, we consider the general classification problem. That is, given a model database $\mathcal{D}$, we would like to determine the probability of misclassifying a distorted version of view $\mathcal{M}_i \in \mathcal{D}$ as any other view, $\mathcal{M}_j \in \mathcal{D}$, such that $TARGET(\mathcal{M}_i) \neq TARGET(\mathcal{M}_j)$. We define $\mathcal{D}_i$ to be the set of view instances that are similar to $\mathcal{M}_i$. As discussed in Section 3.4, the size of $\mathcal{D}_i$ is of order $O((\mid \mathcal{D} \mid -1) \mid \mathcal{M}_i \mid M)$, where M is the average size of target views in $\mathcal{D}$. The probability of misclassifying $\mathcal{M}_i$ as *any* instance $\mathcal{M}_j^{\tau_{j\epsilon}} \in \mathcal{D}_i$, expressed as $\Pr(\mathcal{D}_i | \mathcal{M}_i, o, c)$, can be written as follows:

$$\begin{aligned} \Pr(\mathcal{D}_i | \mathcal{M}_i, o, c) = &\sum_j \Pr(\mathcal{M}_j^{\tau_{j\epsilon}} | \mathcal{M}_i, o, c) - \\ &\sum_j \sum_{k \neq j} \Pr(\mathcal{M}_j^{\tau_{j\epsilon}} \wedge \mathcal{M}_k^{\tau_{k\epsilon}} | \mathcal{M}_i, o, c) + \cdots - (-1)^{|\mathcal{D}_i|} \Pr(\wedge_{j=1}^{|\mathcal{D}_i|} \mathcal{M}_j^{\tau_{j\epsilon}} | \mathcal{M}_i, o, c) \end{aligned} \tag{14}$$

where $\Pr(\wedge_j \mathcal{M}_j^{\tau_{j\epsilon}} | \mathcal{M}_i, o, c)$ is the probability that *every* instance $\mathcal{M}_j^{\tau_{j\epsilon}}$ gets a higher ranking than $\mathcal{M}_i$, when evaluated by the vote-based criterion. Initially, at low occlusion and clutter levels, all the terms in (14) are almost equal to zero, and the probability of correctly identifying the target view is almost one. Then, as these levels increase, the probability of misclassification becomes non-negligible and the value of the first term in (14) increases. However, at this point and for a certain occlusion/clutter range, the values of the remaining terms stay almost zero. This can be explained as follows. In the majority of cases, we can expect little or no overlapping between each pair of instances, $\mathcal{M}_j^{\tau_{j\epsilon}}$ and $\mathcal{M}_k^{\tau_{k\epsilon}}$. Accordingly, there is a "competition" between these instances to: 1) get more votes which can be provided by a finite number of clutter features, and 2) avoid losing votes, which can take place if some of the finite number of occluded features happen to be consistent with either of the two instances. These factors result in a negative correlation between the events, $\Pr(\mathcal{M}_j^{\tau_{j\epsilon}} | \mathcal{M}_i, o, c)$ and $\Pr(\mathcal{M}_k^{\tau_{k\epsilon}} | \mathcal{M}_i, o, c)$. For a certain occlusion/clutter range, this negative correlation is strong enough to result in $\sum_j \sum_{k \neq j} \Pr(\mathcal{M}_j^{\tau_{j\epsilon}} \wedge \mathcal{M}_k^{\tau_{k\epsilon}} | \mathcal{M}_i, o, c)$ as well as the higher-order terms to be almost zero. Beyond this range, the value of this term becomes non-negligible, while the higher-order ones remain of negligible values, and so on.

For our purposes, we use only the first-order approximation of (14), i.e.,

$$\Pr(\mathcal{D}_i|\mathcal{M}_i, o, c) \approx \sum_j \Pr(\mathcal{M}_j^{\tau_{j\epsilon}}|\mathcal{M}_i, o, c). \tag{15}$$

Interestingly, as will be shown in the next section, we have found that the first-order approximation can yield accurate estimates up to relatively high occlusion and clutter levels. Beyond these levels, equation (15) can be interpreted as an *upper bound* on the misclassification probability.

