FINAL REPORT

Multichannel Detection and Acoustic Color-Based Classification of Underwater UXO in Sonar

SERDP Project MR-2416

JUNE 2015

M. R. Azimi-Sadjadi
Colorado State University

Steven Kargl
University of Washington

Distribution Statement A
This document has been cleared for public release

SERDP
DOD ▪ EPA ▪ DOE

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
08/31/2015

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
March 2014 - June 2015

4. TITLE AND SUBTITLE
Multichannel Detection and Acoustic Color-Based Classification of Underwater UXO in Sonar

5a. CONTRACT NUMBER
W912HQ-14-C-0007

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. M. R. Azimi-Sadjadi

Dr. Steven G. Kargl

5d. PROJECT NUMBER
MR-2416

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Electrical and Computer Engineering Colorado State University Fort Collins, CO 80523

Applied Physics Laboratory University of Washington Seattle, WA 98105

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Strategic Environmental Research and Development Program (SERDP)
4800 Mark Center Drive
Suite 17D08
Alexandria VA 22350-3600

10. SPONSOR/MONITOR'S ACRONYM(S)
SERDP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES
N/A

14. ABSTRACT
The Department of Defense (DoD) is currently responsible for clearing many sites that are potentially contaminated with munitions as a result of past training and weapons testing activities. In many cases, these activities occurred near or were performed in shallow water environments where munitions pose threats to public safety and the environment. The objective of this SERDP Exploratory Development (SEED) project is to develop efficient signal processing techniques for the detection and classification of military munitions in shallow underwater environments using data collected from synthetic aperture sonar (SAS) systems.

15. SUBJECT TERMS
Underwater military munitions, detection and classification, signal processing, synthetic aperture sonar (SAS) data.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. M. R. Azimi-Sadjadi
U	U	U	SAR	32	19b. TELEPHONE NUMBER *(include area code)* 970-491-7956

Contents

1 Abstract **4**

2 Objective **4**

3 Background **5**

4 Materials and Methods **6**

4.1 Detection Using Broadband Coherence 6

4.2 Scattering Models Utilized in Classifier Development 7

4.3 AC Data Generation 9

4.4 Classification Using Matched Subspace Classifier 11

4.4.1 Signal Subspace Generation using K-SVD Method 12

4.4.2 Locality Preserving Extension of K-SVD 13

5 Results and Discussion **14**

5.1 Broadband Coherence Detector Results 14

5.1.1 Formation of the Test Statistic 14

5.1.2 Detection Results 16

5.2 MSC Classifier Results 17

5.2.1 Description of Training and Testing Datasets 17

5.2.2 Classification Results and Analysis 20

6 Conclusions and Implications for Future Research/Implementation **23**

6.1 Conclusions and Discussions 23

6.2 Possible Future Research/Implementation 24

List of Acronyms

Acronym	Meaning
AC	Acoustic Color
APL-UW	Applied Physics Laboratory, University of Washington
ATR	Automatic Target Recognition
CCA	Canonical Correlation Analysis
CMAC	Collaborative Multi-Aspect Classifier
DoD	Department of Defense
FE	Finite Element
FFT	Fast Fourier Transform
HMM	Hidden Markov Models
GLRT	Generalized Likelihood Ratio Test
IFFT	Inverse Fast Fourier Transform
LFM	Linear Frequency Modulated
LP-KSVD	Locality Preserving K-Means Singular Value Decomposition
LSAS	Linear-Path Synthetic Aperture Sonar
MSC	Matched Subspace Classifier
MSE	Minimum Squared Error
NSWC	Naval Surface Warfare Center
OMP	Orthogonal Matching Pursuit
PondEX10	Pond Experiment 2010 (APL-UW/NSWC)
ROC	Receiver Operating Characteristic
ROI	Region of Interest
SAS	Synthetic Aperture Sonar
TREX13	Target and Reverberation Experiment 2013 (APL-UW)
UXO	Un-exploded Ordnance

Keywords

Broadband Coherence Test, Multichannel Signal Detection, Cross-Spectral Matrix, Generalized Likelihood Ratio Test, Fast Ray Models, Matched Subspace Classifier, Sparse Coding, Synthetic Aperture Sonar, Unexploded Ordinance.

Acknowledgments

The investigators would like to thank NSWC-Panama City for support and providing the data used in this study.

1 Abstract

The objective of this work is the development of efficient signal processing techniques for the detection and classification of military munitions in shallow underwater environments using data collected from synthetic aperture sonar (SAS) systems. In this final report we first address the problem of detecting the presence of underwater munitions using the multichannel coherence analysis framework. Our detection hypothesis is that the presence of munitions in the sonar backscatter collected from a hydrophone array will lead to higher levels of coherence compared to the backscatter from the seafloor alone. This method has been found to produce excellent detection performance on other sonar datasets. Here, detection results are presented on a sonar dataset which was collected in a relatively controlled and clutter-free environment. Results are presented using standard performance metrics such as probability of detection (P_d), probability of false alarm (P_{fa}), and Receiver Operating Characteristic (ROC) curve characteristics.

The goal of the second part of this work is to develop a robust target classification method that can be applied to the detected contacts to discriminate munitions from non-hostile man made objects and competing clutter. This method is developed based upon the Matched Subspace Classifier (MSC) using multidimensional Acoustic Color (AC) data extracted from the raw sonar returns. Scattering models developed by APL-UW were acquired to generate the required training dataset for various UXO and non-UXO objects. This was done owing to the fact that actual sonar data from a wide range of UXO and non-UXO objects is scarce in realistic situations. Although, it may be somewhat ambitious to expect model data capture all the essential features of these objects for target characterization, it will provide us with clues on how to augment the training datasets using perhaps a limited training samples from sonar returns of actual objects to improve the robustness in different environmental conditions. Our classification hypothesis is that spectral content of the sonar backscatter display unique acoustic signatures providing excellent discrimination between different classes of detected objects. Classification results of the MSC using three different signal subspace construction methods are presented on three real sonar datasets. The first two sonar datasets, PondEX09 and PondEX10, were collected for various underwater objects using a rail system in a pond under relatively controlled and clutter-free conditions. The third dataset, TREX13, was also collected using the same rail system but in the bay area off of the Panama City coast where other factors such as school of fish, water turbulence, seafloor roughness, and target range were more realistic. Results are presented using standard performance metrics such as probability of correct classification (P_{CC}), probability of false alarm (P_{FA}), ROC curve, and confusion matrix of the classifier.

