Table 2: Results on various SAR databases

Model/Test Data Differences	Total no. of model objects	Total no. of test objects	% of total variance explained by first 2 factors	No. of random test queries	Avg. no. of retrievals after indexing	No. of cases with correct ID
Depr. angle differences (*Database* 1)	1621	1351	85.31	200	84 (5.18%)	185 (92.5%)
Configuration differences (*Database* 2)	694	1621	70.84	200	64 (9.22%)	170 (85%)
Articulation differences (*Database* 3)	606	239	78.29	200	96 (15.84%)	139 (69.5%)

target database, whereas, the *mode*, *wdia* and *hdia* fall distant to any other feature and hence form independent descriptors of the database. The other features look smeared in the factor space and so we do not use them in clustering. Thus, we form four independent feature-groups: the first consists of (*mean*, *std*, *max* and *median*), and the rest consist of just one feature each (*mode*, *wdia* and *hdia*, respectively). Figure 3 (right) shows the projections of both objects and the feature-groups in the factor space spanned by the first two dominant factors for the same databases. The centroids of the four feature-groups are used as the seeds for nearest-neighbor clustering of the objects. The results of using our approach are summarized in Table 2.

5 Conclusions

We have presented a database-retrieval oriented approach to model-based object recognition, for large model databases. We have presented detailed results on several real SAR target databases to demonstrate our approach. These are difficult databases since feature invariance may be small (20% to 50% [1]). The technique is general and has other interactive object recognition applications involving large model databases. It also helps interactive discovery and acquisition of new models to update database.

Currently, mutual information based ranking takes more than 99% of the time during recognition (retrieval) phase. The performance of our approach could be further improved by: (a) efficient computation of mutual information between a query and a candidate model object using a coarse-to-fine strategy, (b) overlapped boundaries in nearest-neighbor clustering method, and (c) efficient methods to adapt factors to dynamically changing databases. We are currently investigating along these directions.

References

[1] B. Bhanu, G. Jones, and J. Ahn. Recognizing articulated objects and object articulation in SAR images. In *SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery V*, volume 3370, April, 1998.

[2] T. Pun and D. Squire. Statistical structuring of pictorial databases for content-based image retrieval systems. *Pattern Recognition Letters*, 17:1299–1310, 1996.

[3] J.-P. Benzécri. *Correspondence Analysis Handbook.* Marcel Dekker, New York, 1992.

[4] M. Jambu. *Exploratory & Multivariate Data Analysis.* Academic Press, New York, 1991.

[5] J. F. Barros, W. Martin, P. Kelley, and M. Cannon. Using the triangle inequality to reduce the number of comparisons required for similarity-based retrieval. In *Proc. of IS&T/SPIE – Storage and Retrieval for Still Image and Video Databases*, volume 4, pages 392–403, January, 1996.

[6] A. Berman and L. G. Shapiro. Selecting good keys for triangle-inequality-based pruning algorithms. In *Proc. of IEEE Intl. Workshop on Content-based Access of Image and Video Databases (CAIVD'98)*, pages 12–19, January, 1998.

[7] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. In *Proc. of* 5^{th} *Intl. Conf. on Computer Vision (ICCV'95)*, pages 16–23, 1995.

[8] Y. L. Moon, B. Rajagopalan, and U. Lall. Estimation of mutual information using kernel density estimators. *Physical Review E: Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 52(3):2318–2321, 1995.

Bounding Fundamental Performance of Feature-Based 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

Performance prediction is a crucial step for transforming the field of object recognition from an art to a science. In this paper, we address this problem in the context of a vote-based approach for object recognition using 2-D point features. A method is presented for predicting tight lower and upper bounds on fundamental performance of the selected recognition approach. Performance bounds are predicted by considering data-distortion factors, which are uncertainty, occlusion and clutter, in addition to model structural similarity. Given a statistical model of data uncertainty, the structural similarity between every pair of model objects is computed as a function of the relative transformation between them. Model-similarity information is then used along with statistical data-distortion models to predict bounds on the probability of correct recognition. Validity of the method is experimentally demonstrated using MSTAR public SAR data.

1 Introduction

The problem of object recognition is concerned with identifying and localizing model objects from scene data. It involves searching for a consistent correspondence between scene features, and those of a model object. Performance of such a process depends on a large number of factors, which are associated with either scene data (e.g., sensor noise, missing and spurious features), or model objects (e.g., size of model database, similarity of model objects, articulation of model parts). Predicting the performance of object recognition as a function of these factors is a challenging task.

In this paper, we address the problem of performance prediction in the context of an approach for object recognition using 2-D point features. Such an approach uses a vote-based matching criterion, which ranks object/pose hypotheses based on the number of model features (votes) that are consistent with scene features. We predict recognition performance of this approach, by considering the following factors: *1) Scene-Data Factors:* uncertainty (due to sensor noise and imperfections of the feature-extraction process), occlusion (missing features), and clutter (extraneous features), *2) Model-Object Factors:* similarity (degree of structural overlap between pairs of model objects), and number of model objects (this factor is implicitly considered in our handling of object similarity).

