Target Confirmation Architecture for a Buried Object Scanning Sonar

Daniel D. Sternlicht, R. David Dikeman, David W. Lemonds,
Matthew T. Korporaal, Arthur M. Teranishi

ORINCON Defense
4770 Eastgate Mall
San Diego, CA 92121

Abstract **– Efficient mine clearing operations are essential for maintaining sea lines of communication and for the timely dispatch of military and economic supplies to conflicted areas. To locate stealthy buried mines, a future naval system of systems is under development that incorporates high-resolution acoustic and electromagnetic sensors. This paper describes an evolving target confirmation architecture for this program's buried object scanning sonar that utilizes image and signal classification strategies. Feature extraction from the 3-D sediment volume imagery is described. Image classification using a joint Gaussian Bayesian classifier is demonstrated with synthetic 2-D image classification experiments that employ image clustering and ellipse feature extraction methods. The signal classifier is demonstrated with data collected from mine-like and clutter objects buried in sand. These tests utilize data from different run orientations and transmit angles for training, cross validation and testing, achieved 5-class classification levels of 94% and 2-Class ROC curve knee values of Pcc = 96% and Pfc = 4%, and thus illustrate the buried mine-hunting potential for time-frequency based signal classification.** [1]

I. INTRODUCTION

Efficient mine clearing operations are essential for maintaining sea lines of communication and for the timely dispatch of military and economic supplies to conflicted areas. As demonstrated in Gulf War II, Explosive Ordnance Disposal units, to a large extent, rely on marine mammal and diver teams to locate and identify mines [1]. For automated localization of stealthy buried mines, a future naval system of systems is under development that incorporates high-resolution acoustic and electromagnetic sensors [2]. This paper describes an evolving target confirmation architecture for this program's buried object scanning sonar.

To locate buried mines, an adaptive acoustic classification architecture was described [3] that fuses object classification clues extracted from sub-bottom seafloor imagery [4] created with a prototype 14-kHz center-frequency broadband Buried Object Scanning Sonar (BOSS) [5] with signal classification algorithms [6] that operate on echo time-series data collected with this system. The concept is outlined in the block diagram of Fig. 1. Objects buried below the sediment-water interface are localized using detection algorithms (module 1). The voxel characteristics associated with the object are evaluated with image classifiers (module 3) and focused echo time-series are evaluated with algorithms trained to recognize acoustic scattering characteristics in the time-frequency plane (module 2). Outputs from both the image and signal classifiers are to be combined using an algorithm fusion engine (module 4).

Section II of this paper describes the system and data; Section III presents feature extraction for 2-D and 3-D imagery and investigates image classification using a joint Gaussian Bayesian classifier; Section IV presents results of training and testing the proposed signal classification engine with data from several target field transits; Sections V and VI describe the overall acoustic data target confirmation architecture and summarize results, limitations, and follow-on efforts.

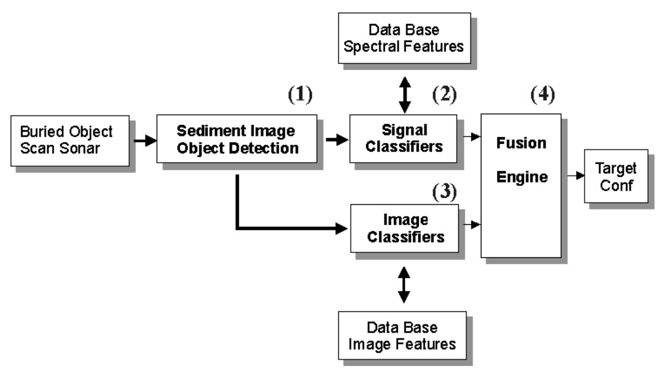

Fig. 1. Buried mine classification block diagram.

II. SYSTEM AND DATA

The prototype BOSS system [5] was developed in 1999-2000 by Florida Atlantic University, EdgeTech Inc., and Sea Engineering Inc. under funding from the Office of Naval Research (ONR) and the National Defense Center of Excellence for Research in Ocean Science (CEROS). The classification methods discussed in this paper are currently being developed with data collected from that system; however, they are intended for use with a new system currently under construction and described in [7]. The BOSS system [5] was designed to achieve imaging resolutions of 4 to 8 cm for objects buried 1 to 2 m below the seafloor. Using a 6-element line transmitter, the system transmits a 2-msec, 5-23-kHz Linear FM signal, and receives using a 32-element planar array.

A controlled test range was designed and installed by Sea

[1] This work was sponsored by the Office of Naval Research under contract # N00014-02-C-0185.

Engineering Inc. at their Waimanalo, Hawaii research facility [8]. Test range targets include objects buried within one meter of the sand substrate sediment-water interface, at various aspects and offsets to a centerline track. The objects summarized in Table I, and more thoroughly described in [3,4], were buried in February 2001. Data used in this paper were collected in April 2001. The BOSS system was towed at 0.7 knots and used a 20 Hz ping repetition rate to create 2-D vertical slice images that are stacked together to form a 3-D volume representation of the seafloor (Refs. 3,5 and Sec. III).

For demonstration of scale and testing of image classification algorithms (Sec. III), Fig. 2 shows a synthetic floormap image of these objects. This simulation employs elliptic and rectangular geometries and a signal to random noise ratio of 15 dB.

TABLE I
CATALOGUE OF TEST OBJECTS

Object Number	**Description**	**Burial Depth** (m)
T1	Concrete block, 45.7 cm x 45.7 cm	0.3
T4	Aluminum sphere, sand-filled 45.7 cm diameter	0.3
T7	Steel scuba tank, concrete-filled 70.0 cm x 17.8 cm	0.3
T9	Minelike-shape, concrete-filled steel cylinder, 45.7 cm x 198.1 cm	0.5
T10	155mm Ordnance Shell, 15.5 cm x 91.4 cm	0.3

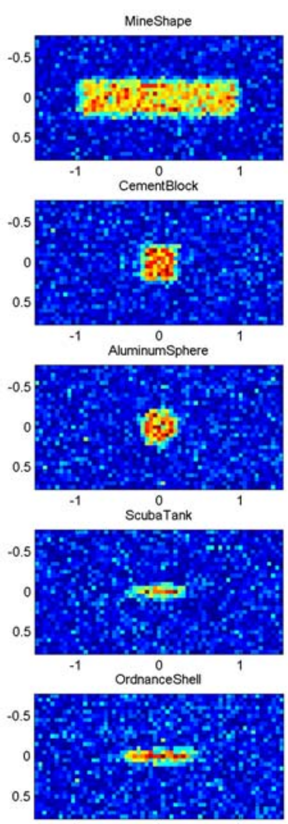

Fig. 2. Synthetic Floor Map Imagery for Table I Buried Objects

III. IMAGE CLASSIFICATION

This section describes the application of object detection, clustering, and image feature extraction algorithms to synthetic and real data.

