Side Scan Sonar Object Classification Algorithms

M.F. Doherty, J.G. Landowski, P.F. Maynard,
G.T. Uber, D.W. Fries and F.H. Maltz

Lockheed Missiles & Space Co., Inc.
Astronautics Division - Marine Systems
Santa Clara, California

Abstract

Autonomous underwater vehicles require the capability to understand their environment. This understanding, coupled with the operational goals of the vehicle, determines the subsequent actions of the vehicle. Environmental understanding is realized through the vehicle's sensors and a priori knowledge. This paper focuses on our investigation of automatic interpretation of side scan sonar data for the purpose of detecting and classifying undersea mines.

The test data set is a series of side scan sonar images taken from a U. S. Navy acoustic sensor under optimal conditions (flat, sandy bottom). Groundtruth is available for acoustic images with eight unique types of mine targets. The interpretation of the data is performed in two stages. The first stage, preprocessing and target ***detection***, uses an adaptive thresholding algorithm coupled with an adaptive averaging technique to locate objects of interest in the sonar image. The second stage, ***classification***, performs a binary classification of whether each detected object is, or is not, a mine. The classification is achieved using an attribute-based decision tree. An approach for a third stage, ***identification*** of the mark and mod of the classified bottom mines, is also presented.

The results to date, as well as plans for the future, are discussed.

1.0 Introduction

There are a variety of missions for autonomous undersea vehicles (AUVs) which require an automatic recognition of objects and topographic details within the immediate vicinity of the vehicle. Our research is aimed at providing the technology to support mine countermeasure missions for AUVs. In particular, we are interested in developing algorithms which will automatically detect and recognize ocean floor "bottom" mines using side scan sonar data.

Our general approach is to utilize existing well-known techniques found in the automatic target recognition[1] and image understanding[2] domains, and apply these techniques to the side scan sonar's acoustic data. Implicit in this approach is the desire to build algorithms which are easily adapted to take advantage of higher resolution data, and which are suitable for detecting and recognizing objects other than, but similar in dimension to, bottom mines. This paper describes in detail the experimentation which has been initiated this year at Lockheed, and our preliminary results following several months of effort.

2.0 Experimental Data

Our test data set consists of 33 images, each with an identified bottom mine target. Each image is a digital sampling of acoustic data from a side scan sonar, with 512 rows corresponding to sampling over time of 512 acoustic beams. There are 1024 time samples (columns of data) per image, with 8 bits per sample (pixel). High data values indicate high energy acoustic returns, and low data values indicate low energy acoustic returns. (Thus, the mine targets appear as bright blobs and downrange shadows appear as dark areas.) The data was collected over a flat sandy bottom, with the bottom mines resting proud of the bottom (not buried). Given an estimate of the total range of the sensor in meters, an examination of the water column and some trigonometry, one can derive a good estimate of the sensor altitude above the sea floor for each image. A typical sensor altitude for this imagery is 20 meters. The ocean bottom in this data set has long linear features which appear to be shallow narrow trenches, or tracks. These have a fairly random orientation, and sometimes coincide with the physical positioning of the mine. In addition, a significant percentage (74%) of the imagery contains surface noise and saturation, as well as the more typical motion blur and environmental noise.

3.0 Detection

Detection, for the purposes of this paper, is defined to be the automatic specification of the locations of objects of interest in the image. These objects may or may not be bottom mines. Our detection strategy incorporates two principal algorithms, adaptive thresholding and shadow analysis on data smoothed by adaptive averaging.

3.1 Adaptive thresholding.

This method for the automated detection of undersea bottom mines using acoustic imagery is predicated on two requirements: two or more contiguous pixels in the image must be brighter than the mean graylevel in the surrounding neighborhood, and there must be a shadow area that extends downrange from the bright pixels that is lower than the mean graylevel in the surrounding neighborhood. Ideally, the bright pixels would exceed a threshold derived from the graylevel histogram and set to κ standard deviations above the mean. However, due to the extreme variability of the data within and between images, simple thresholding proved infeasible. Some data sets were heavily saturated (graylevel 255) over large portions of the image; in others the brightest pixels were under graylevel 200. Some images were very homogeneous while others exhibited large changes in variance as a function of downrange position.