Problem Definition: Our performance-prediction problem can be defined as follows. We are given: *a)* a set of model objects, $\mathcal{M} = \{\mathcal{M}_i\}$, where each object $\mathcal{M}_i$ is represented by a set of 2-D point features, $\{F_{ik}\}$, that are discretized at some resolution, *b)* statistical models for data distortion (uncertainty, occlusion, and clutter), and *c)* a class of applicable transformations, $\mathcal{T}$ (e.g., translation, rigid, affine). Our objective is to predict tight lower and upper bounds on the probability-of-correct-recognition (PCR) plot, as a function of occlusion and clutter rates (assuming a fixed uncertainty model). 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 vote-based recognition algorithm used. Thus, it sets an upper bound on performance that is achievable by any recognition algorithm that uses the same matching criterion.

*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.

Relevant Research and Our Contributions: Performance analysis of object recognition has not received much attention in the literature. Some research efforts analyze the problem of discriminating objects from random clutter, or clutter-to-object misclassification (e.g., [2, 3, 6]). The problem of object-to-object misclassification has been analyzed by considering either data uncertainty and model similarity (with partial handling of occlusion) [4], or data uncertainty, model similarity and random clutter [5]. The *main contribution* of this paper is the development of a method for predicting tight bounds on fundamental performance of object-to-object misclassification, using a vote-based criterion, by simultaneously considering data uncertainty, occlusion, clutter and model similarity. Previous approaches consider only a subset of these factors. We present experimental results to validate the proposed method using MSTAR public SAR data.

2 Overview of Our Method

In this section, we first present a qualitative description of the effects of data distortion and model similarity on PCR, and then we outline the proposed method.

2.1 Effects of Data Distortion and Model Similarity on Performance

Figure 1 shows a schematic diagram of $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, which is an *instance* of object $\mathcal{M}_j$ at pose $\tau \in \mathcal{T}$ relative to $\mathcal{M}_i$. In this figure, the intersection region denotes "similar" features in $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, while the other two regions denote "dissimilar" features in both objects. The size of the intersection region, denoted by $S(\mathcal{M}_i, \mathcal{M}_j^\tau)$, represents the degree of similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$. Let us assume that we are given a scene of object $\mathcal{M}_i$. Initially, without any kind of distortion, objects $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$ get $\mid \mathcal{M}_i \mid$, and $S(\mathcal{M}_i, \mathcal{M}_j^\tau)$ votes, respectively. The vote difference $\Delta V(\mathcal{M}_i, \mathcal{M}_j^\tau) = \mid \mathcal{M}_i \mid - S(\mathcal{M}_i, \mathcal{M}_j^\tau)$ represents the "cushion" which protects $\mathcal{M}_i$ from being misclassified as $\mathcal{M}_j^\tau$ in the presence of data distortion. Accordingly, the likelihood of misclassifying $\mathcal{M}_i$ as $\mathcal{M}_j^\tau$ is directly proportional to the degree of similarity between them. This supports the intuitive observation that distinct objects are less likely to be misclassified than those which are similar.

The qualitative effects of data-distortion factors on $\Delta V(\mathcal{M}_i, \mathcal{M}_j^\tau)$ can be outlined as follows. *1) Uncertainty:* intuitively, uncertainty "blurs" the structure (features) of object $\mathcal{M}_i$, thus making it look more similar to $\mathcal{M}_j^\tau$. This is manifested in increasing $S(\mathcal{M}_i, \mathcal{M}_j^\tau)$, and, accordingly, decreasing $\Delta V(\mathcal{M}_i, \mathcal{M}_j^\tau)$. *2) Occlusion:* occlusion of features which distinguish $\mathcal{M}_i$ from $\mathcal{M}_j^\tau$ (region $\mathcal{M}_{i/j}$ in Figure 1) decreases $\Delta V(\mathcal{M}_i, \mathcal{M}_j^\tau)$. On the other hand, occlusion of features in $\mathcal{M}_i$ which are similar to those in $\mathcal{M}_j^\tau$ (region $\mathcal{M}_{ij}$ in Figure 1) tends to keep $\Delta V(\mathcal{M}_i, \mathcal{M}_j^\tau)$ fixed, since both objects would lose the same number of votes. *3) Clutter:* clutter decreases $\Delta V(\mathcal{M}_i, \mathcal{M}_j^\tau)$ by increasing the number of votes for $\mathcal{M}_j^\tau$, while keeping that for $\mathcal{M}_i$ fixed. We note that the ambiguous case of a clutter feature that happens to coincide with an occluded feature in $\mathcal{M}_i$ is interpreted as no-occlusion/no-clutter in this paper. The effect of clutter is magnified in the presence of high data uncertainty, since this would increase the chances of a clutter feature to be considered as "consistent" with an unmatched feature in $\mathcal{M}_j^\tau$.