Image clusters are formed by intensity thresholding and outlier removal, followed by an object growth algorithm that is guided by measures of inter-cluster gaps through an optimal search strategy.

Features extracted from these clusters are passed to a joint Gaussian Bayesian target classifier [4,9] that measures feature distributions, transforms the original feature distributions into Gaussian distributions, computes feature covariance to exploit feature correlation, and classifies the detected object using a log-likelihood ratio test:

$$\mathrm{LLR} = \frac{1}{2}\left\{ \left(\mathbf{x}_{test} - \mathbf{x}_c\right)^{\mathrm{T}} \mathbf{R}_c^{-1} \left(\mathbf{x}_{test} - \mathbf{x}_c\right) - \left(\mathbf{x}_{test} - \mathbf{x}_m\right)^{\mathrm{T}} \mathbf{R}_m^{-1} \left(\mathbf{x}_{test} - \mathbf{x}_m\right) \right\} + \frac{1}{2}\ln\frac{|\mathbf{R}_c|}{|\mathbf{R}_m|}$$

where

$\mathbf{x}_{test}$ = transformed measured feature vector

$\mathbf{x}_m, \mathbf{x}_c$ = mean vectors of classification features for target and clutter training populations

$\mathbf{R}_m, \mathbf{R}_c$ = covariance matrix of target and clutter training data sets

The LLR is based on a joint least-squares covariance-weighted estimate of the difference in test detection from target and clutter mean training vectors. The value of the LLR is compared to a threshold to determine the test detection's membership in the target or clutter population.

The 5-cm image resolution of the prototype BOSS system creates buried target images that are roughly ellipsoidal in shape. We thus employ 3-D eigenanalysis to extract the following geometric and image intensity features from clustered 3-D sediment volume images: ellipsoid axes lengths, object volume, cumulative energy, and energy density (cumulative energy / volume).

A. Synthetic 2-D Data

Preliminary classification tests were conducted with 2-D synthetic image data, such as shown in Figs. 2 and 3. In Fig. 3, synthetic images of two ellipse shapes are displayed at 5 cm pixel resolution and a signal to noise ratio (SNR) of 10 dB. The large ellipse has major and minor axes lengths of 2.25 m and 0.45 m, and the small ellipse has major and minor axes lengths of 1.5 m and 0.375 m respectively. The images are convolved with a smoothing filter to blur edges, and noise is modeled simply as a standard normal process. The *a priori* probabilities of target and clutter are set equal.

The ellipses major and minor axes lengths fluctuate independently for each of 1000 realizations by a specified variance. This fluctuation loosely mimics the size differentials observed in real data. Figure 4 illustrates an example of the thresholding and clustering outputs for a large ellipse exemplar. The five extracted features for these 2-D experiments consist of: ellipse axes lengths, object area, cumulative energy, and energy density (cumulative energy / area).

The classifier performance was tested on ellipse realizations that varied normally in the dimensions of their major and minor axes by standard deviations of 8%, 11%

and 15%. For 11% variance, the cumulative distribution functions (CDF) for the major axis length feature are plotted in Fig. 5, where it is observed that the 0.5 values match closely the specified lengths of the objects. The separation between the clutter (Large Ellipse) and target (Small Ellipse) distributions along this feature is evident, resulting in a receiver operating characteristic (ROC) curve with a probability of correct classification (Pcc) and probability of false classification (Pfc) pair of 90% and 2% respectively, as illustrated by the blue curve of Fig. 6.

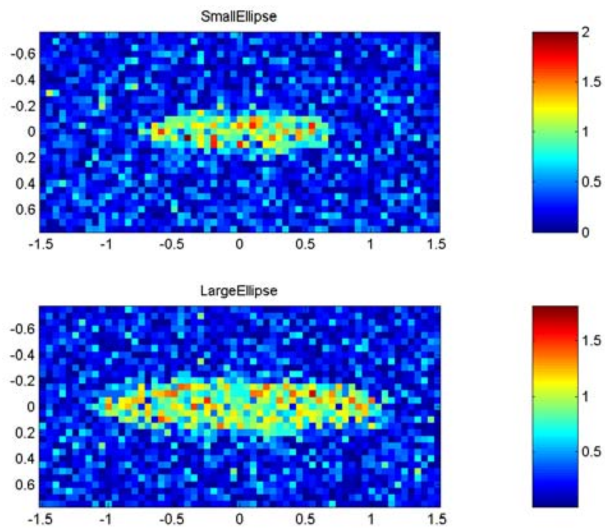

Fig. 3. Synthetic Image Exemplars of Two Ellipse Shapes

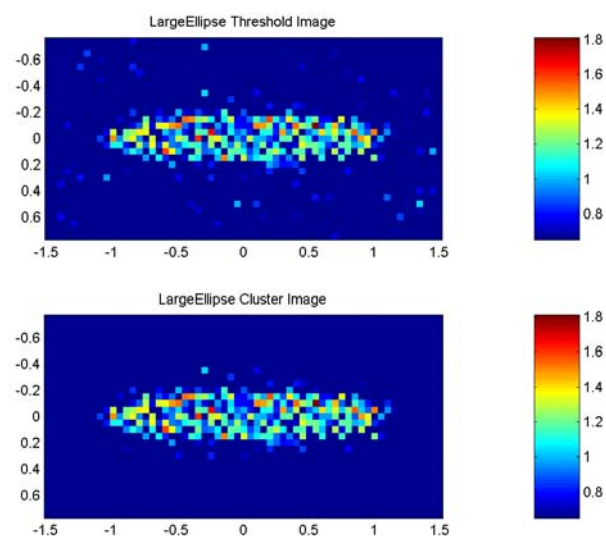

Fig. 4. Threshold and Cluster Outputs for Large Ellipse Exemplar.

