Performance Prediction and Validation for Object Recognition*

Michael Boshra and Bir Bhanu
Center for Research in Intelligent Systems
University of California, Riverside, California 92521
{michael,bhanu}@cris.ucr.edu
http://www.cris.ucr.edu/

Abstract

This paper addresses the problem of predicting fundamental performance of vote-based object recognition using 2-D point features. It presents a method for predicting a tight lower bound on performance. Unlike previous approaches, the proposed method considers data-distortion factors, namely uncertainty, occlusion, and clutter, in addition to model similarity, simultaneously. The similarity between every pair of model objects is captured by comparing their structures as a function of the relative transformation between them. This information is used along with statistical models of the data-distortion factors to determine an upper bound on the probability of recognition error. This bound is directly used to determine a lower bound on the probability of correct recognition. The validity of the method is experimentally demonstrated using synthetic aperture radar (SAR) data obtained under different depression angles and target configurations.

1 Introduction

Model-based object recognition is concerned with identifying and localizing objects from scene data. It is performed by extracting features from the scene data and searching for a consistent correspondence between scene features and those of a model object. Performance of such a process depends on: *1)* distortion of the scene data (e.g., sensor noise, missing and spurious features), *2)* nature of the model objects (e.g., degree of similarity between objects, object articulation), and *3)* criterion used to evaluate scene/model matches (e.g., number of votes, probabilistic likelihood). It has been a challenge to model these factors in a single approach for performance prediction, which is a crucial step for the advancement of the field.

In this paper, we present a method for predicting a tight lower bound on object recognition performance. We assume that both scene data and model objects are represented by 2-D point features. In addition, the matching criterion is assumed to be vote-based; i.e., a specific object/pose hypothesis is evaluated based on the number of model features that are consistent with at least a single scene feature. The proposed method considers the following: *1) Data-Distortion Factors:* uncertainty (positional), occlusion (missing features), and clutter (spurious features), and *2) Model Factor:* structural similarity (degree of structural "overlap" between pairs of model objects).

Our performance-prediction problem can be formally defined as follows. We are given: *1)* a set of model objects, $\mathcal{DB} = \{\mathcal{M}_i\}$, where each object $\mathcal{M}_i$ is represented by a set of 2-D point features, $\{F_{ij}\}$, that are discretized at some resolution, *2)* statistical models for uncertainty, occlusion, and clutter, and *3)* a class of applicable transformations, $\mathcal{T}$ (e.g., translation, rigid, affine), which is a discretized set since we are dealing with discretized features. Our objective is to predict a lower bound on the *probability-of-correct-recognition* (PCR), as a function of occlusion and clutter rates, assuming a fixed uncertainty model and a vote-based matching criterion. Note that recognition involves determining *both* object identity and pose.

The remainder of the paper is organized as follows. The next section summarizes related research and our contributions. Section 3 presents an overview of the method. Modeling of data-distortion factors is outlined in Section 4. Object-similarity issues are discussed in Section 5. Section 6 presents the derivation of a lower bound on PCR. In Section 7, we demonstrate the validity of the proposed method by comparing predicted PCR plots with actual ones. Finally, Section 8 provides conclusions.

2 Related Research

Traditionally, object recognition performance is empirically determined through extensive experimentation. Formal analysis of performance has not re-

*This work was supported in part by DARPA/AFOSR grant F49620-97-1-0184; the contents and information do not reflect positions or policies of the U.S. Government.

0-7695-0149-4/99 $10.00 © 1999 IEEE

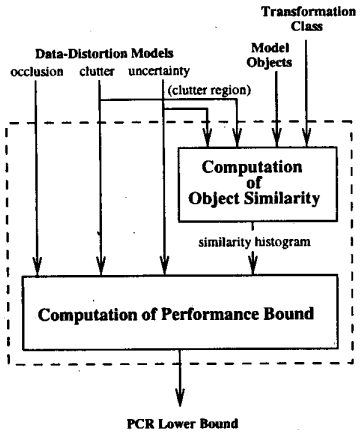


Figure 1: Proposed performance-prediction method.

ceived much attention in the literature. Some research efforts analyze the problem of discriminating objects from random clutter (e.g., [3, 4]). Other efforts address the problem of object recognition performance, considering data uncertainty and object similarity (implicitly) [2], or uncertainty and similarity, with partial handling of occlusion [5].