5. EXPERIMENTAL RESULTS

In this section, we validate the proposed performance modeling method by comparing predicted PCI plots with ones that are determined experimentally. Our recognition system is based on the concept of geometric hashing. A hash table is constructed which maps every pair of scattering centers in a model view to a tuple, which consists of target type, azimuth, and location of the reference scattering center. This tuple is stored in the entry that corresponds to the relative locations between the scattering centers, which are translation-invariant. In the recognition phase, hash-table keys are formed by determining the relative distances between scattering center pairs in the scene view. Data uncertainty is accommodated by considering the four neighbors of the scattering centers in the key formation process. The hash table is accessed using these keys to retrieve a number of tuples. These tuples are used to generate votes about the target type, azimuth and 2-D location in the scene data, which are accumulated in a 4-D vote table. The target corresponding to the entry with the highest number of votes is selected. Assuming the 4-neighbor uncertainty region, this algorithm examines nearly all the target, azimuth, and discrete location space, and so its performance is almost optimal.

The PCI plot is predicted as follows:

1. For each pair of model views $\mathcal{M}_i$ and $\mathcal{M}_j$, $TARGET(\mathcal{M}_i) \neq TARGET(\mathcal{M}_j)$, we determine the similarity between them as explained in Sections 3.3 and 3.4. The statistics of the similarity between considered pairs of views is stored in a 3-D *similarity histogram*. Each entry in this histogram corresponds to the number of occurrences of the triple $(\mid \mathcal{M}_i \mid, \mid \mathcal{M}_j^{\tau_\epsilon} \mid, S_\epsilon(\mathcal{M}_i, \mathcal{M}_j^{\tau_\epsilon}))$. In general, the size of this histogram is of the order $O(V^2W)$, where V is the range of model view sizes (number of features), and W is the maximum size of a model view. In our experiments, we choose a fixed number of the strongest scattering centers per view, N. Accordingly, the similarity histogram in our case, which is 1-D, is of size N.

2. For given occlusion and clutter rates, the average probability of misclassification is determined as follows. For each non-zero entry in the similarity histogram, we compute the view-instance misclassification probability, defined in equation (13), assuming the parameters corresponding to that entry. The resulting probabilities are then weighted by the values of the corresponding entries, added, and then divided by the total number of views in the database. Notice that this is equivalent to computing the general-case misclassification probability, defined in equation (15), for all model views and then taking the average.