The classifier employs transformed versions of the five image features to produce a composite LLR score and corresponding ROC curve (black curve in Fig. 6), where ellipse axes variance is 11%. The classifier yields a high Pcc to Pfc ratio that spans the entire ROC curve, combining the contributions of the individual classification features to maximize overall classification performance. For ellipse axes variance of 11%, a Pcc of 90% has an associated Pfc of 0.8%. For ellipse axes variance of 15%, classification performance is significantly degraded; a Pcc of 90% has an associated Pfc of 5.7%. However, for reduced variation in ellipse dimensions (variance 8%) a Pcc of 90% has an associated Pfc of < 0.1 %.

Inspection of the ROC curves in Fig. 6 reveals that, for this scenario, the length, area, and energy features contribute more useful information than width and energy density features. Prudent selection of feature types employed may thus improve the composite performance of the classifier.

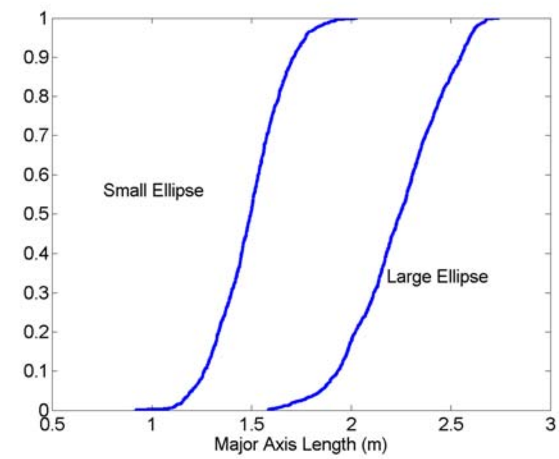

Fig. 5. Cumulative Distribution Functions for Major Axis Length

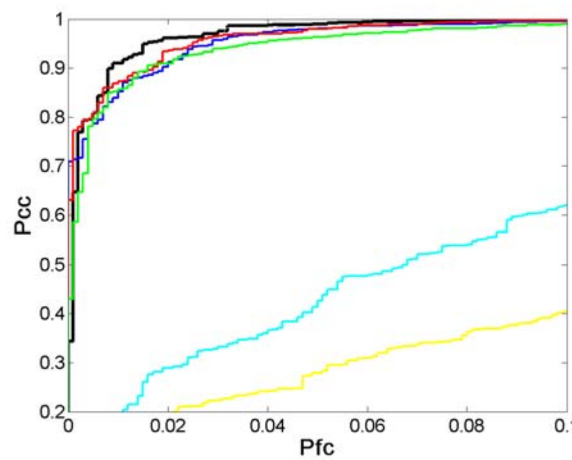

Fig. 6. Receiver Operating Characteristics for Joint Gaussian Bayesian Image Classifier (Black) and for Individual Features. Blue: Major Axis Length; Cyan: Minor Axis Length; Red: Area; Green: Energy; Yellow: Energy Density

This classification method was also applied to the 2-D synthetic floor map imagery of Figure 2, where the mine-like shape was designated the target and the remaining four objects were designated clutter. With variance in axes lengths set to 10% for the rectangular and elliptic images, SNR set to 15 dB to match the expected beamformer performance of the prototype BOSS system, and *a priori* probabilities of target and clutter set equal, the classifier perfectly separated the mine-like shape from the clutter objects. This result is not surprising as the mine-like shape is considerably larger than the clutter objects.

B. 3-D BOSS Imagery

In the prototype BOSS system, a vertical slice image is produced for each sonar transmission (Fig. 7). Real-time object detection is carried out by 1-D integration of pixel energy per voxel over multiple transmissions, and subsequent comparison to a threshold [3,5]. Image feature extraction consists first of subjecting alerted regions in the vertical slice imagery to thresholding and image clustering; then, over the extent of the object a list of 3-D locations and associated intensities is compiled for input to the feature extraction routine. Fig. 8 presents the cluster output for the cement block and mine-like shape vertical slice images of Fig. 7. Table II lists the outputs of the feature extractor run in 2-D mode, illustrating, for these examples, the general agreement of extracted features with displayed imagery.

TABLE II
EXTRACTED 2-D FEATURES

Feature	Cement Block	Mine shape
Axis Length 1	0.46 m	0.85 m
Axis Length 2	0.43 m	0.19 m
Area	0.155 m^2	0.127 m^2
Energy	817412	2131539
Energy Density	5.27 x 10^6 / m^2	1.69 x 10^7 / m^2

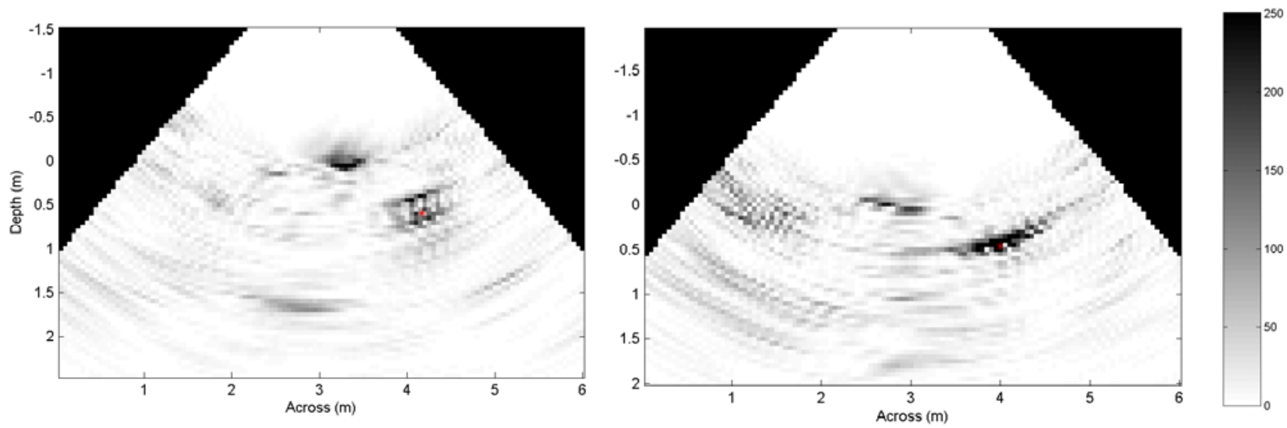

Fig. 7. Cement Block (left) and Mine-like Shape Vertical Slice Images with Detected Centroids (red dot).