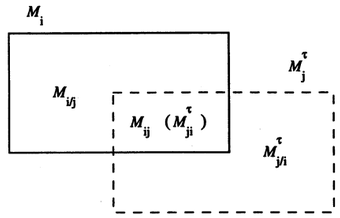


Figure 1: A schematic diagram of $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$ showing similar and dissimilar features. The intersection region denotes the similar features in $\mathcal{M}_i$ ($\mathcal{M}_{ij}$) or $\mathcal{M}_j^\tau$ ($\mathcal{M}_{ji}^\tau$), while the other two regions denote dissimilar features in $\mathcal{M}_i$ ($\mathcal{M}_{i/j}$), and $\mathcal{M}_j^\tau$ ($\mathcal{M}_{j/i}^\tau$).

The above qualitative analysis is quantified in our proposed method, outlined in the next section, in order to determine tight bounds on PCR performance.

2.2 Outline of the Method

A block diagram of the proposed method is shown in Figure 2. Its major components are described below.

Data-Distortion Models: We model data distortion as follows: *1) Uncertainty:* since we are dealing with discretized features, the actual location of a scene feature is described by a probability mass function (PMF). In this paper, we assume that this distribution is uniform for the sake of simplicity. Extension of the proposed method to handle arbitrary uncertainty PMF's is possible. Further, we assume that the uncertainty PMF's associated with scene features are independent. *2) Occlusion:* we assume that each subset of features is equally likely to be occluded as any other subset that is of the same size.

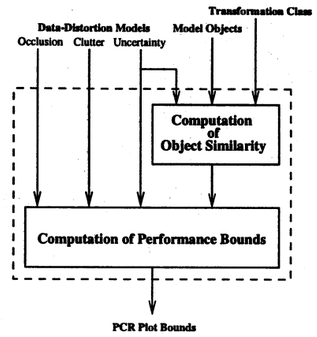


Figure 2: Our performance-prediction method.

3) Clutter: clutter features are assumed to be uniformly distributed within some area surrounding the object.

Computation of Object Similarity: This stage computes similarity between all pairs of model objects. The structural similarity between $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$ can be defined as the number of votes that $\mathcal{M}_j^\tau$ would get, given an "uncertain" version of $\mathcal{M}_i$. Thus, in addition to the degree of structural overlap, our notion of object similarity is dependent upon the amount of uncertainty in the data as well. Similarity between every pair of model objects is computed for all possible relative poses, as determined by the applicable transformation class $\mathcal{T}$, and then accumulated in *similarity histograms.*

Computation of Performance Bounds: This stage utilizes object similarity information and data-distortion models, within a statistical framework, in order to calculate PCR bounds of an object, for given occlusion and clutter rates. Averaging the results for all model objects, and repeating the process for a variety of occlusion and clutter rates, we can predict bounds on the PCR plot.

3 Computing Object Similarity

In this section, we formally define a similarity measure between model objects, and outline how similarity histograms are constructed.

3.1 Definition of Object Similarity

We introduce a sequence of definitions that lead to formally defining a similarity measure between a pair of objects.

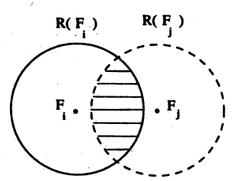


Figure 3: An illustration of feature/feature similarity.

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. Since we are assuming uniform data uncertainty, the uncertainty PMF associated with feature F_j can be represented by an *uncertainty region*, $R(F_j)$. In such a case, F_i is considered consistent with F_j, if $F_i \in R(F_j)$.

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. It is easy to show that $S_{ff}(F_i, F_j)$ is proportional to the area of the region corresponding to the intersection of $R(F_i)$ and $R(F_j)$ (see Figure 3). In particular,

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

where $A(R)$ is the area of region R.

Feature/Object 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 $\mathcal{M}_i$ is consistent with F_j. Formally,

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

Object/Object Similarity: The similarity between objects $\mathcal{M}_i$ and $\mathcal{M}_j^\tau$, $S_{oo}(\mathcal{M}_i, \mathcal{M}_j^\tau)$ or simply $S(\mathcal{M}_i, \mathcal{M}_j^\tau)$, can be defined as the number of features in $\mathcal{M}_j^\tau$ that have a consistent uncertain measurement of a feature in $\mathcal{M}_i$. It can be easily seen that $S(\mathcal{M}_i, \mathcal{M}_j^\tau)$ is a random variable (see Figure 4). We approximate the distribution of such a variable by a binomial distribution that is defined as follows:

$$\begin{aligned} P_S(s; \mathcal{M}_i, \mathcal{M}_j^\tau) &= \Pr[S(\mathcal{M}_i, \mathcal{M}_j^\tau) = s] \\ &= B_S(s; N(\mathcal{M}_i, \mathcal{M}_j^\tau), P(\mathcal{M}_i, \mathcal{M}_j^\tau)) \end{aligned}$$

where

$$\begin{aligned} N(\mathcal{M}_i, \mathcal{M}_j^\tau) &= \mid \{F_{jk}^\tau : S_{of}(\mathcal{M}_i, F_{jk}^\tau) > 0\} \mid, \\ P(\mathcal{M}_i, \mathcal{M}_j^\tau) &= \frac{\sum_k S_{of}(\mathcal{M}_i, F_{jk}^\tau)}{N(\mathcal{M}_i, \mathcal{M}_j^\tau)}, \\ B_X(x; n, p) &= C(n, x) p^x (1-p)^{n-x}, \end{aligned}$$