2 Objective

The objective of this work is the development of automatic target recognition (ATR) algorithms for the detection and classification of military munitions in shallow underwater environments using data collected from SAS systems. Specifically, one of the technical questions that will be answered in this work is whether or not multichannel detection [1] using the Generalized Likelihood Ratio Test (GLRT) can detect munitions with sufficiently high probability of detection and localization while at the same time providing high false alarm rejection capability. For this problem, the hypothesis is that the presence of munitions in the sonar backscatter collected from a hydrophone array will lead to higher levels of coherence compared to the backscatter from the seafloor alone. This increase in coherence can give one an indication of which areas of the field may contain possible munitions and subsequent classification and further analysis may then be conducted in those areas. A new version of this detector which looks for high coherence in two frequency bands of the received data was developed and tested in this study. As the detector must be applied to the entire target field, it is important that the developed methods are computationally efficient as well as providing high

detection rates at a sufficiently low false alarm rate. In this report, we present detection results of the new version of the multichannel detection technique when applied to the PondEX10 dataset consisting of sonar backscatter from multiple munitions.

The specific objective of the second part of our work is the development of efficient methods for the classification of military munitions in shallow underwater environments using data collected from SAS systems. Specifically, the technical question that is addressed in this work is whether or not the Matched Subspace Classifier (MSC) [2] can successfully be trained on model-based sonar data of various UXO-like objects and then be applied to real sonar datasets to discriminate munitions with sufficient accuracy from other natural or man-made competing objects. The motivation here is related to the fact that collecting real UXO data in realistic settings is a difficult, if not impossible, task. Thus, by using physical models to construct a signal subspace spanned by the acoustic response of a particular target over a range of aspect orientations, our hypothesis is that the matched subspace classifier in [2] can provide an effective classification method that remains robust to changes in target aspect. Moreover, by keying in on specific target responses, the classifier will correspondingly exhibit low false alarm rates. The development of systems that can be trained on model-based data with guaranteed performance on real data could provide a significant contribution toward solving this difficult problem. Throughout the course of this research, proof-of-concept of the proposed methods will be provided by applying the trained classifier to one or more representative real datasets for evaluation purposes. These include, but are not limited to, multiple datasets from the PondEX09 and PondEX10 exercises as well as a more recently acquired dataset referred to as TREX13.

The comprehensive testing of the MSC and its application to the classification of munitions using SAS systems is another primary objective of this work. To this end, we test the hypothesis that the spectral features captured in the AC data, extracted from the sonar backscatter from various objects, display unique features providing excellent discrimination between different classes of detected objects. In this final report, we will present new classification results of the MSC when trained via K-SVD [3] and LP-KSVD [4] methods on synthetic AC data that has been generated via a fast ray model. These classifier systems are then applied to the entire PondEX and TREX13 datasets to test the generalization ability of the classifier.

3 Background

The Department of Defense (DoD) is currently responsible for clearing many sites which are potentially contaminated with munitions as a result of past training and weapons testing activities. In many cases, these activities occurred near or were performed in shallow water environments where munitions pose serious threats to public safety and the environment. The ultimate goal of this work is to provide the DoD with improved detection and characterization techniques as it strives to find safer and more cost-effective technologies for underwater munitions remediation. This seed research responds to SERDP SON MRSEED-14-01 in *Wide Area and Detailed Surveys* for rapid and highly efficient detection and classification of underwater UXOs found in contaminated sites.

Various methods have also been developed for modeling the acoustic response of objects with geometries typically observed in mine and UXO-hunting applications and using this information for the purposes of classification. In [5], the authors considered SAS imaging of simple targets by combining models for reverberation, acoustic penetration, and target scattering into a unified model. This is then used to generate pings suitable for SAS simulations over a range of environmental and experimental conditions. Experimentally measured target scattering from proud and buried targets are then used to validate the model through several simulations. In [6], the authors analyze experimental results from a SAS data set collected in a fresh water pond. These measurements were conducted to investigate discrimination capabilities based on the acoustic response of targets for

underwater UXO applications. Results from this study show that it is possible to use the acoustic template as a fingerprint to uniquely identify a given target. In [7], it was further shown that these acoustic-color (AC) features are useful for discriminating similarly shaped targets. In [8] and [9], the PI and his team developed new coherent-based feature extraction and SAS-like acoustic color for detection and classification of underwater mines and UXO objects using the Canonical Coordinate Analysis (CCA) framework. New feature and decision-level fusion algorithms were also developed in [10] and [11] using a hidden Markov model (HMM) and a Collaborative Multi-Aspect Classifier (CMAC) to improve classification of mine-like objects while reducing the false alarms when multi-pings/aspects sonar data are available.

4 Materials and Methods

In this section, we first give a brief review of the theory that motivates the detection algorithm studied in this project. More specifically, we will begin by reviewing the theory of multichannel detection using broadband coherence [1], [12]. A discussion on the fast ray model [13], [14] used in generating synthetic AC and SAS datasets as well as a review of the theory that motivates the classification algorithm studied in this project are then provided. This includes the general theory of the MSC classifier [2] as well as the two subspace learning methods, K-SVD and LP-KSVD, which were utilized in training the signal subspaces for the MSC.

4.1 Detection Using Broadband Coherence

In this section, we give a brief review of the theory of multichannel coherence [1], [12] and discuss the technical approach used to apply this theory for the detection of underwater munitions. Consider a set of L random data matrices $\{X_i\}_{i=1}^{L}$ with each matrix

$$X_i = \begin{bmatrix} x_i[0,0] & x_i[0,1] & \cdots & x_i[0,N-1] \\ x_i[1,0] & x_i[1,1] & \cdots & x_i[1,N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x_i[M-1,0] & x_i[M-1,1] & \cdots & x_i[M-1,N-1] \end{bmatrix} \in \mathbb{C}^{M\times N} \tag{1}$$

representing a 2-dimensional, zero-mean random process captured at sonar platform i. In this particular application, the random variable $x_i[m,n]$ represents the nth temporal sample collected from the mth hydrophone element in an array from the ith frequency band. Stacking the columns of X_i to form the vector $\mathbf{x}_i = \text{vec}\,(X_i)$, the composite vector $\mathbf{z} = \left[\mathbf{x}_1^T \ \cdots \ \mathbf{x}_L^T\right]^T$ has covariance matrix

$$R = E\left[\mathbf{z}\mathbf{z}^H\right] = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1L} \\ R_{12}^H & R_{22} & \cdots & R_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ R_{1L}^H & R_{2L}^H & \cdots & R_{LL} \end{bmatrix} \in \mathbb{C}^{LMN\times LMN}$$

with $R_{ik} = R_{ki}^H = E\left[\mathbf{x}_i\mathbf{x}_k^H\right] \in \mathbb{C}^{MN\times MN}$. This matrix captures all space-time second-order information within and between the random vectors $\{\mathbf{x}_i\}_{i=1}^{L}$.