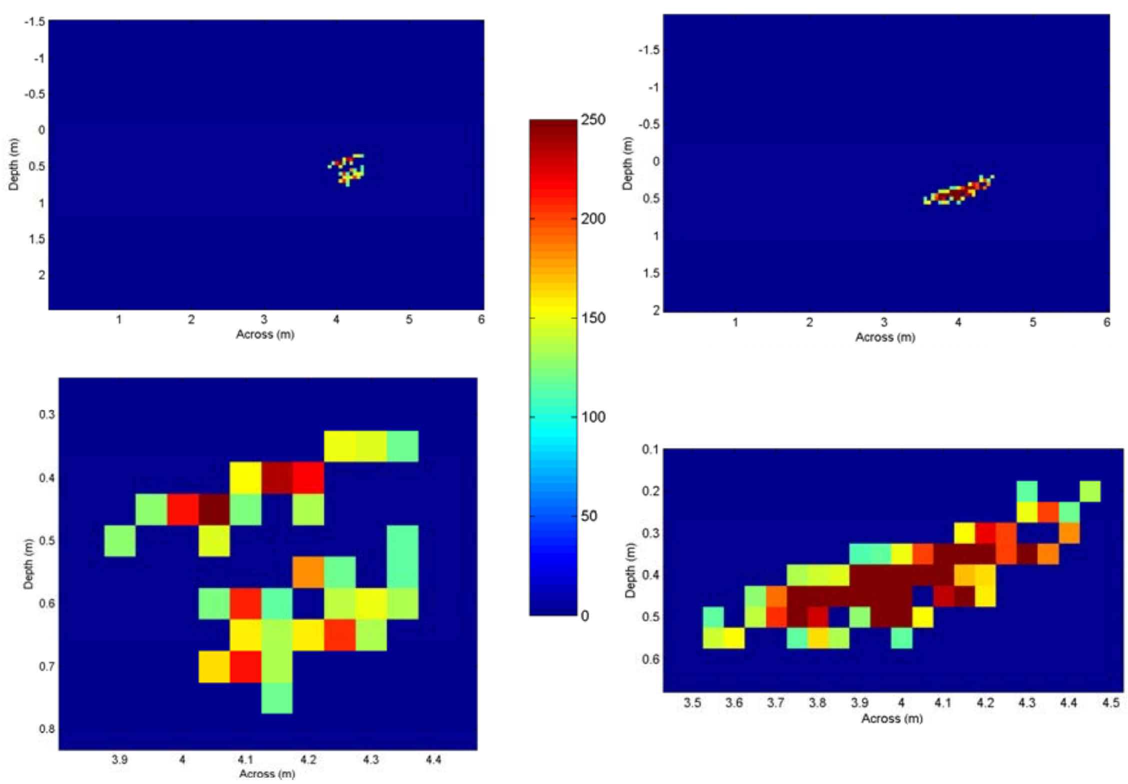

Fig. 8. Cluster Outputs for Vertical Slice Images. Left: Cement Block; Right: Mine-like Shape.

Figure 9 shows 3-D and maximum intensity projection (MIP) representations of the cement block and mine-like shape clusters across the extent of the detected objects. The cement block image demonstrates the dimensions expected from Table I. The image dimensions for the mine-like shape, however, differ considerably from geometric specifications. The short length and the evident gap are likely due to disadvantageous acoustic scattering geometries, and the excessive image width is likely a function of receiver aperture limitations. The relatively narrow width in the depth direction illustrates the advantage of wideband signal compression.