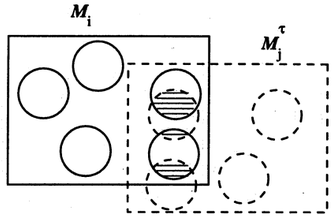


Figure 4: An illustration of object/object similarity. In the example shown, $S(\mathcal{M}_i, \mathcal{M}_j^\tau) \in [0, 2]$.

and $C(a,b) = \frac{a!}{(a-b)!\ b!}$. Note that, in the special no-uncertainty case, $S(\mathcal{M}_i, \mathcal{M}_j^\tau)$ is a constant number, which is equal to $N(\mathcal{M}_i, \mathcal{M}_j^\tau)P(\mathcal{M}_i, \mathcal{M}_j^\tau)$.

3.2 Constructing Similarity Histograms

Figure 5 shows the algorithm used to compute similarity information. In that algorithm, each model object is compared with all model objects in the database, including itself, to compute the similarity information. The output of this process is accumulated in a pair of similarity histograms. One of them, ***all-similarity*** histogram, accumulates similarity information for all object pairs. The other one, ***peak-similarity*** histogram, accumulates similarity information for only object/instance pairs $(\mathcal{M}_i, \mathcal{M}_j^\tau)$ corresponding to ***peaks*** of the expected-similarity function, $N(\mathcal{M}_i, \mathcal{M}_j^\tau)P(\mathcal{M}_i, \mathcal{M}_j^\tau)$, in the space of applicable transformations $\mathcal{T}$. As will be shown in Section 5.2, we use all- and peak-similarity histograms to predict lower and upper bounds on the PCR plot, respectively.

A similarity histogram is of three dimensions, corresponding to object size, instance size, and degree of similarity. The value of an entry, (m_i, m_j, s_{ij}), in all-(peak-) similarity histogram represents the ***expected*** total number of all (peak) instances of size m_j that have ***exactly*** s_{ij} consistent features with objects of size m_i. Note that, unless there are identical or very similar model objects, entry (m, m, m) in either all- or peak-similarity histogram will contain the number of model objects with size m. This is simply because computing the similarity of an object, $\mathcal{M}_i$, with a copy of itself (i.e., computing $S(\mathcal{M}_i, \mathcal{M}_i)$) will result in a deterministic similarity distribution $B_S(s; \mid \mathcal{M}_i \mid, 1)$, which will increment the value of entry $(\mid \mathcal{M}_i \mid, \mid \mathcal{M}_i \mid, \mid \mathcal{M}_i \mid)$ by one. Diagonal entries are set to 0, for the purpose of performance prediction.

```
Initialize all-/peak-Similarity histograms (AS/PS).
for each model object M_i do
  for each model object M_j do
    for each τ ∈ T do
      Calculate N(M_i, M_j^τ) and P(M_i, M_j^τ).
      for each possible similarity value s do
        Set p_s = B_S(s; N(M_i, M_j^τ), P(M_i, M_j^τ)).
        Increment AS(| M_i |, | M_j^τ |, s) by p_s.
        if M_j^τ is a peak instance
          Increment PS(| M_i |, | M_j^τ |, s) by p_s.
      end
    end
  end
end
```

Figure 5: Similarity-computation algorithm.

4 Vote Analysis

In this section, we determine the number of votes for $\mathcal{M}_i$, and $\mathcal{M}_j^\tau$, for given levels of data-distortion factors. The results obtained in this section form the basis for determining PCR bounds, which are derived in the next section.

Let $\mathcal{D}_i$ be a distorted image of $\mathcal{M}_i$ that is obtained by

1. Occluding a subset of features, $\mathcal{O} \subset \mathcal{M}_i$,
2. Perturbing each unoccluded feature in $\mathcal{M}_i$, uniformly, within uncertainty region $R(\cdot)$, such that no perturbed feature is consistent with any feature in $\mathcal{O}$, and
3. Adding a set of clutter features, $\mathcal{C}$, none of them are consistent with any feature in $\mathcal{O}$.

