# An Efficient Recognition Technique for Mine-like Objects using Nearest-Neighbor Classification

Sudha Reese, Naval Undersea Warfare Center, 176 Howell Street, Newport, RI
Gita Sukthankar, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA
Rahul Sukthankar, Intel Research Pittsburgh and Carnegie Mellon University, Pittsburgh, PA

*Abstract*- Broadband active sonars (15-150 kHz) capture morphological characteristics of underwater objects and are used by dolphins to recognize targets in cluttered environments, motivating their use in underwater mine countermeasure applications. However, the data from broadband sonars is very high-dimensional (typically 1400), requiring classification algorithms that can operate in these spaces with limited training data. Standard statistical approaches such as probability density estimation are often ill-suited to this task. This paper presents a new algorithm for mine-like object recognition that has shown promising results on in-water tests. The technique employs a nearest-neighbor classifier in conjunction with a non-metric similarity function and synthetic augmentation of the training data. We present experimental results comparing our method to a standard algorithm (LDA/PCA) indicating that our approach addresses some deficiencies in existing techniques for the mine-like object recognition problem.

# 1 Introduction

Mine warfare, including the detection and classification of undersea mines, has become extremely important to the U.S. Navy. Sophisticated sea mines can be deployed at a relatively insignificant cost to cause huge problems for a battle group because of the difficulties associated with their detection and classification. Moreover, in certain scenarios, mine countermeasure operations must be performed rapidly to allow naval platforms to reach their destinations in a timely manner. Although the task of finding mine-like objects in clutter has received recent attention [1,2], little has been published on the problem of discriminating between mine-like objects (MLO) and non-mine bottom objects of similar size and shape (NOMBO). Our approach to the problem is motivated by the observation that echo highlights from broadband active sonars are an important perceptual cue that enable dolphins to recognize targets in challenging underwater environments [7]. Since echo returns from broadband sonar capture morphological information (including size, shape and scattering characteristics), it should be possible to discriminate between MLO and NOMBO by analyzing the contact echo response. Successful approaches must overcome several challenges. First, broadband sonar signals are very high-dimensional, and naïve pattern classification approaches tend to fail. Second, since acquiring this data is time-consuming and expensive, the classifier does not have access to large training sets. Finally, the appearance of MLO and NOMBO varies with aspect and range, and cannot easily be expressed using view-invariant models.

This paper describes an application of pattern recognition techniques to the mine recognition problem with propitious results using realistic in-water data. We employ a variant of memory-based learning that has shown promising results in the face recognition domain [4]. The salient features of this algorithm are: (1) a non-parametric representation of the data (nearest-neighbor classifier); (2) the use of a non-metric similarity function (instead of the traditional Euclidean distance metric); (3) the augmentation of the training data using synthetically-generated exemplars.

The remainder of the paper is organized as follows. Section 2 describes the mine-like object recognition problem. Section 3 gives an overview of a baseline classification algorithm (LDA/PCA). Section 4 presents our algorithm and results from experiments that evaluate the technique on in-water data, and explores generalization. Section 5 discusses these results and provides some hypotheses regarding the poor accuracy of PCA/LDA. Finally, Section 6 concludes the paper with a brief discussion of future research.

# 2 Statistical Pattern Recognition for Mine Recognition

Statistical pattern recognition schemes employ a training set of representative data to learn characteristics of data patterns that are envisioned in the real world [3]. Some algorithms attempt to estimate a probability density function for each of the classes (i.e. MLO and NOMBO), and classify an observed echo by observing how well it fits each of the learned class models. Other algorithms build discriminative functions that highlight the differences between MLO and NOMBO classes and classify the observed echo on the basis of this function. When the probability density function for each class is assumed to be of a known form (e.g., Gaussian) with unknown parameters, the classifier is said to employ a parametric model. Conversely, when the classes are represented by a sampling of the data, the classifier is termed to be non-parametric. Most statistical pattern recognition algorithms are challenged by high-dimensional data. For parametric models, this is because the number of parameters grows quickly with dimension; for non-parametric models, the problem is that the sampled points populate the observation space very sparsely. Furthermore, distances in very high-dimensional spaces cease to provide discriminative power since distances between any two points in such spaces are approximately equal. For these reason, almost all statistical pattern recognition techniques apply a pre-processing step (such as PCA) to project high-dimensional data on to a lower-dimensional subspace before classification. In this paper, we compare the performance of a typical parametric discriminative classifier (LDA) against a non-parametric nearest-neighbor algorithm.