The *main contribution* of this paper is the development of a performance-prediction method that simultaneously considers uncertainty, occlusion, clutter and object similarity. As mentioned above, previous approaches consider only a subset of these factors. The performance predicted by our method is fundamental, since it is obtained by analyzing the amount of information provided by both scene data and model objects, *independent* of the particular algorithm used for vote-based recognition (e.g., alignment, vote accumulation, tree search). The proposed method is validated using synthetic aperture radar (SAR) data obtained under different depression angles and target configurations (variations of the same basic model).

3 Overview of Our Method

Figure 1 shows a block diagram of the proposed method. Its major components are described below.

• **Data-Distortion Models:** Data distortion factors are statistically modeled as follows: *1) Uncertainty:* the actual location of a scene feature is described by a probability mass function (PMF), which is assumed to be uniform. *2) Occlusion:* it is assumed to be uniform, where each subset of features is equally likely to be occluded as any other subset that is of the same size. *3) Clutter:* clutter features are assumed to be uniformly distributed within some area surrounding the object.

• **Computation of Object Similarity:** The objective of this stage is to compute the similarity between all pairs of model objects. The similarity of an object, $\mathcal{M}_j$, to another one, $\mathcal{M}_i$, is defined as the number of votes that $\mathcal{M}_j$ would get given an "uncertain" version of $\mathcal{M}_i$, as a function of the relative transformations between the two objects. Notice that the space of relative transformations is defined by the applicable transformation class $\mathcal{T}$. It can be seen that, in addition to structural object overlap, our definition of object similarity depends on data uncertainty as well. Obviously, this leads to a probabilistic similarity function. The output of this stage is a *similarity histogram*, which accumulates information derived from the similarity functions of all pairs of model objects.

• **Computation of Performance Bound:** The objective of this stage is to compute a lower bound on PCR. It is obtained by determining an upper bound on the probability that votes for any erroneous object/pose hypothesis reach or exceed those for the correct hypothesis, given specific levels of data distortion. As shown in Figure 1, this stage utilizes both object similarity information provided by the previous stage, and data-distortion models.

4 Data-Distortion Models

In this section, we formally define the three sources of data distortion considered in this paper.

• **Uncertainty:** The effect of uncertainty is to perturb positions of object features according to some PMF. Since we are assuming uniform distribution, this PMF can be represented by an *uncertainty region*, $R_u(\cdot)$. An uncertain version of a model object, $\mathcal{M}_i$, can be represented as

$$\mathcal{D}_u(\mathcal{M}_i, R_u(\cdot)) = \{P_u(R_u(F_{ij}))\}$$

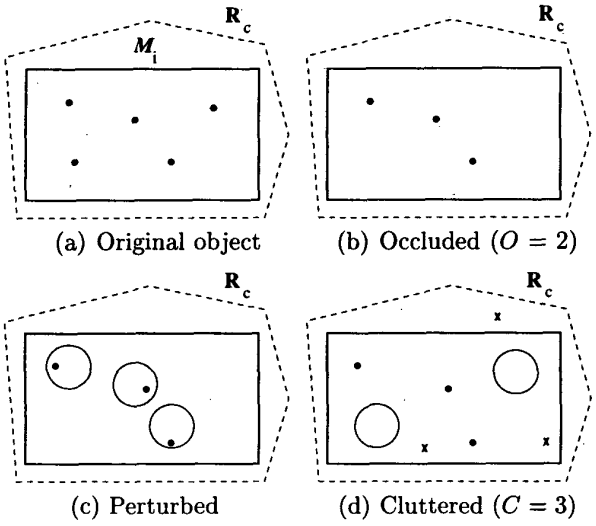
where $P_u(R)$ is a function that returns a randomly-selected feature within region R.

• **Occlusion:** The effect of occlusion is to eliminate some object features. An occluded version of object $\mathcal{M}_i$ can be represented as follows:

$$\mathcal{D}_o(\mathcal{M}_i, O) = \mathcal{M}_i - \mathcal{P}_o(\mathcal{M}_i, O)$$

where $\mathcal{P}_o(\mathcal{M}_i, O)$ is a function that returns a randomly-selected subset of O features from $\mathcal{M}_i$. Development of more sophisticated occlusion models that capture the spatial dependency among features is a subject of future research.