We now assume we are given an experiment producing P *iid* realizations $\{\mathbf{x}_i[p]\}_{p=1}^{P}$ of the random vector associated with the ith frequency band. In the contexts of this problem, these P independent copies represent multiple pings collected from each frequency band. All P realizations may then be used to form the data matrix

$$\mathcal{Z} = [\mathbf{z}[1] \ \cdots \ \mathbf{z}[P]] = \begin{bmatrix} \mathbf{x}_1[1] & \cdots & \mathbf{x}_1[P] \\ \vdots & \ddots & \vdots \\ \mathbf{x}_L[1] & \cdots & \mathbf{x}_L[P] \end{bmatrix} \in \mathbb{C}^{LNM\times P} \tag{2}$$

where $\mathbf{z}[p] = \left[\mathbf{x}_1^T[p] \ \cdots \mathbf{x}_L^T[p]\right]^T$. Using the Generalized Likelihood Ratio Test (GLRT) to test the null hypothesis that $R = \text{blkdiag}\{R_{11}, \ldots, R_{LL}\}$ involves forming an estimate of the composite covariance matrix

$$\hat{R} = \frac{1}{P}\mathcal{Z}\mathcal{Z}^H = \frac{1}{P}\sum_{p=1}^{P}\mathbf{z}[p]\mathbf{z}^H[p] = \begin{bmatrix} \hat{R}_{11} & \hat{R}_{12} & \cdots & \hat{R}_{1L} \\ \hat{R}_{12}^H & \hat{R}_{22} & \cdots & \hat{R}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{1L}^H & \hat{R}_{2L}^H & \cdots & \hat{R}_{LL} \end{bmatrix}$$

and computing the following likelihood ratio [1]

$$\Lambda = \frac{\det \hat{R}}{\det \hat{D}} = \frac{\det \hat{R}}{\prod_{i=1}^{L} \det \hat{R}_{ii}} \tag{3}$$

Finally, assuming that all L data sequences are jointly 2-dimensional WSS and using results on determinants of asymptotically large block-Toeplitz matrices, similar arguments as that given in [12] shows that, as M, $N \to \infty$, the likelihood ratio given in (3) can be extended to the 2-dimensional frequency domain

$$\Lambda^{\frac{1}{MN}} \to \exp\left\{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \ln \frac{\det \hat{S}(e^{j\theta}, e^{j\phi})}{\prod_{i=1}^{L} \hat{S}_{ii}(e^{j\theta}, e^{j\phi})} \frac{d\theta\, d\phi}{4\pi^2}\right\} \tag{4}$$

with $\hat{S}(e^{j\theta}, e^{j\phi})$, $-\pi < \theta \leq \pi$, $-\pi < \phi \leq \pi$

$$\hat{S}(e^{j\theta}, e^{j\phi}) = \begin{bmatrix} \hat{S}_{11}(e^{j\theta}, e^{j\phi}) & \hat{S}_{12}(e^{j\theta}, e^{j\phi}) & \cdots & \hat{S}_{1L}(e^{j\theta}, e^{j\phi}) \\ \hat{S}_{12}^*(e^{j\theta}, e^{j\phi}) & \hat{S}_{22}(e^{j\theta}, e^{j\phi}) & \cdots & \hat{S}_{2L}(e^{j\theta}, e^{j\phi}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{S}_{1L}^*(e^{j\theta}, e^{j\phi}) & \hat{S}_{2L}^*(e^{j\theta}, e^{j\phi}) & \cdots & \hat{S}_{LL}(e^{j\theta}, e^{j\phi}) \end{bmatrix} \in \mathbb{C}^{L\times L} \tag{5}$$

an estimated composite power spectral density matrix. Here, $\hat{S}_{ik}(e^{j\theta}, e^{j\phi})$ represents a quadratic estimate of the cross-power spectrum between channels i and k at frequency θ and wavenumber ϕ. Thus, the likelihood ratio becomes a frequency/wavenumber dependent Hadamard ratio integrated over the Nyquist band. This likelihood ratio can then be compared to a threshold (determined experimentally based upon some training data) to decide whether or not UXO's are present. Moreover, the test statistic given in (4) is computationally efficient as the estimated cross-spectral matrix in (5) can be computed using a Fast Fourier Transform (FFT) of the waveforms received by each channel. More specifically, if we let $X_i^{(p)}(e^{j\theta}, e^{j\phi}) \in \mathbb{C}$ represent the two dimensional FFT of the pth realization of the data matrix given in (1) at frequency θ and wavenumber ϕ, then the cross-power spectrum may be computed as

$$\hat{S}_{ik}(e^{j\theta}, e^{j\phi}) = \frac{1}{P}\sum_{p=1}^{P} X_i^{(p)}(e^{j\theta}, e^{j\phi}) \left(X_k^{(p)}(e^{j\theta}, e^{j\phi})\right)^* \tag{6}$$

4.2 Scattering Models Utilized in Classifier Development

In order to tackle the task of constructing template signals to reliably represent the various UXOs in our MSC system, the work of the Co-PI on modeling scattering from objects at a water-sediment interface has been utilized [13]-[15]. The scattering model allows for monostatic SAS data sets to be simulated via a fast ray model that combines an acoustic ray approximation for propagation in a fluid-filled halfspace with scattering from a target in a number of conditions and media. This fast modeling is beneficial for generating large datasets for the MSC signal subspace construction. In this section we will discuss this scattering model as well as its utility in our classifier's development.

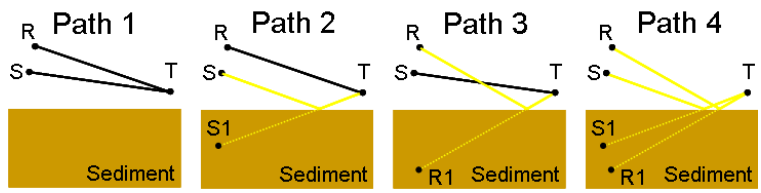


Figure 1: The path 1 is a direct path. The paths 2 and 3 interact with the sediment once and scatter from the target in a bistatic direction. Path 4 is a back-scattering path with two bottom interactions.