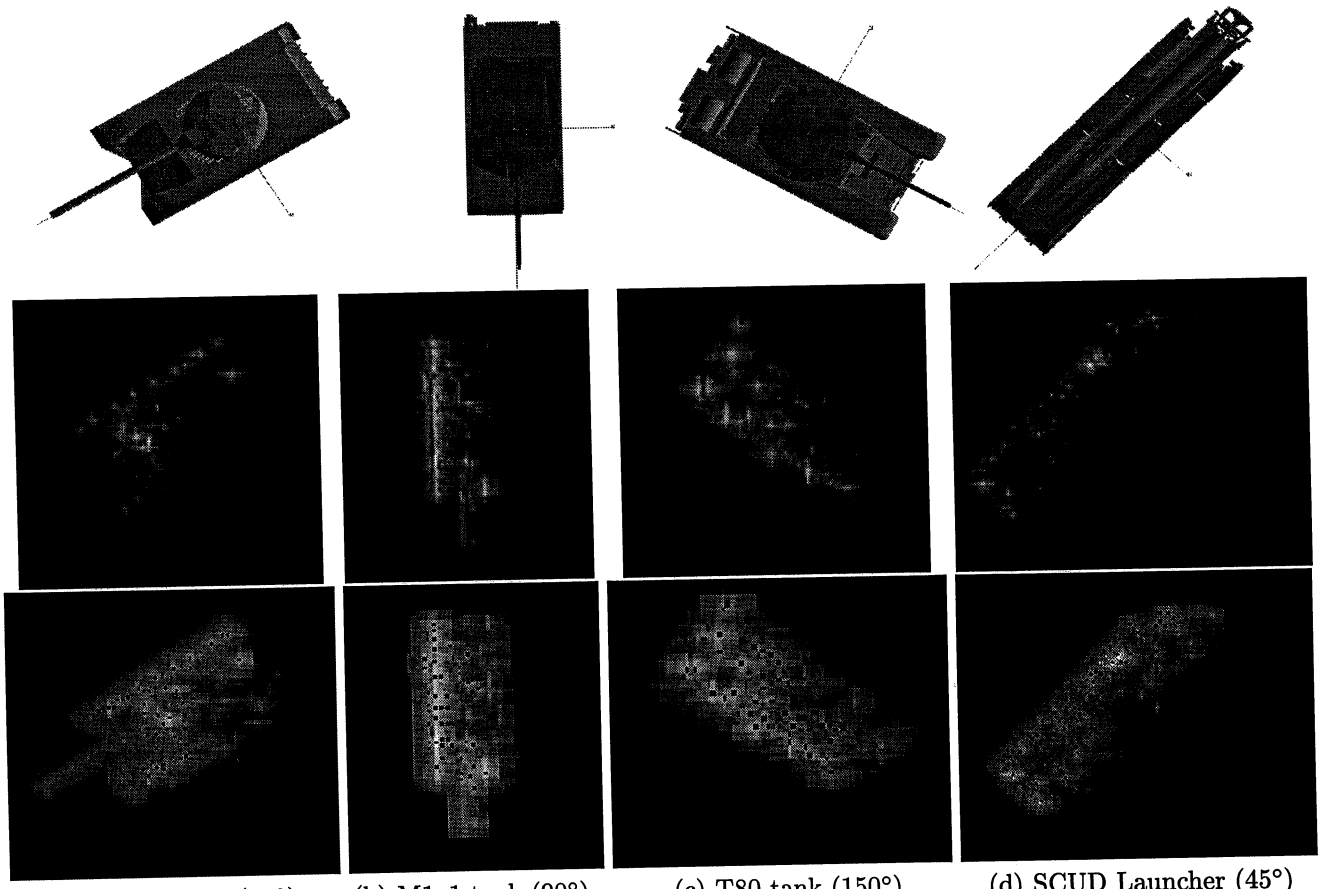
3. The process in the previous step is repeated for various occlusion and clutter rates to construct the PCI plot.

5.1. Experiments Using XPATCH Data

We use the XPATCH data to test the performance of recognition under occlusion and clutter; without taking data uncertainty into consideration. These data are generated using XPATCH3 radar signature prediction code (six inch resolution, 15° depression angle, 90° squint angle).

Model Data: A model database is selected which consists of four targets: T72, T80, and M1a1 tanks, and a SCUD missile launcher. Fig. 5 shows these targets, along with sample views and extracted scattering centers. Each target is represented by 360 views in steps of 1° azimuth, and so the total size of the model database is 1440 views.

Test Data: Two view sizes are selected for our experiments: $N = 20$ and 40. For each view size, we generate test data by introducing varying amounts of clutter and occlusion to the model views. A view is occluded by randomly selecting a feature and eliminating it, along with a number of its nearest neighbors depending upon the occlusion rate. Clutter is simulated by adding a number of random features, depending upon the clutter rate, within the bounding box of the target. This process is repeated four times for each model view, and so the size of a test data



(a) T72 tank (30°) (b) M1a1 tank (90°) (c) T80 tank (150°) (d) SCUD Launcher (45°)

Figure 5. Examples of model targets, corresponding XPATCH SAR images and extracted scattering centers, superimposed on 8-bit SAR images.

set for specific occlusion and clutter rates is 5760 views. Test sets are generated for various occlusion and clutter rates which are assumed to be the same in our experiments.

Results: The PCI plot is experimentally estimated by using the model data to build the hash table, and then supplying the test data to the recognition system. The PCI plot is predicted using the method outlined above. Both plots are shown in Figs. 6(a) and 6(b) for $N = 20$ and 40, respectively. From these figures, we observe that the predicted PCI plots closely match the ones obtained experimentally. In particular, they accurately predict the occlusion/clutter range within which recognition is almost perfect (up to 65% and 75% for $N = 20$ and 40, respectively), and the sharp degradation in performance beyond this range. Further, we observe that the predictions become conservative beyond the knee point in the plots. This is expected since, as mentioned in Section 4.2, the expression for the misclassification probability, equation (15), is interpreted as a lower bound when the higher-order terms in equation (14) become of non-negligible values, which is the case when the recognition performance degrades.