Therefore, the brightest pixels in each neighborhood are found by adapting the threshold cutoff value to each neighborhood. A moving window is passed over the image in half-window increments. A histogram operation is performed, and the histogram bins are summed from graylevel 255 toward graylevel 0 until the sum is equal to or greater than a histogram count parameter, δ. Setting the threshold at the corresponding graylevel guarantees that at least δ pixels within the window are available to be chosen as objects of interest. The center pixels of this window are then sequentially examined, and the graylevels of pixels with graylevels equal to or greater than the threshold are written to the output image; zeros are written for those pixels that are less than the threshold. Moving the window creates an overlap that averages the boundary conditions between high and low variance areas and creates contiguity between the processed centers. At the end of the scan line and at the top and bottom of the image, the processed area is extended to the image edge.

3.2 Shadow Analysis

The thresholded image may have several thousand "pixels of interest". This number is reduced prior to the computationally expensive steps of shadow analysis and mine identification by rejecting all single pixels. A candidate group of pixels is accepted as an object of interest only if an area of low pixel intensity (whose principal axis is oriented downrange) is present immediately downrange. Automated labeling of shadow versus non-shadow using the raw data is not feasible because of the confounding effects of the sonar sidelobes. The sidelobes tend to break up the signal and to scatter data points, creating random speckle that obscures edges and creates bright spots in the middle of valid shadows and dark spots in the middle of valid objects.

Adaptive averaging is used to preprocess windows created around candidate objects of interest. It smooths graylevels within low gradient change areas and preserves high gradient transitions. Averaging is performed by obtaining the Sobel edge image of the window, moving a smaller window over the Sobel image, and determining which pixel has the lowest value (the smoothest area in the raw data). The center output pixel of the window is replaced by the mean of a still smaller window in the raw data centered over the Sobel low value pixel.

Visual examination of valid shadow patterns indicated that shape analysis of the shadow in the adaptively averaged window via the method of moments would be infeasible due to the high probability that the shadow would intersect other low graylevel features and radically alter its shape. The adaptive averaging algorithm guarantees that if there is a valid object and shadow (in the original data), then there will also be a contiguous region of low intensity pixels directly downrange from the object (in the smoothed image). The key question is to determine if the candidate shadow region is low enough in intensity to qualify. An average graylevel from the adaptive averaged image uprange of the object is compared with the average graylevel of a rectangle centered on the expected area of the shadow. If the shadow average is more than α graylevels less than the average uprange of the object, it is accepted as a valid shadow. As α is decreased, more shadows are accepted; as α is increased, more are rejected. Visual inspection of many of the accepted non-mine objects indicates that their similarity to mine targets is very high.

4.0 Classification

The second stage of the processing is classification, which is theoretically defined as the labeling of each detected object of interest as either "MINE" or "NOT MINE". However, the limited resolution of the particular side scan sonar sensor used in these experiments does not permit distinction between actual mines and non-mine objects which could have shapes and dimensions similar to mines. Therefore, for the purposes of this paper, classification is defined as the labeling of each detected object of interest as either "MINE-LIKE OBJECT" or "NON-MINE-LIKE OBJECT".

The input to this algorithm is a set of objects of interest determined by the detection stage. Classification is performed in four steps: locating the downrange end of the shadow, computing the ratio of object to shadow average intensity, computing the diameter of the object, and classifying based upon these two measured attributes.

An estimate of the median downtrack position of the object is made given the parameters passed by the detection algorithm. A scan is then performed in the downrange direction to find the start of the object's shadow (characterized by a sharp negative gradient in the pixel intensities). Once found, a moving average of the shadow intensity values is calculated over λ pixels. The downrange edge of the shadow is located when the moving average exceeds a threshold parameter value (μ), indicating a rise in the intensity values.

The ratio of object to shadow intensity is calculated by averaging the high energy graylevel values of the object pixels, and dividing by the low energy graylevel values of the shadow pixels. The object diameter is calculated using a similar triangle ratio of the sensor altitude and distance to the end of the shadow equated with the object height and length of shadow.