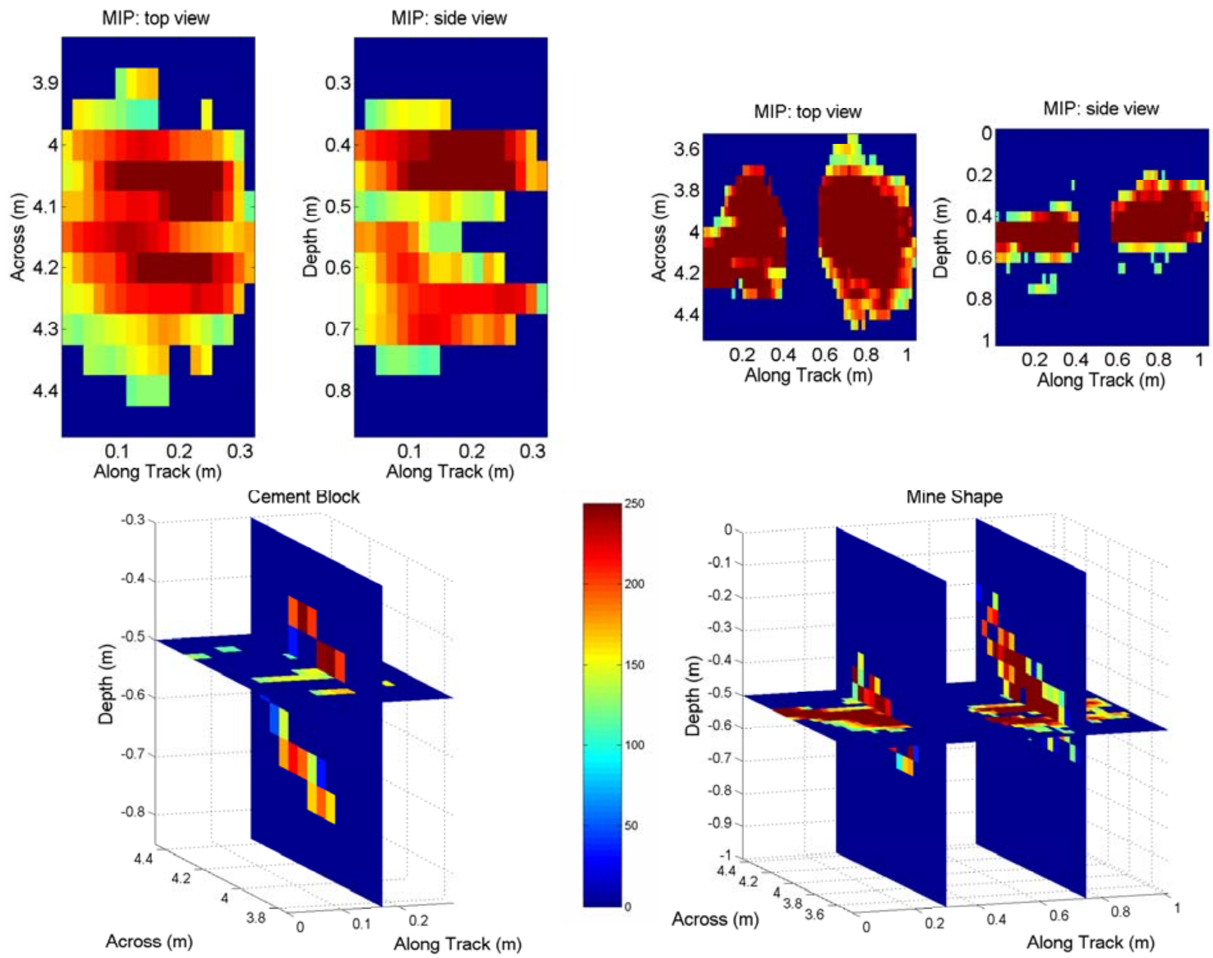

Fig. 9. MIP and 3-D Cluster Images for the Buried Cement Block (Left) and Mine-like Shape (Right)

Table III lists the outputs of the feature extractor run in 3-D mode; again, illustrating the general agreement of extracted features with displayed imagery. The extracted cement block dimensions agree roughly with Table I specifications, and the geometric features extracted for the mine-like shape correspond to its respective imagery. From a classification perspective, the geometric features for the mine shape identify it as an object that is much larger than the cement block. Furthermore, the energy statistics show that the mine-like shape backscatters considerably more energy per unit volume.

C. Discussion

Image classification for objects in this prototype BOSS data set is limited due to the spatial fluctuation of object projections and the saturation of intensity values for many of the man-made metallic objects (evident in the mine-like shape images). Accordingly, we have concentrated on the fundamental image features described. This year's deployment of a new BOSS system, utilizing a larger receive aperture and reflection tomography image creation techniques, should yield a rich data set for testing a variety of related features [10]. Continuing work includes further development of the joint Gaussian Bayesian classifier for 2-D and 3-D image scenarios, and automation of optimal feature selection methods.

TABLE III
EXTRACTED 3-D FEATURES

Feature	Cement Block	Mine shape
Axis Length 1	0.47 m	1.17 m
Axis Length 2	0.39 m	0.69 m
Axis Length 3	0.31 m	0.19
Volume	0.029 m^3	0.080 m^3
Energy	1045287	63647332
Energy Density	3.56×10^8 / m^3	7.80×10^8 / m^3

IV. SIGNAL CLASSIFICATION

Objects alerted in the target detection stage are localized within the images and passed to a focusing utility that employs the spatial coordinates to extract high-SNR target-echo time-series. Currently, signal classification is achieved using the Hybrid Time Frequency Processor Maximum Likelihood Perceptron Classifier [3,6], illustrated in Fig. 10 and summarized in Sec. IV.B.

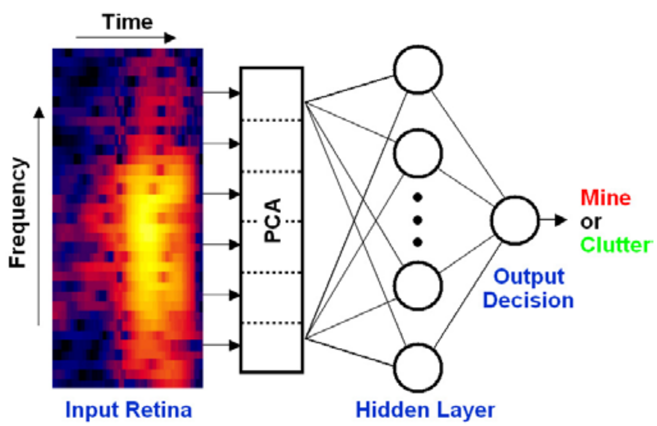

Fig. 10. Hybrid Time Frequency Processor Maximum Likelihood Perceptron Classifier

In previous work [3,4] signal data from a single sonar vehicle pass were used to train the classification engine, and data from a separate pass (in the same direction) were used for cross-validation and algorithm testing. In this section we report 5-class and 2-class classification results for a more challenging experiment. The signal classifier was: trained with echoes collected during a forward pass, for which the sonar's transmit angle was 8^o with respect to vertical; cross-validated with data from a separate forward pass (8^o transmit angle); and tested with data from a backward pass (0^o transmit angle).

A. Signal Preprocessing

Echo time series were amplitude filtered, temporally aligned and normalized. For each target, received signals with peak amplitudes 6 dB lower than the peak amplitude of the strongest signal were removed. Remaining signals were normalized by the time-series amplitude standard deviation. The complexity of envelope returns associated with some targets required that we investigate a number of envelope alignment methods discussed in [11,12], which employ alignment indices derived from measures of envelope amplitude threshold minima, amplitude peak, cumulative half-energy and group delay. A threshold minima index most effectively aligned signals, while maintaining consistent temporal energy distributions required for time-frequency processing.

Figures 10 and 11 display the post-processed training echoes for the buried mine-like shape and cement block objects. Each figure shows the waterfall plot and raster image for echo envelope time series (left and center) and a composite image of the coherent time-series spectra (right), illustrating the temporal and spectral differences between echoes associated with each object.