Preliminary results of the algorithms have been obtained using in-water data provided by S. P. Pitt of the Acoustic Research Laboratory of the University of Texas. The data was collected using wideband active sonar, with a 15-100 kHz Linear Frequency Modulated insonifying waveform and a 7 millisecond pulse-length. The target returns are time-series with a 0.5 MHz sampling rate. The dataset contains returns from four mine-like objects (MLO) and two non-mine bottom objects (NOMBO), all located at the same distance from the active sonar. Echo returns for each object were collected from approximately 1700 aspects (spanning 360°). Figure 1 shows a subset of this data for a particular MLO and NOMBO. In each plot, the strength of the echo return (brighter pixels indicating stronger returns) is plotted against range to target on the horizontal axis, and aspect (viewing angle) on the vertical axis. Each data point is a horizontal slice through this plot, and consists of a 1400 element vector, and the aspect at which the particular observation was made is not revealed to the classifier. From these plots, it is clear that the appearances of the MLO and NOMBO change significantly with aspect, but one can see that the echo structure of the MLO exhibits different scattering characteristics. This paper focuses on single-ping classifications .

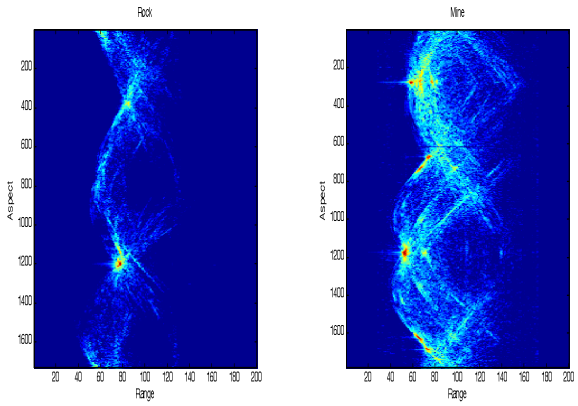


Figure 1: Echo intensity against range and aspect for a typical mine-like object (MLO-1, right) and a similar non-mine bottom object (NOMBO-1, left).

# 3 Baseline Algorithm: Linear Dimensional Analysis

Fisher's Linear Discriminant Analysis (LDA) is a popular discriminative classifier. It finds the optimal linear mapping that best separates the data points into the classes, by maximizing the ratio of between-class scatter[1] to within-class scatter (see [3] for details). The between-class $S_b$ and within class $S_w$ scatter matrices are defined as:

$$S_b = \sum_{i=1}^{c} N_i(\mu_i - \mu)(\mu_i - \mu)^T \qquad S_w = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

where $\mu_i$ is the mean of each class $X_i$, $N_i$ is the number of samples in class $X_i$, and $c$ is the number of classes. LDA chooses the transformation A that maximizes the ratio of the determinant of the between-class scatter matrix to the determinant of the within-class matrix:

$$A_{opt} = \arg\max_{A} \frac{|AS_bA^T|}{|AS_wA^T|}.$$

This is achieved by solving a generalized eigenvalue problem as described in [3]. Note that the number of parameters that LDA must estimate scales poorly with the dimension of the data space, and applying the standard LDA algorithm directly to a high-dimensional data set requires a prohibitively large number of training examples. For this reason, most practical implementations of LDA first pre-process the data using a dimensionality-reduction technique, such as principal components analysis (PCA) [3]. More sophisticated versions of LDA incorporate the dimensionality-reduction directly into the algorithm, and we employ one such algorithm [5] as a benchmark in this paper. Table 1 summarizes classification results using LDA/PCA on several pair-wise recognition tasks. Half of the data from each set was used for training the classifier, and the other half was used for testing.

Table 1: LDA/PCA Classifier Mine Recognition Results

| LDA/PCA Classifier | False Positives | False Negatives | Accuracy |
|---|---|---|---|
| **MLO-1 v NOMBO-1** | 542/865 | 179/865 | 58.3% |
| **MLO-1 v NOMBO-2** | 0/870 | 0/870 | 100.0% |
| **MLO-2 v NOMBO-2** | 210/840 | 0/840 | 87.5% |
| **MLO-3 v NOMBO-2** | 54/890 | 129/890 | 89.7% |
| **MLO-4 v NOMBO-1** | 0/865 | 35/865 | 98.0% |

We observe that the LDA classifier gives reasonable results for most of the discrimination tasks but performs very poorly on MLO-1 v NOMBO-1. This is because MLO-1 is designed to resemble a single, large rock (NOMBO-1). However, a close examination of Figure 1 reveals that the two objects exhibit different scattering characteristics,. This motivates our search for an automated solution that can reliably distinguish between them.