• **Clutter:** The effect of clutter is the addition of spurious features. We assume that these features are uniformly distributed within a *clutter region*, R_c, which can be of an arbitrary shape (e.g., convex hull of object features, bounding box, etc). An ambiguity takes



(a) Original object (b) Occluded ($O = 2$)

(c) Perturbed (d) Cluttered ($C = 3$)

Figure 2: An example of the distortion process. Notice that: 1) $R_u(\cdot)$ is a circle centered at the feature's true location, 2) clutter features are shown as crosses, and 3) there are no clutter features inside the uncertainty regions of occluded features (shown in (d)).

place if a clutter feature happens to fall within the uncertainty region of an occluded one, since it can not be differentiated from the no-occlusion/no-clutter case. In order to resolve this ambiguity, we restrict clutter features to lie outside the uncertainty regions of occluded ones. Formally, a cluttered version of $\mathcal{M}_i$ can be defined as follows:

$$\mathcal{D}_c(\mathcal{M}_i, C, R_c, R_x) = \mathcal{M}_i \cup \mathcal{P}_c(C, R_c - R_x)$$

where $\mathcal{P}_c(C, R)$ is a function that returns C randomly generated features within region R. In our context, R_x is the union of the uncertainty regions associated with the occluded features.

In general, a distorted version of object $\mathcal{M}_i$, $\mathcal{D}_i(R_u(\cdot), O, C, R_c)$, is obtained by occluding O features in $\mathcal{M}_i$, perturbing unoccluded ones within uncertainty region $R_u(\cdot)$, and then adding C clutter features within clutter region R_c (see Figure 2). Formally,

$$\mathcal{D}_i(R_u(\cdot), O, C, R_c) = \mathcal{D}_c(\mathcal{D}_u(\mathcal{D}_o(\mathcal{M}_i, O), R_u(\cdot)), C, R_c, R_x)$$

where $R_x = \cup_j R_u(F_{ij})$, $\forall F_{ij} \in (\mathcal{M}_i - \mathcal{D}_o(\mathcal{M}_i, O))$. We refer to $R_c - R_x$, or simply R'_c, as the *effective clutter region.*

5 Computation of Object Similarity

In this section, we formally define a similarity measure between model objects, and explain the method used to construct the similarity histogram.

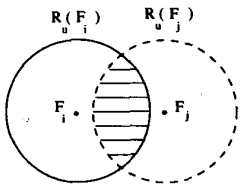


Figure 3: An illustration of feature/feature similarity.

5.1 Definition of Object Similarity

We introduce a sequence of definitions that lead to a quantitative measure of the similarity between a pair of model objects.

• **Feature Consistency:** A feature, F_i, is said to be consistent with another feature, F_j, if F_i can be interpreted as an uncertain measurement of F_j. It can be easily seen that the condition for consistency in our case is $F_i \in R_u(F_j)$.

• **Feature/Feature Similarity:** The similarity between features F_i and F_j, $S_{ff}(F_i, F_j)$, is defined as the probability that an uncertain measurement of F_i is consistent with F_j. The degree of similarity is proportional to the extent of overlap between $R_u(F_i)$ and $R_u(F_j)$ (see Figure 3). Specifically,

$$S_{ff}(F_i, F_j) = \frac{A(R_u(F_i) \cap R_u(F_j))}{A(R_u(F_i))}$$

where $A(R)$ is the area of region R. We refer to feature pairs with overlapping/non-overlapping regions as *similar/dissimilar* pairs.

• **Object/Feature Similarity:** The similarity between a model object, $\mathcal{M}_i$, and a feature, F_j, $S_{of}(\mathcal{M}_i, F_j)$, is defined as the probability that an uncertain measurement of *any* feature in $\mathcal{M}_i$ (i.e., a feature in $\mathcal{D}_u(\mathcal{M}_i, R_u(\cdot))$) is consistent with F_j. Formally,