The number of votes for $\mathcal{M}_i$, given scene $\mathcal{D}_i$, is

$$V(\mathcal{M}_i; \mathcal{D}_i) \quad = \quad \mid \mathcal{M}_i \mid - \mid \mathcal{O} \mid . \qquad (1)$$

On the other hand, the number of votes for $\mathcal{M}_j^\tau$, $V(\mathcal{M}_j^\tau; \mathcal{D}_i)$, is a random variable, which is described as follows:

$$\Pr[V(\mathcal{M}_j^\tau; \mathcal{D}_i) = v] = \sum_s P_S(s; \mathcal{M}_i, \mathcal{M}_j^\tau) \Pr[V(\mathcal{M}_j^\tau; \mathcal{D}_i, s) = v] \quad (2)$$

where $V(\mathcal{M}_j^\tau; \mathcal{D}_i, s)$ is the number of votes for $\mathcal{M}_j^\tau$ given scene $\mathcal{D}_i$, and assuming that $S(\mathcal{M}_i, \mathcal{M}_j^\tau) = s$. Votes for $\mathcal{M}_j^\tau$ are obtained from two sources: object $\mathcal{M}_i$ (due to similarity), and clutter features $\mathcal{C}$ (due to random coincidence). Accordingly, we can express

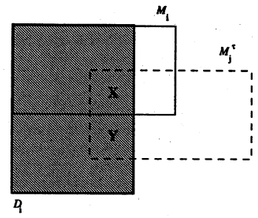


Figure 6: Components of the votes for $\mathcal{M}_j^\tau$ due to similarity (X) and clutter (Y), given a distorted scene, $\mathcal{D}_i$, of object $\mathcal{M}_i$.

$V(\mathcal{M}_j^\tau; \mathcal{D}_i, s)$ and its associated PMF as

$$V(\mathcal{M}_j^\tau; \mathcal{D}_i, s) = X + Y, \text{ and}$$
$$\Pr[V(\mathcal{M}_j^\tau; \mathcal{D}_i, s) = v] = \sum_x \Pr[X = x] \Pr[Y = v - x] \quad (3)$$

where X and Y are random variables that correspond to similarity and clutter votes for $\mathcal{M}_j^\tau$, respectively (see Figure 6).

The PMF's of X and Y are determined based on the statistical data-distortion models mentioned in Section 2.2. Since we are assuming a uniform occlusion model, it can be shown that X is described by a hypergeometric distribution,

$$\Pr[X = x] = H_X(x; s, \mid \mathcal{M}_i \mid - \mid \mathcal{O} \mid, \mid \mathcal{O} \mid) \qquad (4)$$

where $H_U(u; n, a, b) = \frac{C(a,u)C(b,n-u)}{C(a+b,n)}$. Next, let us determine the PMF of Y. It can be shown that, given m model features and n clutter features, the probability that u model features are consistent with clutter features is bounded by $[G_U(u; m, n, 0), G_U(u; m, n, 1)]$ where

$$G_U(u; m, n, b) = C(m, u)P(n, u)p^u\,(1 - (m - bu)p)^{n-u},$$

$P(a,b) = \frac{a!}{(a-b)!}$, $p = \frac{A(R(\cdot))}{A(I)}$, and I is the clutter region. Mapping this PMF to our case, we can bound the PMF of Y:

$$\Pr[Y = y; b] = G_Y(y; \mid \mathcal{M}_j^\tau \mid - s, \mid \mathcal{C} \mid, b) \qquad (5)$$

where $b = 0$ or 1 denotes lower or upper bounds, respectively (henceforth, this notation will be used to denote lower and upper bounds). From (3), (4) and (5), we can bound the PMF of $V(\mathcal{M}_j^\tau; \mathcal{D}_i, s)$:

$$\begin{aligned} P_V(v; \mathcal{M}_j^\tau, \mathcal{D}_i, s, b) &= \Pr[V(\mathcal{M}_j^\tau; \mathcal{D}_i, s) = v; b] \\ &= \sum_x (H_X(x; s, \mid \mathcal{M}_i \mid - \mid \mathcal{O} \mid, \mid \mathcal{O} \mid) \\ &\quad G_Y(v - x; \mid \mathcal{M}_j^\tau \mid - s, \mid \mathcal{C} \mid, b)). \end{aligned}$$

From (2), we can bound the PMF of $V(\mathcal{M}_j^\tau; \mathcal{D}_i)$:

$$\begin{aligned} P_V(v; \mathcal{M}_j^\tau, \mathcal{D}_i, b) &= \Pr[V(\mathcal{M}_j^\tau; \mathcal{D}_i) = v; b] \\ &= \sum_s (P_S(s; \mathcal{M}_i, \mathcal{M}_j^\tau) \\ &\quad P_V(v; \mathcal{M}_j^\tau, \mathcal{D}_i, s, b)). \end{aligned}$$

5 Computing Performance Bounds

In this section, we derive bounds on the PCR of object $\mathcal{M}_i$, given distorted scene $\mathcal{D}_i$, and determine average PCR bounds.