Under typical operation for a short-range SAS platform, air-water scattering paths can be ignored, because paths that interact with the air-water interface are either removed by time-gating the received signals or are naturally suppressed by the directivity of the source and receiver. In addition, the separation distance between the actual source and receiver is much smaller than the distance between the interface and the target, so the source and receiver can be considered to be co-located. Under these conditions, only the four ray paths that are shown in Figure 1, associated with the actual source and receiver and their images in the sediment, contribute to the scattered pressure. In this model, a source, receiver, and target are at locations $\mathbf{r}_s$, $\mathbf{r}_r$, and $\mathbf{r}_t$ respectively and a single image source is located at $\mathbf{r}_{s_i}$ with image receiver at $\mathbf{r}_{r_i}$. The source, receiver, and target are denoted by S, R, and T, respectively; while $S1$ and $R1$ in Fig. 1 are the image source and image receiver. To distinguish path 2 and path 3, the source and receiver are shown at distinct locations; and with our assumption of co-located source and receiver, paths 2 and 3 are reciprocal and paths 1 and 4 are back-scattered. With the specification of an image source and image receiver, the scattering from a target has been reduced to a superposition of 4 free-field scattering problems. Under operational conditions, the distance associated with each path satisfies $d \gg \lambda$ where λ is the wavelength of the pressure. The scattered steady-state pressure can then be written as

$$p_s = p_0 A\left(\mathbf{k}_i, \mathbf{k}_s, \omega\right) \frac{\exp(ikr)}{r} \tag{7}$$

where p_0 is the amplitude of the incident pressure, A is the scattering amplitude, $\mathbf{k}_i$ and $\mathbf{k}_s$ are unit vectors associated with the direction of the incident and scattered fields, ω is the angular frequency, r is the range from the target to a field point in a target-centered coordinate system, and $\exp(ikr)/r$ is a spherically diverging wave, respectively. The scattering amplitude contains all the information concerning the material properties of the target (e.g., density) and the directionality of the scattered field. A scattering amplitude can be determined from an analytic solution to a scattering problem (e.g., scattering from a spherical target), direct measurement from an actual target, or numerical simulation (e.g., a finite element (FE) model for a given target as was used in Figures 4(c)).

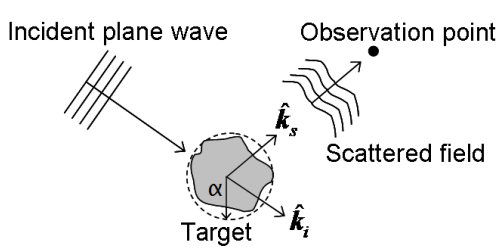


Figure 2: Free-field scattering assumes a portion of an incident wave is scattered to a distant observation. Here, α represents a characteristic dimension of the target.

Combining the ray model paradigm with free-field scattering as given in (7), the spectrum of

the total scattered pressure can be written as

$$P(\omega) = \left[\frac{A_1(\omega)}{d_1 d_2} e^{i\omega t_1} + \frac{V(\theta_g) A_2(\omega)}{d_2 d_3} e^{i\omega t_2} + \frac{V(\theta_g) A_3(\omega)}{d_1 d_4} e^{i\omega t_3} + \frac{V^2(\theta_g) A_4(\omega)}{d_3 d_4} e^{i\omega t_4} \right] r_0 P_{src}(\omega) \quad (8)$$

with $d_1 = |\mathbf{r}_s - \mathbf{r}_t|$, $d_2 = |\mathbf{r}_t - \mathbf{r}_r|$, $d_3 = |\mathbf{r}_{s_i} - \mathbf{r}_t|$, and $d_4 = |\mathbf{r}_t - \mathbf{r}_{r_i}|$. The time delays are then $t_1 = (d_1 + d_2)/c$, $t_2 = (d_2 + d_3)/c$, $t_3 = (d_1 + d_4)/c$, and $t_4 = (d_3 + d_4)/c$; with c being the speed of sound in water. The pressure spectrum $P_{src}(\omega)$ represents the frequency spectrum of the transmitted wave packet from the source, and r_0 is a reference distance associated with the calibration of the source (typically, $r_0 = 1$ m). The scattering amplitudes $A_k(\omega)$ in (8) depend on the locations of the sources, receivers, and target. Note the indices of A_k correspond to the path enumeration described before. The reflection coefficient at the water-sediment interface, which is represented by $V(\theta_g)$ and is a function of the grazing angle θ_g, is defined as follows

$$V(\theta_g) = \frac{\rho \sin(\theta_g) - (\kappa^2 - \cos^2(\theta_g))^{1/2}}{\rho \sin(\theta_g) + (\kappa^2 - \cos^2(\theta_g))^{1/2}} \quad (9)$$

where $\rho = \rho_2/\rho_1$ and $\kappa = (1 + \nu\delta)/\nu$ with $\nu = c_2/c_1$. Here c_1 and ρ_1 are the speed of sound and density for the water, and c_2, ρ_2, and δ are the speed of sound, density, and loss parameter for the sediment. An inverse Fourier transform of $P(\omega)$ thus gives a generated sonar signal that includes the four primary acoustic paths for a target near an interface.

4.3 AC Data Generation

In order to create synthetic AC data for classifier training, raw sonar returns generated by the above method were processed. Generation of AC data amounts to forming the intensity of the returned spectral power over the entire range of aspect angles that are modeled in a linear path SAS (LSAS) run. This is accomplished by the following procedure: First, Finite Element (FE) model [15] is implemented to produce scattering amplitude signal for an intended target or non-target objects. These scattering amplitudes are modeled for acoustic transmissions and returns in the frequency range of 0-31 kHz. Next, the half-space model including the four described ray paths in (8) is utilized to generate a raw sonar return dataset by generating the modeled returns of a target using the IFFT of (8) over a pre-generated coordinate set representing the various positions along a linear path making soundings. In this simulation model (as in the PondEX09-10 and TREX data) an LFM incident signal is used which provides spectral information of an object's response in the range of 0-31 kHz. Next, these raw soundings are matched filtered (pulse compressed) with the original transmit signal. Then the FFT is taken of these pulse compressed soundings. The new series is windowed to 0-31 kHz to remove the unused frequency portions and isolate the frequency range of interest. Finally, the aspect dimension of the AC is determined as a function of the path and target locations and the corresponding aspect of each transformed pulse-compressed sounding is plotted against the frequency axis with a cool-hot color mapping that represents spectral energy intensity for a given frequency and aspect. An example of AC data generated for a 21 m simulated LSAS run with a target at a 10 m range and with the source/receiver elevation 3.8 m above the water-sediment interface is given in Figure 3.

As can be seen from this figure, the observed aspect angles are limited to $\approx 86°$ corresponding to the limited aspect orientations that an observing sonar interface experiences in a 21 m linear path sounding an object at a range of 10 m. As an arbitrary trajectory of a SAS platform can be modeled with the fast ray model [13], the model was used to generate LSAS sonar data, with a target located at the center of the LSAS path, at a variety of ranges, with the sonar interface elevated 3.8 m above the sediment, corresponding to the PondEX and TREX data collection procedures. This was repeated for several objects present in the testing data, a list of which and their characteristics is given in Table 2.