$$S_{of}(\mathcal{M}_i, F_j) = 1 - \prod_k (1 - S_{ff}(F_{ik}, F_j)).$$

• **Object/Hypothesis Similarity:** Let $\mathcal{M}_j^\tau$ be a *hypothesis* of object $\mathcal{M}_j$ at pose $\tau \in \mathcal{T}$ relative to $\mathcal{M}_i$. The similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, $S_{oi}(\mathcal{M}_i, \mathcal{M}_j^\tau)$ or simply S_{ij}^τ, can be defined as the number of votes for $\mathcal{M}_j^\tau$ given $\mathcal{D}_u(\mathcal{M}_i, R_u(\cdot))$. That is, S_{ij}^τ is the number of features in $\mathcal{M}_j^\tau$ that have consistent uncertain measurements of features in $\mathcal{M}_i$. Obviously, S_{ij}^τ is a random variable. It can be easily shown that the bounds of S_{ij}^τ are

$$\begin{aligned} \min(S_{ij}^\tau) &= |\{F_{jk}^\tau : S_{of}(\mathcal{M}_i, F_{jk}^\tau) = 1\}| \\ \max(S_{ij}^\tau) &= |\{F_{jk}^\tau : S_{of}(\mathcal{M}_i, F_{jk}^\tau) > 0\}|, \end{aligned}$$

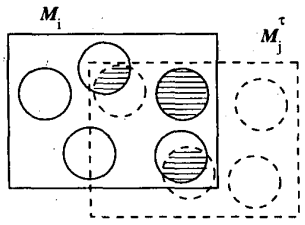


Figure 4: Similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$. Notice that $S_{ij}^\tau \in [1,3]$, and $E(S_{ij}^\tau) \approx 2$.

and its expected value is

$$E(S_{ij}^\tau) \approx \sum_k S_{of}(\mathcal{M}_i, F_{jk}^\tau),$$

where $F_{jk}^\tau \in \mathcal{M}_j^\tau$. An example is shown in Figure 4. The PMF of S_{ij}^τ is approximated by the following binomial distribution:

$$P_{S_{ij}^\tau}(s_{ij}^\tau) = B_{S_{ij}^\tau}(s_{ij}^\tau; N_{ij}^\tau, P_{ij}^\tau)$$

where $P_X(x) = \Pr[X = x]$, $N_{ij}^\tau = \max(S_{ij}^\tau)$, $P_{ij}^\tau = \frac{E(S_{ij}^\tau)}{N_{ij}^\tau}$, $B_X(x;n,p) = K(n,x)p^x(1-p)^{n-x}$, and $K(a,b) = \frac{a!}{(a-b)!\ b!}$.

• **Object/Object Similarity:** The similarity between objects $\mathcal{M}_i$ and $\mathcal{M}_j$ is simply the object/hypothesis similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, S_{ij}^τ, for all possible values of $\tau \in \mathcal{T}$. Thus, object/object similarity can be viewed as a probabilistic function.

5.2 Construction of Similarity Histogram

As discussed in the previous section, the similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$ is described using two parameters, $(N_{ij}^\tau, P_{ij}^\tau)$, which define the associated binomial distribution. For our purpose of performance prediction, two additional parameters are required: 1) the size of $\mathcal{M}_i$, $\mid \mathcal{M}_i \mid$, and 2) the *effective size* of $\mathcal{M}_j^\tau$, which is simply the number of features that can contribute votes to $\mathcal{M}_j^\tau$. Potential vote-contributing features of $\mathcal{M}_j^\tau$ are those which lie inside the clutter region R_c, i.e., $\mathcal{M}_j^\tau \cap R_c$.[1] Accordingly, similarity information is accumulated in a 4-D histogram. Figure 5 shows the algorithm used to build the similarity histogram (SH). Notice that each model object is compared with all objects in the model set, including itself. Accordingly, unless there are identical or extremely similar model objects, each diagonal entry $SH(m,m,m,m)$ will contain the number of model objects of size m. This is because the similarity between an object, $\mathcal{M}_i$, and a copy of itself (i.e., $\mathcal{M}_i^{\mathbf{0}}$, where $\mathbf{0}$ is the origin of $\mathcal{T}$) is a deterministic value, which is $\mid \mathcal{M}_i \mid$. For our purposes, we set these diagonal entries to 0.

[1] In the implementation, we have also included features of $\mathcal{M}_j^\tau$ outside R_c that are similar to features in $\mathcal{M}_i$. Obviously, these similar feature pairs are very close to the boundary of R_c.

```
Initialize similarity histogram SH
for each model object M_i ∈ DB do
  for each model object M_j ∈ DB do
    for each τ ∈ T do
      Inc SH(|M_i|, |M_j^τ ∩ R_c|, N_ij^τ, ⌊N_ij^τ P_ij^τ + 1/2⌋) by 1
    end
  end
end
Set diagonal entries in SH to 0
```

Figure 5: Similarity-computation algorithm.