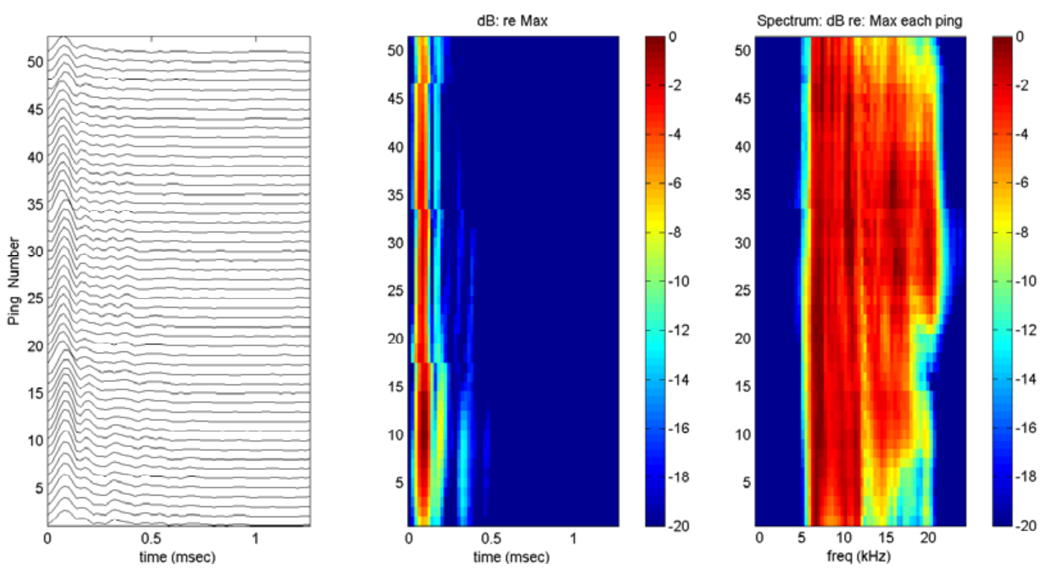

Fig. 10. Mine-like Shape Forward Pass, Transmit Angle 8^o : Echo envelope water fall and raster plots (left and center) and associated time series spectra (right plot)

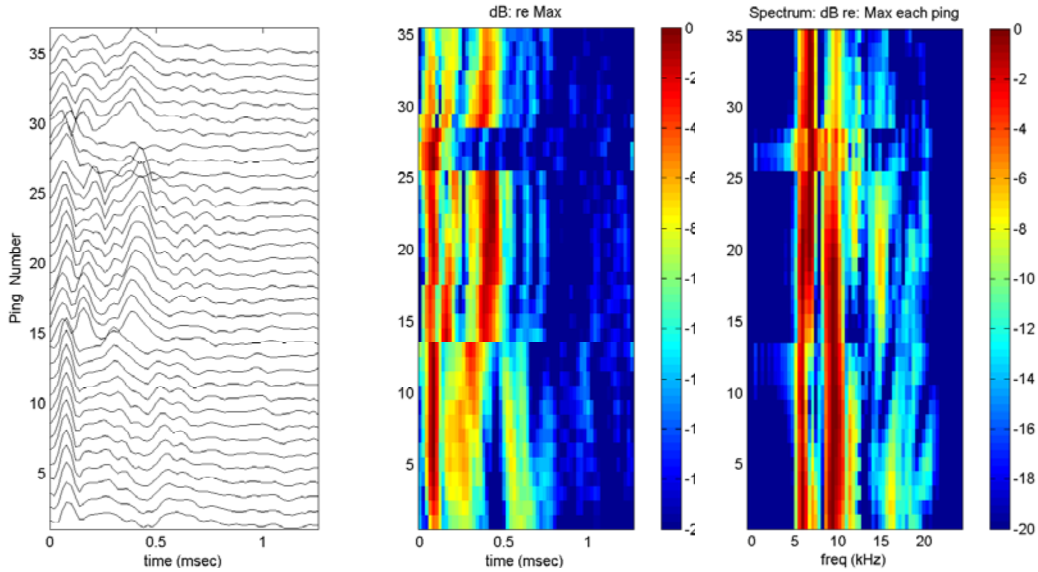

Fig. 11. Cement Block Forward Pass, Transmit Angle 8°: Echo envelope water fall and raster plots (left and center) and associated time series spectra (right plot)

B. Classifier

The Hybrid Time Frequency Processor, Multi-layer Perceptron Classifier [3,6], employs the windowed short-time fast Fourier transform (STFFTs) for joint time-frequency analysis. Spectrogram exemplars for each of the test field objects of Table I are shown in Fig. 12. Using linear principal component analysis (PCA) the resulting 2-D spectrogram is compressed into a small feature vector (Fig. 10), which retains a large proportion of the original energy. These projected features are subsequently used as inputs to an MLP artificial neural network (ANN), which is trained to map inputs into class declarations.

For 5-class tests, the classes are designated: (1) mine-like shape; (2) cement block, (3) aluminum sphere; (4) scuba tank; (5) ordnance shell. For 2-class tests, the mine-like shape is target class (1), and the remaining objects are assigned to the clutter class (2).