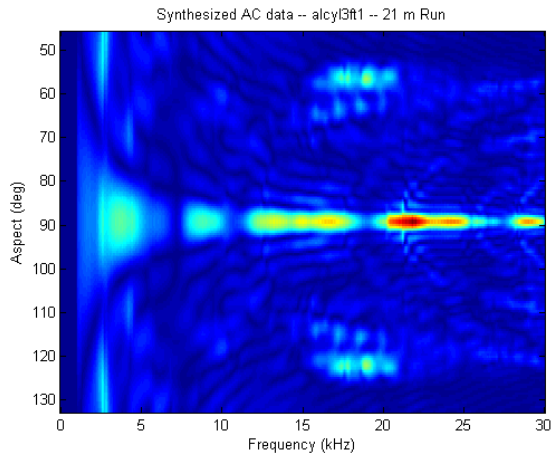


Figure 3: Acoustic Color data for 3 ft Aluminum Cylinder using LSAS Fast Ray Model

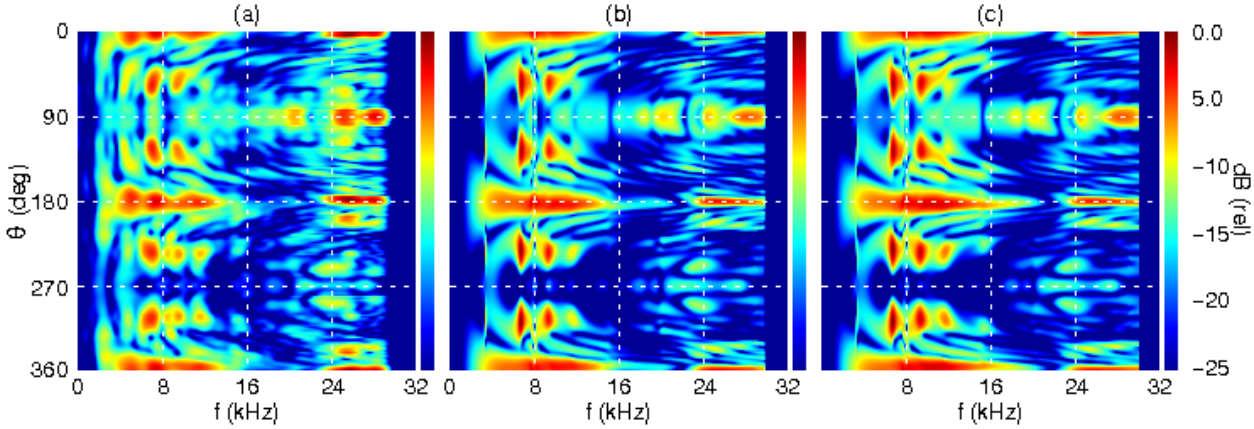


Figure 4: AC data of a bullet-shaped aluminum UXO replica at 10 m range generated via (a) Data collected during PondEX10 (SERDP MR-1665). (b) FEM and Kirchoff-Helmholtz integral result. (c) Fast ray model method with a scattering amplitude derived from the scattered pressure computed in (b).

AC data generated for template signals via the described scattering model was then utilized as training data for an implementation of our MSC classifier. The major benefit of utilizing the ray model developed in [13]-[15] is that, after a free-field scattering amplitude for a desired object is collected or modeled via FE methods, the regeneration of the ray model, simulating various aspects and orientations, is far simpler than re-running these variations with the same slow FE method [15]. A great deal of spectral information for object discernment seems to be present in these AC plots, but, as can be observed from Figure 4(c), the fast ray model generated AC's seem to preserve a great deal of the expected spectral information of a modeled object when compared to the AC data created from real sonar soundings of the same object in Figure 4(a).

The procedure for generating testing AC data for real raw sonar data from the PondEX10 experiment, such as that shown in Figure 4(a), is very similar to the procedure performed on the synthesized sonar data described above. A raw sonar data time series collected from the PondEX10 experiment is first matched filtered with a replica of the transmit signal that was used in the experiment and this new pulse compressed data is further filtered to remove returns from the neighboring objects from those of the object of interest. This filtering utilizes a reversible SAS imaging process, a spatial filtering process using a 2-D Tukey window [16], and a pseudo-inverse

filtering. This inverse filter maps the SAS image back to the pulse compressed version that now has less interference from neighboring objects. These filtered pulse compressed signals are then transformed to the frequency domain via FFT. This process is repeated for all aspects of LSAS and the power spectrum is then generated and plotted to display AC for each object in the real sonar data sets.

4.4 Classification Using Matched Subspace Classifier

In this section, we give a review of the theory behind the MSC. This classifier operates on the multidimensional AC data vectors described above. This is done with the assumption that such vectors, each belonging to a given class, can be formed using some linear combination of basis vectors associated with that class (i.e. signal observations take the form of the classical linear model). In our case, these basis vectors consist of the acoustic frequency responses (or AC) of an object at a given aspect angle. The MSC considered in this work is a special instance of the Weighted MSC discussed in [2], with identity weighting matrix. For signal subspace training, we use the standard singular value decomposition (SVD) [17], K-SVD [3] and LP-KSVD [4] signal subspace construction methods. The standard SVD is used as a benchmark as it is the suggested signal subspace choice for Matched Subspace Detectors/Classifiers [18]. The last two signal-specific dictionary learning methods are presented in Sections 4.4.1 and 4.4.2.

Our classification problem is referred to as an $M-$ary classification problem in which the following hypothesis is to be tested: Our observation contains image sources from M possible classes and one is uniquely most likely to be generating the signal. For a simple consideration of classifying 'UXO' vs. 'non-UXO', $M = 2$.

We will begin by defining the $M-$ary classification problem by considering $m = 0 \ldots M-1$ hypotheses each satisfying the signal model,

$$\mathbf{Y} = \mathbf{H}_m\mathbf{\Theta} + \mathbf{N} \ \in \mathbb{R}^{N\times Q}. \tag{10}$$

Here, our observation matrix is represented by $\mathbf{Y}$, a sum of the signal subspace component $\mathbf{H}_m\mathbf{\Theta}$, where $\mathbf{H}_m \in \mathbb{R}^{N\times K}$ is a matrix whose columns are basis vectors that span the subspace associated with the mth object class, an unknown parameter matrix $\mathbf{\Theta} \in \mathbb{R}^{K\times Q}$, and an additive zero-mean noise matrix $\mathbf{N} \in \mathbb{R}^{N\times Q}$. The distribution of the noise matrix is unknown, and hence it is not possible to derive the Maximum Likelihood Estimates (MLE) of the unknown parameters [2]. We can equivalently express (10) for a single observation vector $\mathbf{y}_q$ (or q^{th} column of $\mathbf{Y}$) as,

$$\mathbf{y}_q = \mathbf{H}_m\boldsymbol{\theta}_q + \mathbf{n}_q \tag{11}$$

where $\mathbf{n}_q$ is the qth column of matrix $\mathbf{N}$ and $\boldsymbol{\theta}_q$ is the qth column of $\mathbf{\Theta}$. The model-based AC subspaces $\mathbf{H}_m$ can be constructed for the mth object class utilizing different subspace reconstruction methods. Here, we used the SVD [17], K-SVD [3] and and LP-KSVD [4] dictionary construction methods.