6 Computation of Performance Bound

In this section, we determine the number of votes for $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, given a distorted version of $\mathcal{M}_i$, and derive a lower bound on PCR.

6.1 Vote Analysis

Given a distorted version of model object $\mathcal{M}_i$, $\mathcal{D}_i(R_u(\cdot), O, C, R_c)$ or simply $\mathcal{D}_i$, it can be easily shown that the number of votes for $\mathcal{M}_i$ is

$$V_i = \mid \mathcal{M}_i \mid - O. \tag{1}$$

On the other hand, the number of votes for $\mathcal{M}_j^\tau$ is a random variable, V_j^τ. In order to simplify the process of estimating the PMF of V_j^τ, we assume that: 1) the uncertainty regions associated with the features of each of $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$ are non-overlapping, 2) there is a one-to-one correspondence between similar features in $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, and 3) the similarity between every pair of similar features is the average object/feature similarity P_{ij}^τ. These assumptions result in a "uniform" view of the similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$. As an example, Figure 6 illustrates the uniform view of similarity corresponding to the object/hypothesis pair shown in Figure 4.

The votes for $\mathcal{M}_j^\tau$ are obtained from two sources: 1) object $\mathcal{M}_i$, due to similarity, and 2) clutter features, due to random coincidence. Accordingly, we can express the number of votes for $\mathcal{M}_j^\tau$, V_j^τ, as

$$V_j^\tau = V_s + V_c \tag{2}$$

where V_s and V_c are random variables corresponding to similarity and clutter votes for $\mathcal{M}_j^\tau$, respectively. A schematic diagram depicting these variables is shown in Figure 7. The number of similarity votes (V_s) depends on: 1) the number of ***unoccluded*** features, N_o,

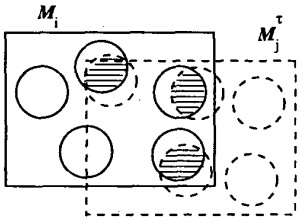


Figure 6: A conceptual view of the uniform similarity for the object/hypothesis pair shown in Figure 4. Notice that feature/feature similarity for each of the three similar feature pairs is $P_{ij}^{\tau} \approx \frac{2}{3}$.

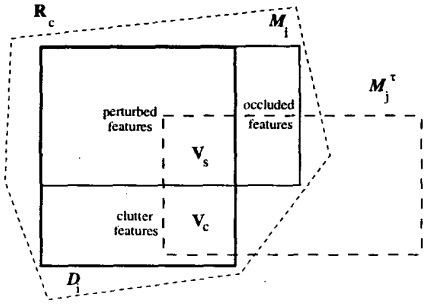


Figure 7: Components of the votes for $\mathcal{M}_j^{\tau}$ due to similarity (V_s) and clutter (V_c), given $\mathcal{D}_i$, a distorted version of object $\mathcal{M}_i$.

among the N_{ij}^{τ} features in $\mathcal{M}_i$ that are similar to features in $\mathcal{M}_j^{\tau}$, and 2) the degree of similarity between their features, represented by P_{ij}^{τ}. Accordingly, we can express the PMF of V_s as

$$P_{V_s}(v_s) = \sum_{n_o} P_{N_o}(n_o) P_{V_s}(v_s; n_o) \tag{3}$$

where $P_{V_s}(v_s; n_o) = \Pr[V_s = v_s; N_o = n_o]$. From (2) and (3), we can express the PMF of V_j^{τ} as

$$P_{V_j^{\tau}}(v_j^{\tau}) = \sum_{n_o} P_{N_o}(n_o) \sum_{v_s} P_{V_s}(v_s; n_o) P_{V_c}(v_j^{\tau} - v_s; n_o, v_s) \tag{4}$$

where $P_{V_c}(v_c; n_o, v_s) = \Pr[V_c = v_c; N_o = n_o, V_s = v_s]$.