5.2. Experiments Using MSTAR Data

We use publicly available MSTAR data to test the performance of recognition under data uncertainty, occlusion and clutter.

Model Data: The selected model database consists of three targets: T72, BMP2 and BTR. The model views are taken at a depression angle of 17°. Fig. 7 shows sample SAR views of the three targets, along with the corresponding regions of interest (ROI's) with extracted scattering centers superimposed on them. The database consists of 697 views: 231 for the T72, and 233 for each of the BMP2 and the BTR.

Test Data: The strongest 30 scattering centers are used to represent model and test views (i.e., $N = 30$). Two different types of test data are obtained:

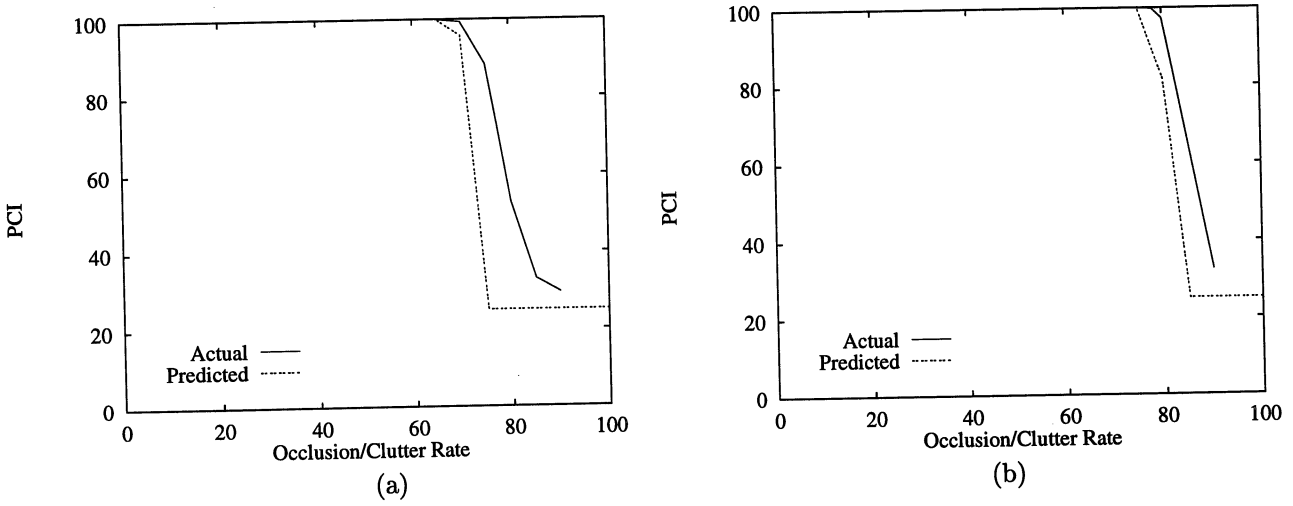
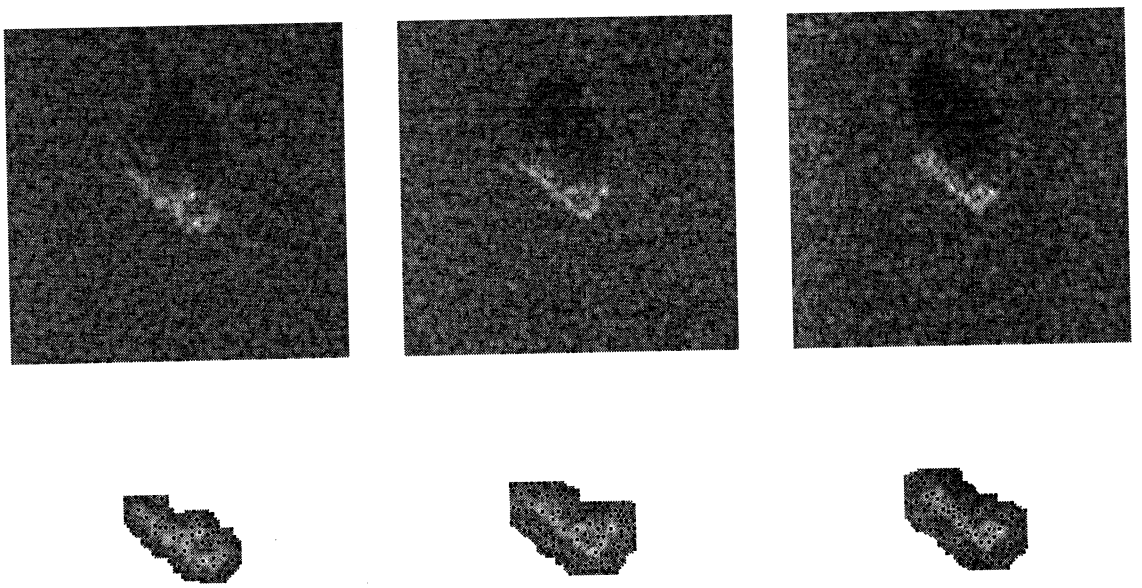