The core idea behind the weighted MSC is the implicit suppression of large amplitude residuals by weighting each of the row residual terms in the discriminant function

$$J_m = \text{tr}\{(\mathbf{Y} - \mathbf{H}_m\mathbf{\Theta})^T\mathbf{W}(\mathbf{Y} - \mathbf{H}_m\mathbf{\Theta})\} = ||\mathbf{W}^{\frac{1}{2}}(\mathbf{Y} - \mathbf{H}_m\mathbf{\Theta})||_F^2, \ \forall m \in [1, ..., M] \tag{12}$$

where $\mathbf{W}$ is a diagonal matrix with weights along the diagonal, corresponding to row-weighting, and $||\mathbf{A}||_F^2$ represents the squared Frobenius norm of matrix $\mathbf{A}$ which is $||\mathbf{A}||_F^2 = \text{tr}\{\mathbf{A}\mathbf{A}^T\}$. For a given weight matrix $\mathbf{W}$, the weighted least-squares estimate of $\mathbf{\Theta}$ under the mth hypothesis is found using

$$\hat{\mathbf{\Theta}} = (\mathbf{H}_m^T\mathbf{W}\mathbf{H}_m)^{-1}\mathbf{H}_m^T\mathbf{W}\mathbf{Y}. \tag{13}$$

Using $\hat{\mathbf{\Theta}}$ in (13) we can get the value of the discriminant function in (12) at this solution i.e.

$$J_m = \text{tr}\{(\mathbf{Y}^T\mathbf{W}(\mathbf{I}-\mathbf{E}_m)\mathbf{Y})\} \tag{14}$$

where $\mathbf{E}_m$ is an oblique projection onto subspace $\langle\mathbf{H}_m\rangle$ i.e. $\mathbf{E}_m = \mathbf{H}_m(\mathbf{H}_m^T\mathbf{W}\mathbf{H}_m)^{-1}\mathbf{H}_m^T\mathbf{W}$. This discriminant function can be reinterpreted as the orthogonal projection of the weighted data matrix $\mathbf{Z} = \mathbf{W}^{1/2}\mathbf{Y}$ onto subspace $\langle\mathbf{W}^{1/2}\mathbf{H}_m\rangle$, or equivalently, $\mathbf{z}_q = \mathbf{W}^{1/2}\mathbf{y}_q$, for single vector observation $\mathbf{y}_q \in \mathbf{Y}$:

$$J_m = \text{tr}\{(\mathbf{Z}^T(\mathbf{I}-\mathbf{P}_{W^{1/2}H_m})\mathbf{Z})\} \tag{15}$$

where $\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{H}_m} = \mathbf{W}^{1/2}\mathbf{H}_m(\mathbf{H}_m^T\mathbf{W}\mathbf{H}_m)^{-1}(\mathbf{W}^{1/2}\mathbf{H}_m)^T$ is the projection matrix onto the subspace spanned by the columns of matrix $\mathbf{W}^{1/2}\mathbf{H}_m$. This weighted MSC classifier assigns a class label to an observation vector $\mathbf{y}_k$ based on the following criterion,

$$m^* = \arg\min_{m\in[1,M]} J_m = \arg\min_{m\in[1,M]} \{\mathbf{z}_q^T\left(\mathbf{I}-\mathbf{P}_{W^{\frac{1}{2}}H_m}\right)\mathbf{z}_q\}. \tag{16}$$

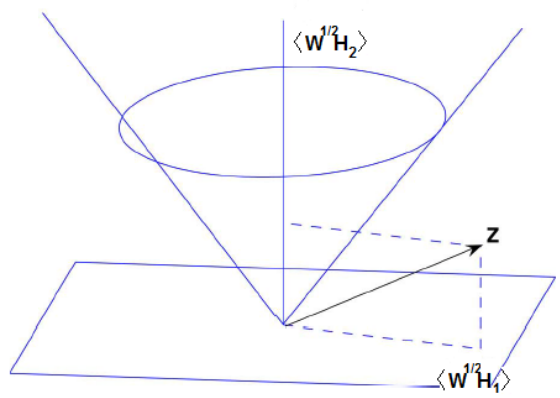


Figure 5: Weighted Matched Subspace Classifier geometric perspective [2]

As Figure 5 and equation (16) indicate, this classifier measures the energy in each of the subspaces $\langle\mathbf{W}^{1/2}\mathbf{H}_m\rangle$ and selects the class label corresponding to the subspace that contains the largest amount of energy. If the noise vector $\mathbf{n}_q$ is assumed to be normal with covariance matrix $E\left[\mathbf{n}_q\mathbf{n}_q^H\right] = \sigma^2\mathbf{I}$, then the criterion given in (16) also corresponds to the decision that minimizes the probability of assigning an erroneous class label. It is important to note that for $\mathbf{W} = \mathbf{I}$, this MSC reduces to the standard Matched Subspace Classifier/Detector [18], which builds our classification approach throughout this work.

4.4.1 Signal Subspace Generation using K-SVD Method

Using the fast-ray model provided by APL-UW, a large database of AC signals was created matching aspect resolution and frequency resolution of those generated for the real sonar data. Using this abundance of model signals, the K-SVD [3] subspace construction method was implemented forming various $\mathbf{H}_m$'s for the two classes. A brief description of this method is given in this subsection.