The PMF of N_o and the conditional PMF's of V_s and V_c are determined based on the statistical data-distortion models outlined in Section 4, and the uniform model of similarity described at the beginning of this section:

• **PMF of N_o:** Assuming uniform occlusion and similarity models, we can describe N_o by the following hypergeometric distribution:

$$P_{N_o}(n_o) = H_{N_o}(N_{ij}^{\tau} - n_o; O, N_{ij}^{\tau}, \mid \mathcal{M}_i \mid - N_{ij}^{\tau})$$

where $H_X(x; n, a, b) = \frac{K(a,x)K(b,n-x)}{K(a+b,n)}$.

• **Conditional PMF of V_s:** Based on the assumptions of uniform models for uncertainty and similarity, it can be shown that the conditional PMF of V_s is represented by the following binomial distribution:

$$P_{V_s}(v_s; n_o) = B_{V_s}(v_s; n_o, P_{ij}^{\tau}).$$

• **Conditional PMF of V_c:** The effective clutter region R_c' can be split into two sub-regions, R_{V_c}' and $R_c' - R_{V_c}'$, such that a clutter feature falling within the first (second) sub-region will (will not) result in a vote for $\mathcal{M}_j^{\tau}$. Region R_{V_c}' is the union of uncertainty regions associated with features in $\mathcal{M}_j^{\tau} \cap R_c$ that are *not* already contributing similarity votes. Due to our modeling of clutter, features in $\mathcal{M}_j^{\tau} \cap R_c$ that are similar to *occluded* features in $\mathcal{M}_i$ effectively have "truncated" uncertainty regions. Under the assumption of uniform similarity, it can be shown that: 1) the area of a truncated uncertainty region is $A(R_u(\cdot))(1 - P_{ij}^{\tau})$, and 2) the number of features in $\mathcal{M}_j^{\tau} \cap R_c$ with full (truncated) uncertainty regions is $\mid \mathcal{M}_j^{\tau} \cap R_c \mid - v_s - N_{ij}^{\tau} + n_o$ $(N_{ij}^{\tau} - n_o)$. Accordingly, the conditional PMF of V_c can be approximated by the following binomial distribution:

$$P_{V_c}(v_c; n_o, v_s) \approx B_{V_c}(v_c; C, \frac{A(R_{V_c}')}{A(R_c')}) \tag{5}$$

where

$$\begin{aligned} A(R_{V_c}') &= A(R_u(\cdot))(\mid \mathcal{M}_j^{\tau} \cap R_c \mid - v_s - P_{ij}^{\tau}(N_{ij}^{\tau} - n_o)), \\ A(R_c') &= A(R_c) - A(R_u(\cdot))O. \end{aligned}$$

6.2 Bounding PCR

Let $\mathcal{N}_i$ be the set of object/pose hypotheses associated with $\mathcal{M}_i$ excluding itself. That is,

$$\mathcal{N}_i = \{\mathcal{M}_j^{\tau} : \forall \mathcal{M}_j \in \mathcal{DB}, \text{ and } \forall \tau \in \mathcal{T}\} - \{\mathcal{M}_i\}.$$

The probability of misinterpreting $\mathcal{D}_i$, as any hypothesis in $\mathcal{N}_i$, can be expressed as

$$\Pr[\mathcal{N}_i; \mathcal{D}_i] = \Pr[\exists\ \mathcal{M}_j^{\tau} \in \mathcal{N}_i \text{ such that } V_j^{\tau} \geq V_i]. \tag{6}$$

From (1) and (4), we determine the probability that votes for $\mathcal{M}_j^{\tau}$ reach or exceed those for $\mathcal{M}_i$:

$$\Pr[\mathcal{M}_j^{\tau}; \mathcal{D}_i] = \sum_{v_j^{\tau} \geq \mid \mathcal{M}_i \mid - O} P_{V_j^{\tau}}(v_j^{\tau}). \tag{7}$$

The probability of misinterpretation, $\Pr[\mathcal{N}_i; \mathcal{D}_i]$, can be bounded as follows (refer to(6)):

$$\Pr[\mathcal{N}_i; \mathcal{D}_i] < \sum_{\mathcal{M}_j^{\tau} \in \mathcal{N}_i} \Pr[\mathcal{M}_j^{\tau}; \mathcal{D}_i].$$

The above inequality can be directly used to determine the following lower bound on the probability of correctly recognizing $\mathcal{M}_i$, given $\mathcal{D}_i$:

$$\Pr[\mathcal{M}_i; \mathcal{D}_i] > 1 - \sum_{\mathcal{M}_j^\tau \in \mathcal{N}_i} \Pr[\mathcal{M}_j^\tau; \mathcal{D}_i]. \quad (8)$$