5.1 Performance Bounds for an Object

Let $\mathcal{N}_i$ be the set of object instances associated with $\mathcal{M}_i$. That is,

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

The PCR of $\mathcal{M}_i$ can be expressed as

$$\Pr[\mathcal{M}_i; \mathcal{D}_i] = \Pr[\forall \mathcal{M}_j^\tau \in \mathcal{N}_i : V(\mathcal{M}_j^\tau; \mathcal{D}_i) < (\mid \mathcal{M}_i \mid - \mid \mathcal{O} \mid)].$$

Attempting to estimate $\Pr[\mathcal{M}_i; \mathcal{D}_i]$ appears to be an extremely difficult task, due to the correlation of the random variables $V(\mathcal{M}_j^\tau; \mathcal{D}_i)$, especially between those corresponding to adjacent instances. This is the reason that we are resorting to determining lower and upper bounds on $\Pr[\mathcal{M}_i; \mathcal{D}_i]$, $[\Pr[\mathcal{M}_i; \mathcal{D}_i, 0], \Pr[\mathcal{M}_i; \mathcal{D}_i, 1]]$.

We start by determining lower bound $\Pr[\mathcal{M}_i; \mathcal{D}_i, 0]$. We can bound the probability that $\mathcal{M}_j^\tau$ gets same or larger number of votes than $\mathcal{M}_i$ as

$$Q(\mathcal{M}_j^\tau, \mathcal{D}_i, b) \quad = \quad \sum_{v \geq \mid \mathcal{M}_i \mid - \mid \mathcal{O} \mid} P_V(v; \mathcal{M}_j^\tau, \mathcal{D}_i, b).$$

It can be shown that the sum of $Q(\mathcal{M}_j^\tau, \mathcal{D}_i, 1)$, for all $\mathcal{M}_j^\tau \in \mathcal{N}_i$, forms an upper bound on the probability of misclassifying $\mathcal{D}_i$ as any instance in $\mathcal{N}_i$. Accordingly, we obtain the following lower bound on $\Pr[\mathcal{M}_i; \mathcal{D}_i]$:

$$\Pr[\mathcal{M}_i; \mathcal{D}_i, 0] = 1 - \sum_{\mathcal{M}_j^\tau \in \mathcal{N}_i} Q(\mathcal{M}_j^\tau, \mathcal{D}_i, 1). \qquad (6)$$

Next, we estimate upper bound $\Pr[\mathcal{M}_i; \mathcal{D}_i, 1]$. Our approach is to consider a subset of the instances of $\mathcal{N}_i$, $\mathcal{P}_i \subset \mathcal{N}_i$, which correspond to ***peaks*** of the expected-similarity function (refer to Section 3.2), and assume that the votes for peak instances are independent. The rationale behind this approach can be stated as follows: *1)* An object is more likely to be misclassified as a peak instance, $\mathcal{M}_j^\tau$, than as a neighboring off-peak instance, $\mathcal{M}_j^{\tau+\delta}$, since $\mathcal{M}_j^\tau$ is expected to have more similarity votes than $\mathcal{M}_j^{\tau+\delta}$. *2)* Peaks are generally not too close to each other and so the vote-independence assumption among peak instances, which belong to the same object, is reasonable. *3)* The vote-independence assumption

among peak instances that belong to different objects is arguably reasonable in most practical applications. Under the vote-independence assumption, the consequence of considering only a subset of all possible instances is the estimation of the following upper bound on $\Pr[\mathcal{M}_i; \mathcal{D}_i]$:

$$\Pr[\mathcal{M}_i; \mathcal{D}_i, 1] = \prod_{\mathcal{M}_j^\tau \in \mathcal{P}_i} (1 - Q(\mathcal{M}_j^\tau, \mathcal{D}_i, 0)). \quad (7)$$

5.2 Average Performance Bounds

In order to compute average PCR bounds, it is possible to evaluate equations (6) and (7), for every object, and then taking the average. However, this approach involves unnecessary waste of space, since it would require storing the similarity information associated with each object separately (a pair of 2-D similarity histograms per object). It also involves unnecessary waste of computational time, since PCR bounds would have to be computed for each object separately. We can overcome these problems as follows. From Section 4, it can be observed that the probability of misclassifying one object, $\mathcal{M}_i$, as an instance, $\mathcal{M}_j^\tau$, depends on the sizes of these objects as well as the similarity between them. Let

$$W(|\mathcal{M}_i|, |\mathcal{M}_j^\tau|, s, b) = \sum_{v \geq |\mathcal{M}_i| - |\mathcal{O}|} P_V(v; \mathcal{M}_j^\tau, \mathcal{D}_i, s, b).$$

It can be shown that the average PCR bounds, $[PCR(\mathcal{M}; 0), PCR(\mathcal{M}; 1)]$, can be expressed as