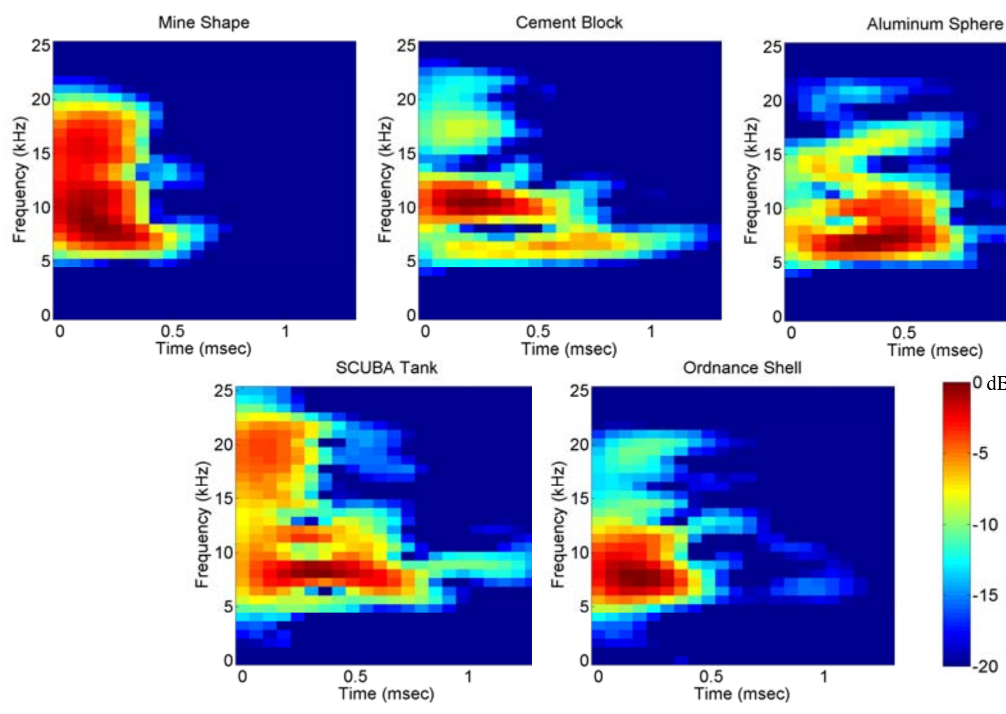

Fig. 12. Spectrogram exemplars for the test field objects of Table I

The classification system, a scaled conjugate gradient back-propagation neural network, maps the 15 principal components into 5 class membership output nodes (target value of 0.9 and non-target value of 0.1) via 15 nodes in a single hidden layer employing hyperbolic tangent sigmoid transfer functions. The output layer employs log sigmoid transfer functions.

C. 5-Class Test

The numbers of preprocessed focused echoes are 162, 95, 82, 95, and 68 for the respective objects. The number of pings used for training, cross-validation and testing are listed in Table IV. The signal classification results for this experiment are shown in the confusion matrix and captions of Figure 13. In this case the composite performance was 94%, however the signals for the mine-like shape were classified at a success rate of 83%. The classifier was more successfully trained to distinguish the mine-like shape from the clutter objects, as demonstrated in Sec. IV.D.

TABLE IV
CATALOGUE OF SIGNALS

CLASS	TRAIN	CV	TEST
1. Mine Shape	51	53	58
2. Cement Block	35	27	33
3. Alum. Sphere	17	32	33
4. Scuba Tank	34	28	33
5. Ordnance Shell	21	19	28

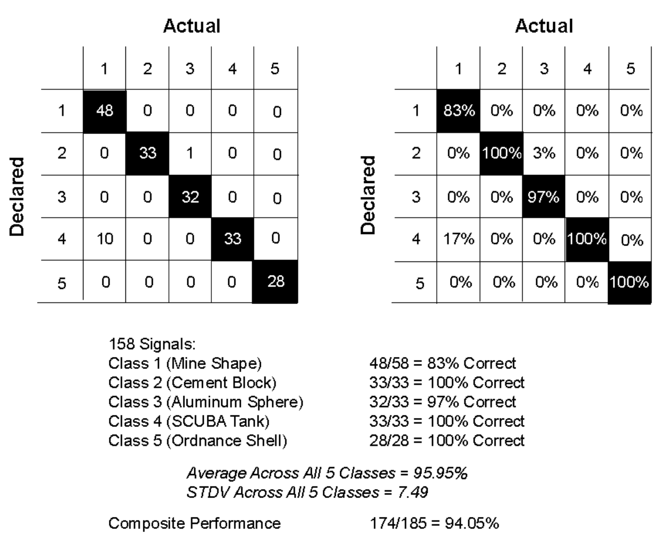

Figure 13. Confusion Matrix for 5-Class Signal Classification Experiment

D. 2-Class Test

For the 2-Class test, the MLP output node is threshold activated to declare the observed object as either mine-like or clutter. The numbers of echoes processed were 162 and 340 for target and clutter respectively. The classifier was trained to yield ROC curve knee characteristics of approximately: Pcc = 96% at Pfc = 4%, as shown in Figure 14.

E. Discussion

Based on the limited number of exemplars and test objects to date, the signal classification experiments described, utilizing different run orientations and transmit angles for training, cross validation and testing, illustrate the potential for time-frequency based signal classification. More realistic assessment will be made with data collected in sea tests scheduled for the coming year.

Several challenges to practical implementation are noteworthy. Logical and consistent alignment of signals is very important, and is complicated by energy scattered at the bottom interface and by neighboring objects. The threshold alignment technique utilized worked well for this data; however, we anticipate that the cluster image centroid, illustrated by the red dots in Fig. 7, will provide a more robust signal alignment index. Furthermore, we are investigating multi-aspect adaptive echo classification approaches that derive wavelet packet sub-band features from the complete time series [13,14], and which should be less sensitive to alignment-induced phase inconsistencies.

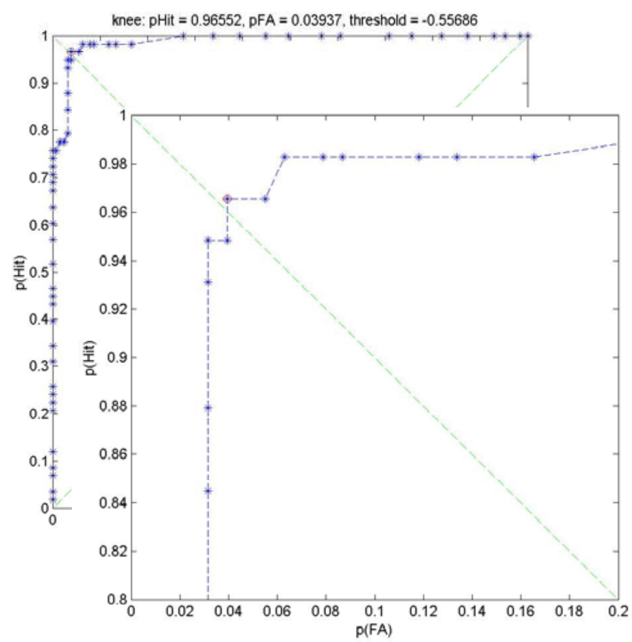

Fig. 14. ROC curve for 2-Class Signal Classification Experiment

V. TARGET CONFIRMATION ARCHITECTURE

The target confirmation architecture being developed for the Buried Object Scanning Sonar comprises the signal and image processing/classification paradigms described in this paper. The processing stream, illustrated in Fig. 15, receives vertical slice imagery and matched filtered echo time series. Per the methods described in Sections III and IV, object detection is achieved by along-track energy integration, and alerted regions in the vertical slice imagery are across-track clustered. This 1-D followed by 2-D approach is designed to adhere to the real-time requirements of the fielded system. The across-track clustering module is the branching point for image and signal classification, making possible the data fusion aspect of our approach.