Next, we calculate a lower bound on average PCR for model-object set $\mathcal{DB}$. From the discussion in the previous section, it can be observed that V_j^τ and, in turn, $\Pr[\mathcal{M}_j^\tau; \mathcal{D}_i]$ depend on only four object-dependent parameters: the sizes of $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, and the similarity parameters $(N_{ij}^\tau, P_{ij}^\tau)$. Let

$$W(a, b, c, d) = \Pr[\mathcal{M}_j^\tau; \mathcal{D}_i],$$

such that $a = \mid \mathcal{M}_i \mid$ and $b = \mid \mathcal{M}_j^\tau \mid$, $c = N_{ij}^\tau$ and $d = \lfloor N_{ij}^\tau P_{ij}^\tau + \frac{1}{2} \rfloor$. The average PCR for model set $\mathcal{DB}$ can be bounded as follows:

$$\text{PCR}(\mathcal{DB}) > 1 - \frac{1}{\mid \mathcal{DB} \mid} \sum_a \sum_b \sum_c \sum_d SH(a, b, c, d) W(a, b, c, d).$$

7 Experimental Results

In this section, we validate our performance-prediction method in the context of a target recognition task using SAR images.

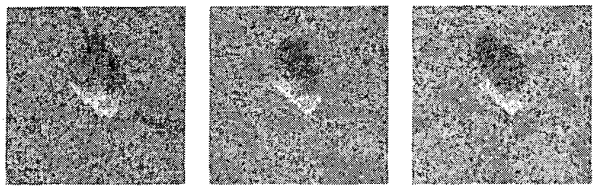
• **Model Data:** The selected model database consists of three military targets: T72, BMP2 and BTR70. Each model target is represented by a number of SAR views, which sample its signature at a variety of azimuth angles, at a specific depression angle. Targets T72, BMP2 and BTR70 are represented by 231, 233, and 233 views, respectively, at depression angle 17°. Examples of these views, which are obtained from the MSTAR public data set, are shown in Figure 8. We treat each model view as an independent object. Scattering centers, peaks in the image, are the point features used for recognition. These features are extracted by comparing the value of each pixel with its eight neighbors. We have chosen the strongest 30 scattering centers to represent both model and test views. Since we are considering a fixed number of scattering centers, the numbers of occluded and clutter features (O and C) in an image are always the same. The space of applicable transformations is 2-D translation (discrete) in the image plane [1].

• **Test Data:** Three sets of test data are selected. The first one, set A, is obtained by introducing distortion to the model views, according to the distortion process described in Section 4. The selected distortion parameters are: $R_u(\cdot)$ = four-neighbor region, R_c = convex hull of view features, $O/C = 9, 10, \ldots, 20$. For a given occlusion/clutter rate, each model view is distorted four times. Accordingly, the total number of test views, for a given O/C, is 4×697. The other two sets, B and C, are variants of the model set obtained by changing depression angle (from 17° to 15°), and configurations (e.g., different number of fuel barrels, different flash lights), respectively. It is known that these changes can significantly distort the view structure. The distortion parameters $R_u(\cdot)$ and R_c are empirically chosen to be the same as those for test set A. The occlusion/clutter rate for each test view is estimated through finding the best match with model views within a difference of ±3° azimuth angles (if no views exist within this range, then the test view is matched with the model view that is nearest in azimuth). Sets B and C are of sizes 581 and 464, respectively.

• **Results:** We compare a predicted lower bound on the PCR plot, as a function of occlusion/clutter rate, with actual plots corresponding to the three test data sets. The actual PCR plots are determined using an uncertainty-accommodating recognition system that examines all of the relevant 4-D problem space (target, azimuth, and translations along the range and cross-range directions). Accordingly, its performance is optimal. The predicted PCR plot is obtained as described in Section 6. Since the shape of R_c is object-dependent (convex hull of object features), $A(R_c)$ is substituted by the average of clutter-region areas corresponding to views of the model set. Furthermore, since the above-mentioned feature-extraction process imposes the constraint that features can not be eight-neighbors, the conditional clutter-vote PMF (5) is generalized to consider such a constraint (see the Appendix). Actual and predicted plots for test sets A, B and C are shown in Figures 9(a), 9(b) and 9(c), respectively. From the results obtained, we observe that the proposed method succeeds in predicting a tight bound on PCR performance. It can also be observed that for sets B and C, the predicted PCR bound is slightly over-optimistic close to the knee of the plot. This is due to arguably-slight differences between the assumed distortion models (including size and shape of the clutter region), and the actual ones. Accurate determination of these models for a given test data set is a subject of future research.