Figure 6. Actual and predicted PCI plots for the XPATCH experiments, considering two different view sizes (N): (a) $N = 20$, (b) $N = 40$. The horizontal axis corresponds to the occlusion and clutter rates which are assumed to be equal.



(a) T72 (132°) (b) BMP2 (132°) (c) BTR (132°)

Figure 7. Examples of MSTAR target images, corresponding ROI's and extracted scattering centers, superimposed on the ROI's.

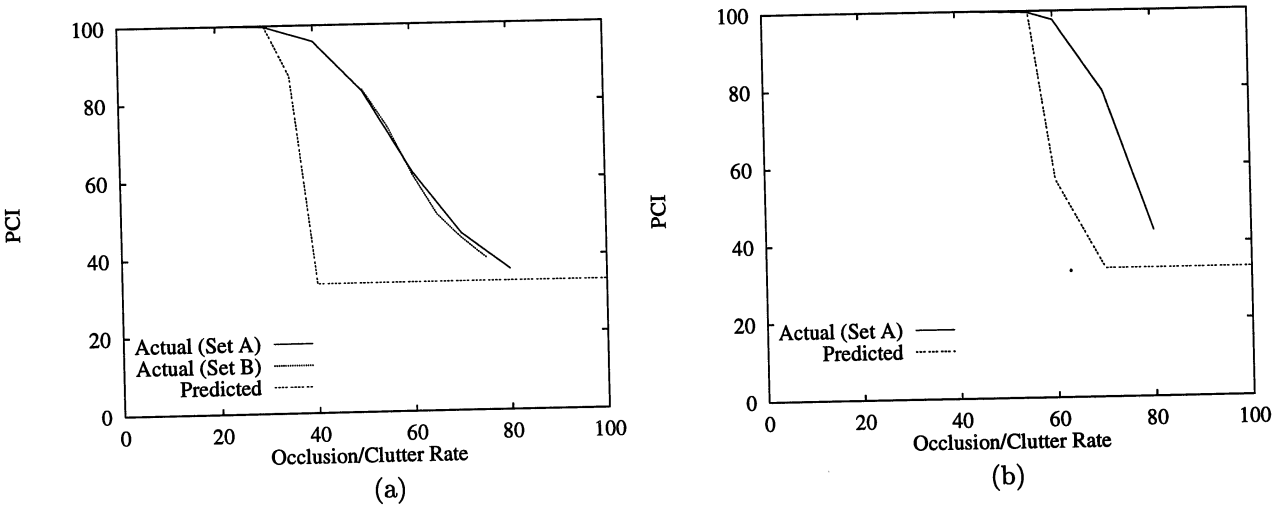


Figure 8. Actual and predicted PCI plots for the MSTAR experiments, considering two different bounds on data uncertainty (ϵ): (a) $\epsilon = 1.0$, (b) $\epsilon = 0.6$. The horizontal axis corresponds to the occlusion and clutter rates which are assumed to be equal.