[1] The scatter matrix quantifies the spread of a set of data points.

# 4 Nearest-Neighbor Algorithm for Mine Recognition

Our technique is motivated by research in face recognition showing that nearest-neighbor classifiers, in conjunction with non-Euclidean similarity measures, achieve high recognition accuracy on reduced-resolution face images [4]. The basic idea is straightforward: elements from the training data are first pre-processed to reduce dimensionality from 1400 to 50 (by averaging using non-overlapping box-filters), and these data points are stored. During testing, each observed vector is averaged in the same way, and then compared to the each element in the stored set using the following similarity function:

$$L_p = \sum_{k=1}^{50} \left| x_i(k) - x_j(k) \right|^p ,$$

where $x_i(k)$ and $x_j(k)$ are the k$^{th}$ components of each data element, and p=0.2 is the norm of the similarity function (see below). The class label corresponding to the best match (stored data element with the smallest similarity score) is returned.

Table 2 summarizes the results of our technique on the experiments discussed above. Note that the nearest-neighbor algorithm achieves excellent accuracy on all five experiments; in particular, it discriminates reliably between MLO-1 v NOMBO-1. We hypothesize that this may be because LDA/PCA is forced to fit a single parametric model to all viewing aspects of the data while our algorithm can represent the different aspects individually with appropriate samples.

Table 2: Nearest Neighbor Classifier Mine Recognition Results

| Nearest-Neighbor | False Positives | False Negatives | Accuracy |
|---|---|---|---|
| **MLO-1 v NOMBO-1** | 0/865 | 1/865 | 99.9% |
| **MLO-1 v NOMBO-2** | 0/870 | 0/870 | 100% |
| **MLO-2 v NOMBO-2** | 0/840 | 0/840 | 100% |
| **MLO-3 v NOMBO-2** | 0/890 | 0/890 | 100% |
| **MLO-4 v NOMBO-1** | 0/865 | 0/865 | 100% |

The following two experiments explore the generalization performance of our technique and explore the effect of changing certain aspects of the algorithm. First, we examine the impact of training and testing the algorithm with data collected under different conditions. The training set contains only exemplars of MLO-1 and NOMBO-1 placed on a smooth gravel seabed. The testing set contains data from the MLO1 objects collected on a rough gravel seabed. Table 3 summarizes the performance of our algorithm. It also shows how changing the norm of the similarity function affects the classification accuracy. Note that the obvious similarity function (Euclidean p=2) performs rather poorly, and the best results are given by a non-metric (p < 1) function. As discussed in [4,8], this is because the Euclidean metric is overly sensitive to outliers in any dimension, while non-metric similarity functions mute the impact of outliers on the matching score.

Table 3: Generalization performance: impact of different training/test conditions and effect of varying similarity functions.

| Nearest-Neighbor: MLO-1 v NOMBO-1<br>Train: rough gravel; test: smooth gravel | False Positives | False Negatives | Accuracy |
|---|---|---|---|
| **p=2.0 (Euclidean norm)** | 0/865 | 267/865 | 84.6% |
| **p=1.0 (Manhattan norm)** | 0/865 | 208/865 | 88.3% |
| **p=0.5** | 0/865 | 158/865 | 90.8% |
| **p=0.2** | 0/865 | 150/865 | 91.3% |

The final experiment examines the generalization performance of the algorithm when the training data is derived from a certain Depression/ Elevation (D/E) viewing angle of the object, and the test data uses a different D/E viewing angle. Table 4 summarizes the results, and also shows the effect of omitting synthetic exemplars from the training set. We note that our technique generalizes well to different D/E viewing angles, but only when the training set is augmented with synthetic exemplars. This is because the augmented data adds samples to sparse regions of the space, thereby reducing sensitivity to small variations in intensity and position.