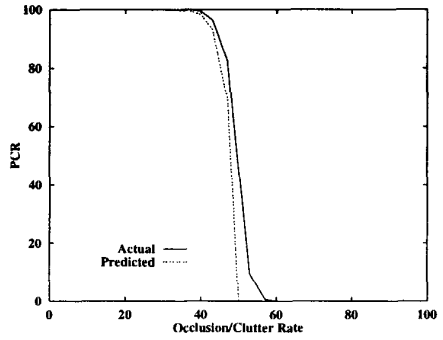
8 Conclusions

We have presented a novel method for predicting a tight lower bound on performance of vote-based object recognition. Performance is predicted by considering uncertainty, occlusion, clutter and object similarity. The method has been validated by comparing

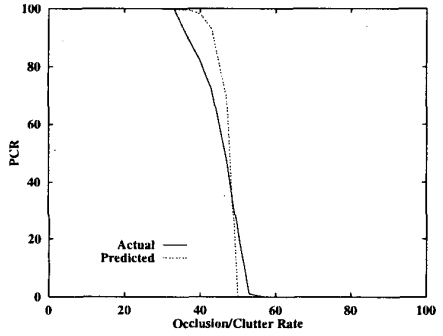


(a) T72 (b) BMP2 (c) BTR70

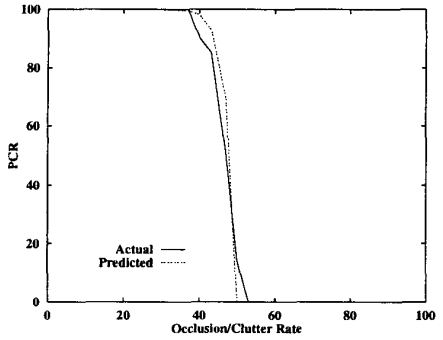
Figure 8: Examples of SAR images.



(a) controlled distortion



(b) depression angle change



(c) configuration differences

Figure 9: Actual and predicted PCR plots for test sets A, B, and C.

experimentally-determined PCR plots with predicted bounds using real SAR data.

Appendix

The feature-adjacency constraint can be defined using a *separation region*, $R_s(\cdot)$. For example, if the point features used correspond to image peaks (as in our experiments, see Section 7), then no two features can be eight neighbors. In such a case, $R_s(\cdot)$ is a 3×3 window centered at the feature's location. Considering the feature-adjacency constraint, we can approximate the conditional PMF of V_c as follows:

$$P_{V_c}(v_c; n_o, v_s) \approx K(C, v_c) \frac{L(R'_{V_c}, v_c) L(R''_c - R'_{V_c}, C - v_c)}{L(R'_c, C)}$$

where

$$\begin{aligned} L(R, n) &= \prod_{i=0}^{n-1} (A(R) - iA(R_s(\cdot))), \text{ and} \\ A(R''_c) &= A(R'_c) - A(R_s(\cdot))(\mid \mathcal{M}_i \mid - O). \end{aligned}$$

Note that if $A(R_s(\cdot)) = 0$, then the above distribution will reduce to the binomial distribution defined in (5).

References

[1] B. Bhanu and G. Jones III. Performance characterization of a model-based SAR target recognition system using invariants. In *SPIE Conf. Algorithms for Synthetic Aperture Radar Imagery IV*, volume 3070, pages 305–321, 1997.

[2] U. Grenander, M. I. Miller, and A. Srivastava. Hilbert-Schmidt lower bounds for estimators on matrix lie groups for ATR. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 20(8):790–802, 1998.

[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] W. W. Irving, R. B. Washburn, and W. E. L. Grimson. Bounding performance of peak-based target detectors. In *SPIE Conf. Algorithms for Synthetic Aperture Radar Imagery IV*, volume 3070, pages 245–257, 1997.

[5] M. Lindenbaum. Bounds on shape recognition performance. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 17(7):665–680, 1995.