The purpose of K-SVD is to create an optimal signal-dependent dictionary that reduces the dimension of a signal vector by representing it as a sparse linear combination of relatively few atoms. More specifically, K-SVD aims to solve a constrained minimization problem to reduce the reconstruction error in a set of training vectors. Let $\mathbf{Y}_m \in \mathbb{R}^{N\times Q}$ be a matrix consisting of class m $(m = 0, 1)$ training data vectors $\mathbf{y}_q^{(m)}$ from Section 5.2.1.1 for $q \in [1\cdots Q]$ as its columns,

$\mathbf{H}_m \in \mathbb{R}^{N\times K}$ be the dictionary matrix to be found, and $\mathbf{X}_m \in \mathbb{R}^{K\times Q}$ be the sparse representation of $\mathbf{Y}_m$ in terms of its dictionary atoms. Note it is desired that the number of non-zero elements of each $\mathbf{x}_q^{(m)}$ be substantially less than N as the dimension should be reduced in this process. The constrained optimization problem [3] is given by,

$$\min_{\mathbf{H}_m,\mathbf{X}_m} \{\|\mathbf{Y}_m - \mathbf{H}_m\mathbf{X}_m\|_F^2\} \qquad \text{subject to ,} \quad \|\mathbf{x}_q^{(m)}\|_0 \leq \tau, \quad \forall q \tag{17}$$

where $||\cdot||_F^2$ is the Frobenius norm of a matrix, and $||\cdot||_0$ is the ℓ_0 norm which counts the non-zero elements of a vector.

During the training, the K-SVD algorithm is composed of two-phases. First, a sparse representation phase where for each $\mathbf{y}_q^{(m)}$ the corresponding $\mathbf{x}_q^{(m)}$ is computed based on a given $\mathbf{H}_m$ using a pursuit method such as Orthogonal Matching Pursuit (OMP) [19]. Second, a dictionary update phase where $\mathbf{h}_k \in \mathbf{H}_m$ is updated based on minimizing the reconstruction error using the SVD of a restricted error matrix $\mathbf{E}_k^R$. The steps of this procedure are outlined in Table 1. These two phases are repeated until convergence through monotonic MSE reduction, a more detailed description of which is given in [3].

Table 1: K-SVD Algorithm

K-SVD Optimal Dictionary Construction Algorithm:
Initialization: Set the dictionary matrix $\mathbf{H}_m^{(0)} \in \mathbb{R}^{N\times K}$ with K randomly selected l_2 normalized columns of $\mathbf{Y}_m$. Set $J = 1$. Repeat following steps until a stopping rule is met.
Sparse Coding Stage: Generate $\mathbf{X}_m$ by computing the sparse representation $\mathbf{x}_q^{(m)}$ for each $\mathbf{y}_q^{(m)}$ based on $\mathbf{H}_m$ using the Fast-OMP pursuit method [19].
Dictionary Update Stage: Each column $\mathbf{h}_k,\ k \in [1, ..., K]$, in $\mathbf{H}_m^{(J-1)}$ is updated by: 1. Compute $k-$exclusive error matrix $\mathbf{E}_k = \mathbf{Y}_m - \sum_{j\neq k} \mathbf{h}_j\mathbf{x}_{j*}$, where $\mathbf{x}_{j*}$ is the jth row of $\mathbf{X}_m$. 2. Define column indices of training data $\mathbf{Y}_m$ that use the kth atom in their reconstruction via $\mathbf{H}_m$: $\omega_k = \{i \mid 1 \leq i \leq Q,\ \ x_{k*}(i) \neq 0\}$. 3. Compute $\mathbf{E}_k^R$ and $\mathbf{x}_k^R$, the restricted error matrix and coefficient vector respectively, by selecting only columns of $\mathbf{E}_k$ corresponding to ω_k indices and likewise for entries of $\mathbf{x}_{k*}$ (i.e. discard zero entries in the row vector). 4. Apply SVD: $\mathbf{E}_k^R = \mathbf{USV}^T$. The updated dictionary column $\hat{\mathbf{h}}_k$ is the first column of $\mathbf{U}$ and the updated coefficient vector $\hat{\mathbf{x}}_k^R = \mathbf{v}_1\mathbf{S}(1,1)$, where $\mathbf{S}(1,1)$ is the first and largest singular value in the SVD of $\mathbf{E}_k^R$.
Set $J = J + 1$ and repeat until convergence.

4.4.2 Locality Preserving Extension of K-SVD

In Locality Preserving K-SVD or LP-KSVD method [4], the goal is to preserve the local geometry of a non-linear manifold while determining the optimal linear support $\mathbf{H}_m$. For our problem, it is assumed that a given AC data point $\mathbf{y}_q$ resides on a non-linear but smooth manifold describing all the AC data points. This method is an unsupervised method which attempts to learn discriminative atoms $\mathbf{h}_k$ by forcing their updates to be made solely from samples in their local neighborhood on the manifold. There are two main benefits to the LP-KSVD approach in comparison with the K-SVD [3]. First, the employed local coding has closed-form solution which makes it more

computationally efficient compared to sparsity driven algorithms such as the OMP used in the K-SVD method. Secondly, the dictionary is optimized for both its capability in representation and its locality preservation property, which is in contrast to sparse coding which only solves for a representational dictionary [4]. In this section, we will give a brief outline of the theory motivating the LP-KSVD method.

To begin, there are two main objectives of LP-KSVD. First, is establishing a compact dictionary $\mathbf{H}_m = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_K] \in \mathbb{R}^{N\times K}$ such that linear combinations of $\mathbf{h}_i$'s can approximate the nonlinear manifold $\mathcal{M} \subset \mathbb{R}^N$. Since we have no access to the true $\mathcal{M}$, $\mathbf{H}_m$ is estimated based on $\mathbf{Y}_m$ (here $\mathbf{Y}_m$ is the training samples for the mth class). The second objective is learning $\mathbf{h}_i$'s as landmark points, which are capable of preserving the locality on $\mathcal{M}$. This dictionary learning problem can be formulated as:

$$\min_{\mathbf{H}_m, \mathbf{X}_m} ||\mathbf{Y}_m - \mathbf{H}_m \mathbf{X}_m||_F^2 \tag{18}$$

$$s.t. \begin{cases} x_{ij} = 0 & \text{if } \mathbf{h}_i \notin \Omega_\tau(\mathbf{y}_j)\ \forall i, j \\ \mathbf{1}^T \mathbf{x}_j = 1 & \forall j \end{cases}$$

where the reconstruction error term measures the fitness of $\mathbf{H}_m$ to $\mathbf{Y}_m$ similar to K-SVD in (17); the matrix $\mathbf{X}_m \in \mathbb{R}^{K\times Q}$ contains Q local reconstruction codes, with $\mathbf{x}_j$ being the code for reconstructing $\mathbf{y}_j$ in terms of $\mathbf{H}_m$, and $\Omega_\tau(\mathbf{y}_j)$ denotes the neighborhood containing τ nearest dictionary atoms of $\mathbf{y}_j$ in terms of Euclidean distance (i.e. ℓ_2 norm). The first constraint dictates that every training sample $\mathbf{y}_j$ can only be re-constructed by its τ nearest-neighbor dictionary atoms. The second constraint allows the reconstruction coefficients to be invariant to translation of the data. The main goal is to choose $\mathbf{h}_i$ to be sufficiently close to $\mathcal{M}$. Moreover, to learn $\mathbf{h}_i$'s as landmark points, it is further required that each $\mathbf{h}_i$ be locally representative with respect to a small patch on the manifold $\mathcal{M}$.