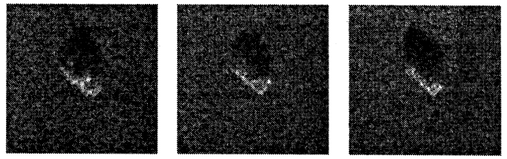
$$PCR(\mathcal{M}; 0) = 1 - \frac{1}{|\mathcal{M}|} \sum_{m_i} \sum_{m_j} \sum_{s} W(m_i, m_j, s, 1) AS(m_i, m_j, s)$$

$$PCR(\mathcal{M}; 1) \approx \sum_{m_i} \frac{N(m_i)}{|\mathcal{M}|} \Big(\prod_{m_j} \prod_{s} (1 - W(m_i, m_j, s, 0))^{\overline{PS}(m_i, m_j, s)}\Big) \Big(1 - \frac{1}{N(m_i)} \sum_{m_j} \sum_{s} \Delta PS(m_i, m_j, s) W(m_i, m_j, s, 0)\Big)$$

where AS and PS are all- and peak-similarity histograms (see the algorithm in Fig. 5), $N(m_i)$ is the number of objects of size m_i in model database $\mathcal{M}$, $\overline{PS}(m_i, m_j, s) = \frac{PS(m_i, m_j, s)}{N(m_i)}$, and $\Delta PS(m_i, m_j, s) = PS(m_i, m_j, s) - \overline{PS}(m_i, m_j, s)$.

6 Experimental Results

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



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

Figure 7: Examples of SAR images from MSTAR public data.

The selected model database consists of three military targets: T72, BMP2 and BTR. 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 BTR are represented by 231, 233, and 233 views, respectively, at depression angle 17°. Examples of these views, which are obtained from MSTAR public data, are shown in Figure 7. For our purposes, we treat each model view as an independent object. Scattering centers, peaks in the image, are the point features used for recognition. They 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 scene data. Since we are considering a fixed number of scattering centers, the occlusion and clutter rates in an image are always the same. The space of applicable transformations is 2-D translation in the image plane [1].

Table 1: Description of test data sets A and B.

	Set A	Set B
Uncertainty	none	4-neighbor region
Test Views	4+/model view	1+/model view
Occlusion/ Clutter Rate	70%, 73.3%, ..., 86.7%, 90%	56.7%, 60% ..., 73.3%, 76.7%

The test data are obtained by introducing distortion to the model views, according to the distortion models described in Section 2.2, and the restrictions outlined at the beginning of Section 4 (when describing distorted scene $\mathcal{D}_i$). The selected clutter region is a rectangular box that is obtained by dilating the bounding box of the given target view, such that the area of the clutter region is nine times that of the target bounding box. Two test data sets, A and B, are generated, assuming no uncertainty, and four-neighbor uncertainty region, respectively. These sets are described in Table 1.

Figures 8 and 9 show experimentally-determined PCR plots, along with predicted lower and upper bounds, for test sets A and B, respectively. The actual PCR plots are determined using an uncertainty-accommodating recognition system, which is based

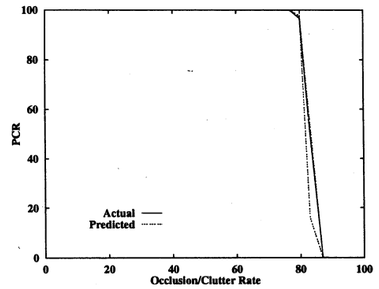


Figure 8: Actual PCR plot and predicted lower and upper PCR bounds for test set A (notice that the predicted upper bound for test set A is almost coinciding with the actual plot).

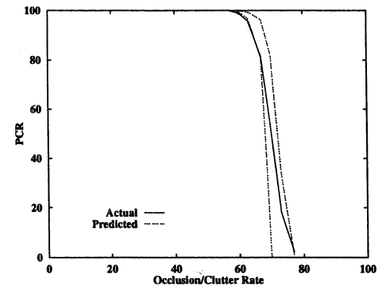


Figure 9: Actual PCR plot and predicted lower and upper PCR bounds for test set B.

on geometric hashing. This system examines almost all of the 4-D problem space (target, azimuth, and translations along the two principal axes), and so its performance is almost optimal. From the results obtained, we observe the following. *1)* The proposed method succeeds in predicting tight lower and upper bounds on PCR performance. The *maximum* horizontal distance between predicted PCR-plot bounds is 3% and 7% for sets A and B, respectively. *2)* The predicted lower bound determines the "breakpoint" in performance very accurately. *3)* The performance in this set of experiments tends to be optimistic, with perfect recognition up to relatively high occlusion/clutter rates. The main reason for such an optimism is our choice of a relatively large clutter area. *4)* In Fig. 9, the predicted lower bound slightly overestimates PCR (by about 1%) at the knee of the plot. This minor discrepancy and its statistical significance is currently under investigation. It appears to be due to approximating the distribution of the similarity function, $S(\cdot, \cdot)$, by a binomial PMF (refer to Section 3.1). Use of a more complicated function which approximates the distribution of $S(\cdot, \cdot)$ more closely would be expected to eliminate this slight overestimation.

The experimental results presented in this section clearly demonstrate the validity of our proposed method.