The final classification step involves three threshold parameters, ρ, η_1 and η_2. These threshold parameters (as well as the others mentioned above) have been empirically determined on a trial-and-error basis by running the algorithms on the data set and observing the quality of the results. This approach is quite feasible since both the data set and the number of parameters require relatively small amounts of processing. A more standard approach, which utilizes parametric and non-parametric statistical classification algorithms (e.g., the Fisher linear discriminant) to devise the optimal selection of parameters given the source data, will be used in the future as the complexity of the algorithms and the resolution of the data increase.

Our trial-and-error experimentation showed that mine-like objects in the test data set have consistently high object/shadow intensity ratios ($> \rho$), and have diameters which consistently fall within a range of values ($\eta_1 < \text{diam} < \eta_2$). Knowing this, a decision tree was implemented. If the average intensity ratio of a detected object exceeds ρ, and if the object diameter falls within the range of η_1 and η_2, then the object is classified by the decision tree as a mine-like object; otherwise, it is classified as a non-mine-like object.

5.0 Identification

The final stage in automatic bottom mine recognition algorithms is that of identification. A principal identification question which often needs to be answered for the purpose of selecting effective mine countermeasures is, "What kind of influence mechanism does this mine have?" A second question of interest is, "Who manufactured this mine?" These questions cannot be answered without a detailed automatic analysis of the geometric shape of the object. It is therefore necessary to construct a database of geometric models of the mines to be identified, in order to match the sensed shape of the object with the known shape of the mine. We propose the use of a "Mark-Mod-OA" taxonomy[3] to aid in the identification of the mine. The mine's shape can be determined by reference to the mine casing dimensions (the Mark, or "Mk") and any bulky shape-altering accessories such as tail fins or nose fairings (specified by the Operational Assembly, or "OA"). The key component of interest is the Modification, or "Mod", which specifies the type of influence firing mechanism (e.g., magnetic vs. acoustic) used by the mine. We therefore define identification to be the processing stage at which the Mark ("Mk"), Modification ("Mod"), and Operational Assembly ("OA") of the bottom mine are determined.

We have recently begun development of a set of bottom mine identification algorithms for use with side scan sonar data. These proposed algorithms will take the "mine-like objects" as input and perform a geometric shape analysis of the sonar reflections of these objects. Although an accurate shape analysis is necessary, it is undoubtedly not sufficient to complete the mine identification process. Once a shape is determined, additional exploitation of the data can be performed (most likely with the use of additional sensor input) to identify the influence mechanism and the source of manufacture. Our proposed processing strategy for initial shape analysis involves four steps: hypothesis generation, model to image matching, statistical verification, and hypothesis ranking.

A hypothetical bottom mine is generated (internally in the computer program) which has the geometric shape of a real mine resting proud of the bottom, with a chosen orientation. The computer then generates the acoustic reflections and shadow which would be cast by this mine (given the altitude of the sensor and the distance of the mine from the sensor). This mine-shadow model, or hypothesis, is then matched to the acoustic imagery at the location where the mine-like object has been classified. An automatic statistical analysis[4] is then performed to determine if the sensed image data "looks like" (has similar intensity patterns to) the acoustic patterns generated by the computer for the hypothetical mine. Hypotheses which fit poorly with the sensed data are immediately rejected. A ranking of the best fit hypotheses is derived from the statistical analysis, and the most highly ranked hypothesis is selected as the correct identification of mine shape and orientation.

Experience and some initial experimentation indicate that an accurate shape analysis using these algorithms will require a side scan sonar sensor which produces as least twice the resolution, if not more, than the sensor which produced the current data set.

6.0 Experimental Results

6.1 Detection Results

A series of experimental runs was conducted over the test imagery. When the detection threshold percentage parameter is adjusted to ensure 100% bottom mine detection (which is appropriate for most mission scenarios), our initial results indicate an average non-mine-like object detection rate of 234 per image. By varying the shadow analysis parameters, processing can be done to reduce the number of detected non-mine-like objects at the expense of some missed actual bottom mine detections. Our initial results indicate that, with the detection algorithms as described, one can achieve an average of 78% detection of bottom mines with an average of only 36 non-mine objects detected per image.