- *Test Set A:* The model views are distorted by adding occlusion and clutter as explained above, and injecting data uncertainty by randomly perturbing the scattering center locations, such that the Euclidean distance between the new location and the original one does not exceed ϵ. Two values of ϵ are considered: $\epsilon = 0.6$ and 1.0. The distortion process is repeated four times for each model view, and so the size of a test data set for specific distortion parameters is 2788 views.

- *Test Set B:* MSTAR views of the same model targets at a slightly different depression angle (15°) are used. This test data set consists of 581 views: 193 for the T72, and 194 for each of the BMP2 and the BTR. Below, we describe the method used to estimate the distortion parameters, (o, c, ϵ), associated with this set. Additional test sets are obtained by adding a variety of occlusion/clutter rates to the original data set, as described above. As in the previous cases, each view is distorted four times, and so the data size for a given occlusion/clutter rate is 2324 views.

Estimation of Distortion Parameters for Original Test Set B: For each original test view at some azimuth angle z, we search for the best match, in the translation space, with the three model views of the same target at azimuth angles z, $z-1$ and $z+1$. In order to accommodate data uncertainty in the matching process, we consider two scattering centers to be consistent if they are 4-adjacent under the translation transformation. The average number of unmatched features in test and model views correspond to clutter and occlusion rates, respectively. Notice that, in our case, these rates are the same since we are using fixed view sizes. The occlusion/clutter rate is estimated to be approximately 49%. Under the assumption of uniform distribution for data uncertainty, it can be shown that the expected value of the Euclidean distance between a measured location of a scattering center and the "ideal" one is 0.707ϵ. Equating this expression with the average Euclidean distance between consistent features in best test/model matches, we obtain an estimate of ϵ, which is found to be approximately 1 pixel.

Results: Figures 8(a) and 8(b) show experimentally-determined and predicted PCI plots for $\epsilon = 1.0$ and 0.6, respectively. As in the XPATCH experiments, we observe that our method accurately predicts the occlusion/clutter range within which recognition is almost perfect (up to 30% and 50% for $\epsilon = 1.0$ and 0.6, respectively). We also observe that, beyond the knee point, the predictions are more conservative than in the XPATCH case. This can be explained as follows. The increase in the data uncertainty level, represented by ϵ, enlarges the uncertainty regions associated with instance features. This will increase the likelihood that instances receive more clutter features, and, subsequently, get more votes than the corresponding target view. Accordingly, the values of the higher-order terms in equation (14) will increase, thus reducing the tightness of the first-order approximation that we are using (equation (15)). Finally, note that our method quantifies the effect of uncertainty on recognition. It highlights the need for accurate feature-extraction methods that can estimate scattering-center locations at sub-pixel accuracy.

6. CONCLUSIONS

A statistical framework has been presented for estimating the probability of target misclassification in the presence of data uncertainty, occlusion and clutter. The method is based on capturing the structural similarity between each pair of model views that do not belong to the same target, assuming a fixed data uncertainty bound. Then, using the computed similarity information, the probability of misclassification is derived assuming statistical models for occlusion and clutter. Validity of the method has been demonstrated by comparing predicted PCI plots with those that are obtained experimentally using both XPATCH and MSTAR data. We are currently extending this method to estimate the probability of misclassifying clutter as target, and vice versa.

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. on Image Processing*, 6(1):1–6, 1997.
2. B. Bhanu and T. Jones. Image understanding research for automatic target recognition. *IEEE AES Systems Magazine*, 8:15–23, 1993.
3. W. E. L. Grimson and D. P. Huttenlocher. On the verification of hypothesized matches in model-based recognition. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 13(12):1201–1213, 1991.
4. M. Lindenbaum. Bounds on shape recognition performance. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 17(7):665–680, 1995.
5. M. Lindenbaum. An integrated model for evaluating the amount of data required for reliable recognition. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 19(11):1251–1264, 1997.
6. J. A. Ratches, C. P. Walters, R. G. Buser, and B. D. Guenther. Aided and automatic target recognition based upon sensory inputs from image forming systems. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 19(9):1004–1019, 1997.
7. K. B. Sarachik. The effect of Gaussian error in object recognition. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 19(4):289–301, 1997.