7 Conclusions

A novel method has been presented for predicting tight lower and upper bounds on the performance of a vote-based approach for object recognition. Performance is predicted by considering data uncertainty, occlusion, clutter and model similarity. Validity of the method has been demonstrated by comparing experimentally-determined PCR plots with predicted bounds. Future work involves developing appropriate data-distortion models for predicting performance of SAR target recognition, under depression-angle changes, configuration differences, and articulation. We also plan to extend this work to predict other performance metrics such as the confusion matrix, and the ROC curve.

References

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

[2] 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.

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

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

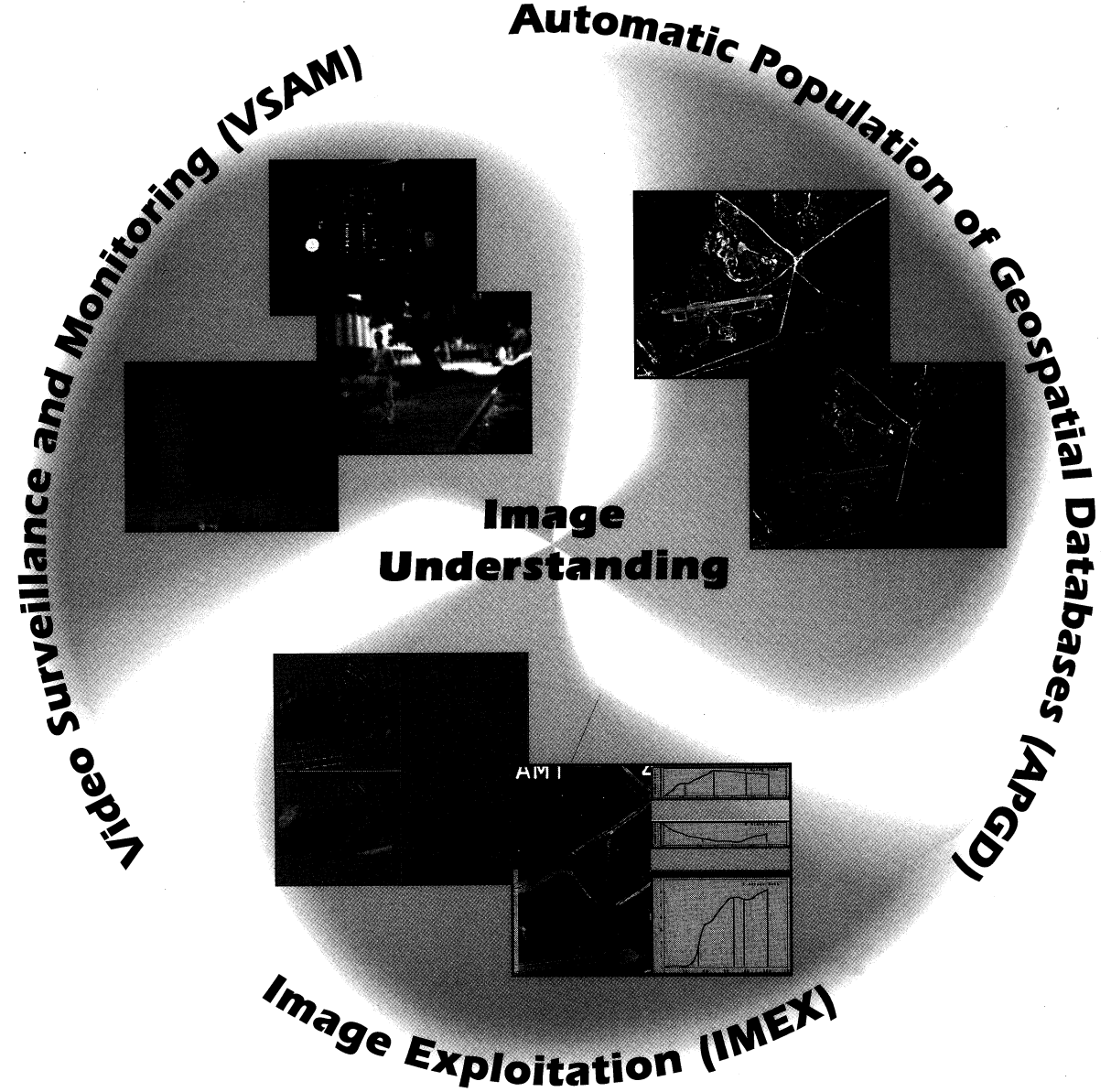
[5] M. Lindenbaum. An integrated model for evaluating the amount of data required for reliable recognition. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 19(11):1251–1264, 1997.

[6] K. B. Sarachik. The effect of Gaussian error in object recognition. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 19(4):289–301, 1997.

Proceedings Volume II

1998 Image Understanding Workshop

Hyatt Regency Monterey
November 20–23
Monterey, California



Hosted by SRI International

Edited by
George E. Lukes

Sponsored by
Defense Advanced Research Projects Agency
Information Systems Office