Table 4: Generalization performance: impact of different training/test conditions and effect of omitting synthetic exemplars from training data.

| Nearest-Neighbor: MLO-1 v NOMBO-1<br>Train: D/E 8.5; Test: D/E: 18.5 | False Positives | False Negatives | Accuracy |
|---|---|---|---|
| **Without synthetic exemplars** | 0/865 | 779/865 | 55.0% |
| **With synthetic exemplars** | 0/865 | 1/865 | 99.9% |

Experiments with other MLO-x/NOMBO-y pairings give similar results (not included in this paper).

# 5 Discussion

In order to understand the behavior of the algorithms, we first examine the acoustic backscatter process from the objects of interest. At these frequencies (15-100 kHz) the broadband echo responses from the contacts consist of many highlights corresponding to the prominent reflectors of the contact. Hence the amplitude and position of these highlights vary with the contact aspect angle and shape.

The nearest-neighbor algorithm is essentially a template-matching procedure that uses a large number of static templates representing the range of anticipated contacts and the aspects associated with each contact. While a brute force correlation algorithm on the original (high-dimensional) space would be computationally intensive, the reduced dimension template matching is manageable. Reducing dimensionality by smoothing also makes the algorithm less sensitive to small registration errors. A traditional template-matching/correlation procedure would employ a Euclidean measure that is sensitive to amplitude and positional outliers. Our use of non-metric similarity functions and additional synthetic prototypes facilitates generalization.

The LDA/PCA algorithm tries to shape the scatter in order to make it more reliable for discrimination. The shape of the scatter is derived from the covariance matrix of the classes. The implied assumption here is the scatter between the classes is greater than the within class scatter. In some of our examples the aspect dependant variation of the vectors of each class can be large. To overcome this problem, each class could be divided into subclasses, where each subclass would correspond to regions in aspect space where the contacts were similar in shape. Classification would then be based on similarity between the observed point and the nearest subclass representation. This hypothesis is supported by our initial experiments, where PCA/LDA trained using 9 such subclasses, with each subtending 40° of aspect, achieves a marked improvement in performance. Breaking a class into subclasses is not necessary for the nearest-neighbor algorithm, since the data is not assumed to fit any particular parametric distribution, and class labels are associated with each individual point.

# 6 Conclusions and Future Work

This paper demonstrates that broadband acoustic returns in the 15-100 kHz band encode information that can be used to discriminate mine-like objects from non-mine bottom objects such as rocks. Although the superior processing gain of broadband sonars provides increased signal-to-noise ratio, the data generated is very high-dimensional compared to that generated by narrow-band sonars. The nearest-neighbor algorithm, coupled with a non-metric similarity function capitalizes on the strong echo highlights produced by broadband sonars while combating the complexity of high-dimensional data. This approach has the potential to be used with forward- or side-looking sonars, as well as synthetic aperture systems in a side-looking configuration. We are extending our work to the problem of buried mine recognition, where we believe that reliable estimation of acoustic color will facilitate classifying buried contacts.

# 7 Acknowledgements

The authors thank Dr. John Tague (Office of Naval Research) and Mr S.P.Pitt of the Acoustic Research Laboratory , University of Texas for providing us with the in-water data. The work of S.Reese was supported by Naval Undersea Warfare Center's internal research funds.

# 8 References

[1] F. B. Shin and D. H. Kil, "Robust mine detection and classification with target physics-derived features and classifiers", in Proceedings of the IEEE International Conference on Oceans 1996, Fort Lauderdale, FL, September 1996.
[2] A. Trucco , "Detection of Objects Buried in the Seafloor by a Pattern-Recognition Approach", IEEE Journal of Oceanic Engineering, Vol. 26, No. 4, October 2001.
[3] R.O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification,* New York: Wiley, 2001.
[4] T. Sim, R. Sukthankar, M. Mullin and S. Baluja, "Memory-based Face Recognition for Visitor Identification", in Proceedings of the 4th International Conference on Face and Gesture Recognition, Grenoble, France, March 2000.
[5] J. Yang, H.Yu and W. Kunz, "An Efficient LDA Algorithm for Face Recognition", International Conference on Automation, Robotics and Computer Vision, ICARCV'2000.
[6] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, July 1997.
[7] Whitlow W.L.Au, *The Sonar of Dolphins*, Springer-Verlag NewYork, 1993
[8] D.W.Jacobs, D Weinshall, Y Gdalyahu, "Classification with Non-Metric Distances: Image Retrieval and Class Representation" Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems 11.