The detection algorithms took approximately 5 CPU seconds to complete on a single acoustic image (512 x 1024 pixels), when running on a Sun 3/60FC workstation. These algorithms are regarded as a significant part of the total computation budget. A complexity analysis will be performed in the near future to determine realtime system requirements using these algorithms as a baseline.

6.2 Classification Results

Our classification algorithm was run on the output of the object detection algorithms which included 100% of the bottom mines found in the imagery. Tests of the "mine-like or non-mine-like" classification algorithm show that the algorithm automatically classifies 29 of the 33 mines as mine-like objects (88% correct classification), and classifies 97% of the non-mine-like objects correctly (roughly 6400 out of 6600). There is an average of 6 false targets for each mine that is classified, and a maximum of 33 false mine targets in one particular image. Table 1 below summarizes the results of the experimentation.

The classification algorithms took approximately 1.5 CPU seconds to complete on a single acoustic image (512 x 1024 pixels), when running on a Sun 3/60FC workstation. These algorithms are currently not considered to be a computational bottleneck; however, highly complex attribute measurements (e.g., cooccurrence texture measures, or cross-correlation pattern analysis) could become a significant computational burden in the future if their inclusion is necessary (due to increased accuracy).

7.0 Conclusions and Future Plans

We conclude that the application of automatic target recognition and image understanding algorithm methodologies to the undersea mine recognition problem using acoustic imagery is a viable technical approach, with great potential utility. We will continue to develop the ability of the detection and classification algorithms to correctly recognize all mine targets in side scan sonar imagery, while continuing to reduce the false target rate. In particular,

Table 1. Classification Results.

Target	Image	Mine Found?	# False Mines
1	4309	YES	0
1	4329	YES	0
1	5119	YES	0
2	4304	YES	1
2	4317	YES	1
2	5113	YES	4
2	5122	YES	5
2	5134	YES	33
2	5216	NO	18
2	5228	YES	5
3	4312	YES	8
3	4313	YES	7
3	4327	YES	8
3	5100	YES	2
3	5128	YES	0
4	4318	YES	1
4	5112	YES	18
4	5123	YES	3
4	5215	YES	5
5	4302	YES	4
5	4331	YES	11
5	5115	YES	13
5	5222	YES	1
6	5106	YES	16
6	5132	YES	6
6	5224	NO	1
7	4315	YES	10
7	4325	NO	3
7	4333	NO	1
7	5218	YES	1
9	4323	YES	0
9	5102	YES	9
9	5117	YES	0

we will investigate the robustness of the algorithms by applying them in the future to acoustic data of targets on other types of ocean bottoms.

8.0 Acknowledgments

We wish to acknowledge the technical and managerial support of Bob Yule and Steve Spratt of LMSC/Marine Systems. We also wish to acknowledge and thank the Naval Coastal Systems Center (NCSC) for providing us the acoustic imagery under contract number NICRAD-87-NCSC-012, "Mine Hunting Technology", and Dr. Lyles Adair of NCSC for his support and encouragement. This research was performed as part of a Lockheed Missiles & Space Co., Inc., Independent Development Project AD-707 "Detection, Classification, and Identification".

9.0 References

1. J. G. Landowski, R. S. Loe, and C. M. Bjorklund, "Feature Extraction for undersampled objects in range imagery", Proc. SPIE Conf. on Digital and Optical Shape Representation and Pattern Recognition, vol. 938, 4-6 April 1988, Orlando, Fl., pp. 289-294.

2. DARPA/ISTO, *Proc. Image Understanding Workshop: May, 1989*, San Mateo, Ca., Morgan Kaufman Publishers, Inc., 1989.

3. L. H. Petersen, Cdr, USN, *NAVSEA Mine Familiarizer*, Yorktown, Virginia, U.S. Naval Mine Engineering Facility (NAVMINENGRFAC), 1982.

4. R. E. Walpole and R. H. Myers, *Probability and Statistics for Engineers and Scientists*, New York, The Macmillan Company, 1972, p. 360.