The image classifier branch collects pixel lists that span the detected object. From this master list, image features are extracted, transformed and used for object classification. The signal classifier branch uses object centroid locations supplied by the clustering module to compute focused echo time series, which are subsequently used for object classification. With the collection of realistic field data for training purposes, fusing the outputs of the two classifiers, one based on spatial characteristics and the other on acoustic color, should provide high confidence confirmation of mine targets.

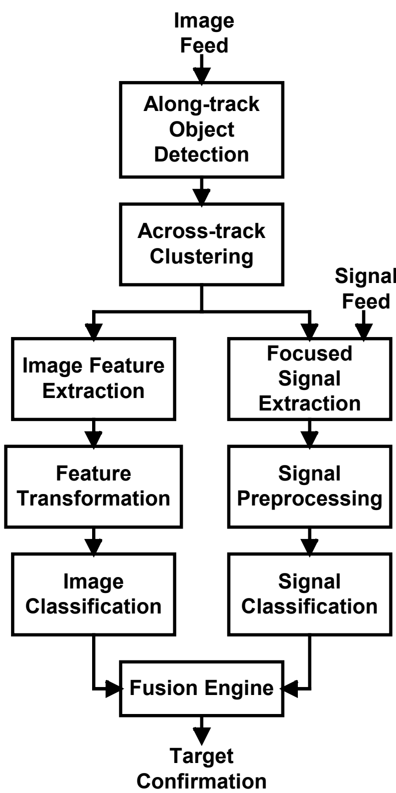

Fig. 15. Proposed Target Confirmation Architecture

VI. SUMMARY

The goal of this R&D effort is to develop image and signal classification methods for identifying mine targets in Buried Object Scanning Sonar (BOSS) data. Section III described synthetic 2-D image classification experiments that employ clustering and ellipse feature extraction methods, followed by a joint Gaussian Bayesian image classification algorithm. The clustering algorithm and an ellipsoid feature extractor were applied to 3-D BOSS imagery, illustrating the efficacy of the approach. Section IV described a signal classification experiment for data collected from mine-like and clutter objects buried in sand. These tests utilized data from different run orientations and transmit angles for training, cross validation and testing, achieved 5-class classification levels of 94% and 2-Class ROC curve knee values of Pcc = 96% and Pfc = 4%, and thus illustrate the buried mine-hunting potential for time-frequency based signal classification. A proposed target confirmation architecture fuses the outputs of the two classifiers; one based on spatial characteristics and the other on acoustic color, and should provide high confidence confirmation of buried mine targets.

References

[1] Schmidt, K. S. (2003), "U.S. Navy Enlists the Help of Marine Mammals to Secure Iraqi Ports," *U.S. Department of State, International Information Programs*. http://usinfo.state.gov/topical/pol/arms/03041801.htm

[2] Clem, T. R. (2002), "Sensor Technologies for Hunting Buried Sea Mines," MTS/IEEE Oceans 2002, pp. 452-460.

[3] Sternlicht, D. D., Lemonds, D. W., Dikeman, R. D., Ericksen, M. and Schock, S. G. (2001), "Detection and classification of buried objects by an Adaptive Acoustic Mine-hunting System," MTS/IEEE Oceans 2001, pp. 212-219.

[4] Sternlicht, D. D., Thompson, A. W., Lemonds, D. W., Dikeman, R. D., Korporaal, M. T. (2002), "Image and Signal Classification for a Buried Object Scanning Sonar," Proceedings MTS/IEEE Oceans 2002, pp. 485-490.

[5] Schock, S. G., Tellier A., Wulf J., Sara, J., Ericksen M.,(2001) "Buried Object Scanning Sonar", *IEEE Journal of Oceanic Engineering*, Vol. 26, No. 4, pp. 677-689.

[6] Okimoto, G. and Lemonds D. (1999) "Principal component analysis in the wavelet domain: New features for underwater object recognition," in *Detection and Remediation Technologies for Mines and Minelike Targets*, *Proceedings of SPIE*, Vol. 3710, pp. 697-707.

[7] Schock, S. G., Wulf, J. (2002) "Sonar for Multi-aspect Buried Mine Imaging," MTS/IEEE Oceans 2002, pp. 479-484.

[8] Sea Engineering Inc. (2000) "Development of a 3-D, Forward/Aft Sweeping, High Resolution Buried Object Imaging System – Phase 2," Report prepared for CEROS.

[9] Godshalk, G.D. and Smith, J.A. (1997), "Evolution of the Low Frequency Active Information Processor (U)", *Navy Journal of Underwater Acoustics*, 47 (1), pp. 155-190, Classification Secret

[10] Dobeck, G. J., Hyland, J. C. and Le'Derick, S. (1997) "Automated Detection/Classification of Sea Mines in Sonar Imagery," in *Detection and Remediation Technologies for Mines and Minelike Targets*, *Proceedings of SPIE*, Vol. 3079, pp. 90-108.

[11] Sternlicht, D. D. and de Moustier, C. P. "Stacking and averaging techniques for bottom echo characterization," *Proceedings: 135th Meeting of the Acoustical Society of America*, pp. 3023-3024, June, 1998.

[12] Sternlicht, D. D. and de Moustier, C. P. "Remote sensing of sediment characteristics by optimized echo-envelope matching," *Journal of the Acoustical Society of America*, In Press.

[13] Azami-Sadjadi, M. R., Jamshidi, A. A. and Dobeck, G. J. (2001) "Adaptive underwater target classification with multi-aspect decision feedback," in *Detection and Remediation Technologies for Mines and Minelike Targets*, *Proceedings of SPIE*, Vol. 4394, 1173-1182.

[14] Robinson, M., Azimi-Sadjadi, M. R., Sternlicht, D. D., Lemonds, D. (2003) "Multi-aspect acoustic classification of buried objects," To appear in MTS/IEEE Oceans 2003.