The proposed LP-KSVD solves (18) iteratively by alternating between the two variables $\mathbf{H}_m$ and $\mathbf{X}_m$. This is done by first fixing $\mathbf{H}_m$ and solve for the best co-efficient matrix $\mathbf{X}_m$ and then, we update $\mathbf{H}_m$ as well as $\mathbf{X}_m$ together. Iterations are terminated if either the objective function value is minimized below some preset threshold or a maximum number of iterations has been reached. The steps for solving the local reconstruction codes and local dictionary optimization are described in details in [4].

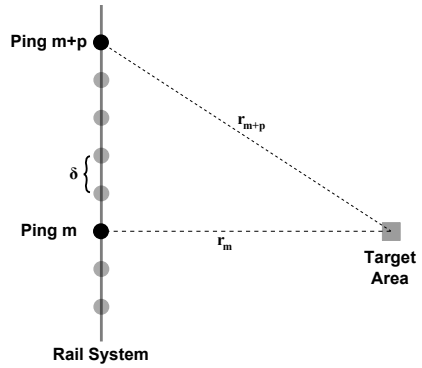
5 Results and Discussion

5.1 Broadband Coherence Detector Results

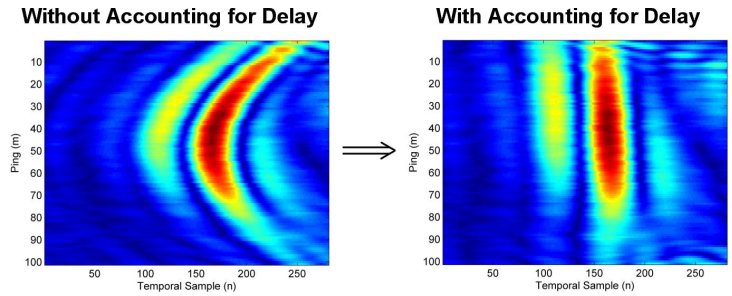
In this section, we provide results of the broadband coherence statistic given in (4) when applied to the PondEx10 dataset collected at the NSWC - Panama City, FL. Moreover, the results of this detector are compared and benchmarked to the Hadamard ratio given in (3) which does not rely on the wide sense stationary assumption when constructing the likelihood ratio. For a review of the dataset used in this section, the reader is referred to [20].

5.1.1 Formation of the Test Statistic

To form detection decisions, each sonar return at a particular ping within the run was first matched-filtered by correlating the received waveform with the LFM transmit signal. Each matched filtered signal was then partitioned into non-overlapping windows of length $N = 281$ (cross-track) corresponding to a range resolution of approximately 0.25m. For every 0.25m in the direction of the rail system (along-track), the time series collected over a $P = 100$ ping window were appropriately



(a) Geometry of Data Collection



(b) Effects of Delay on the Matched Filter Response

Figure 6: The matched filtered signals associated with each ping are delayed to account for increases in path length as the array moves along the rail.

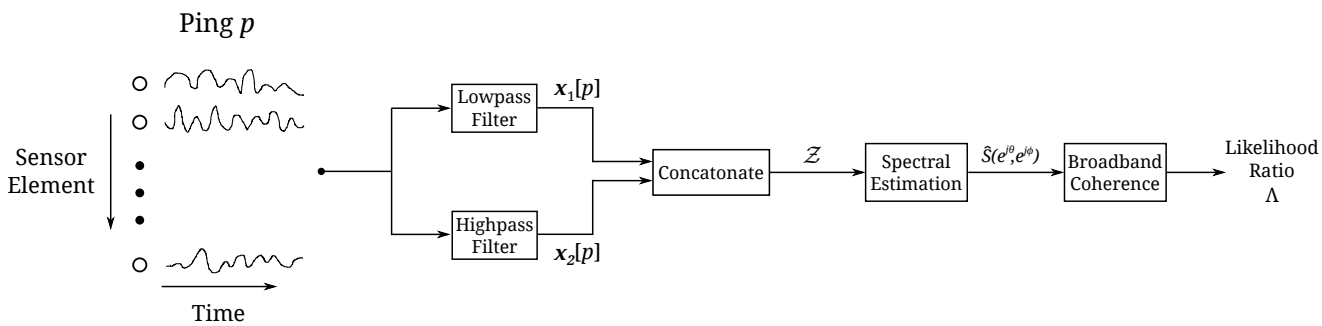


Figure 7: Processing of the data to form the likelihood ratio involves filtering the time series collected by each sensor element into two frequency bands, using all P pings to estimate the cross-spectral matrix $\hat{S}(e^{j\theta}, e^{j\phi})$, and finally computing Λ by integrating coherence over the Nyquist band.

lagged to account for the delay corresponding to the increase in path length as the array moves along the rail system. Figure 6 (a) gives the geometry of this procedure where, if the mth ping exhibits a range of r_m from the area of the target field that is of interest and there exists a ping separation of δ, then the range to the target area at ping $m+p$ is simply $r_{m+p} = \sqrt{r_m^2 + (p\delta)^2}$. This increase in path length can then be used to determine the relative delay between each ping and one may then account for this accordingly. Figure 6 (b) gives an example of the matched filtered response from a target both before and after accounting for relative time delays. After accounting for time delay, Figure 7 gives a block diagram of the subsequent processing steps used in the formation of the likelihood ratio for each 0.25×0.25 m location within the target field. As shown in this figure, the time series for each sensor element and for each ping is first filtered into $L = 2$ non-overlapping frequency bands using two 8th-order Butterworth filters, one a lowpass filter to remove high frequency content and the other a highpass filter to likewise remove low frequency content. This filtering is done independently for all $M = 5$ sensor elements in the array and the resulting time series are vectorized to form the vector $\mathbf{x}_i[p]$ for the pth ping. This is done for all $P = 100$ pings and the data from both frequency bands is concatenated to form the data matrix $\mathcal{Z}$ given in (2). A 2-dimensional Fast Fourier Transform (FFT) is then applied to the data from both frequency bands and the estimated cross-spectral matrix $\hat{S}(e^{j\theta}, e^{j\phi})$ given in (6) is computed. The likelihood ratio given in (4) is finally formed by numerically integrating the frequency/wavenumber Hadamard ratio using a trapezoidal rule approximation.

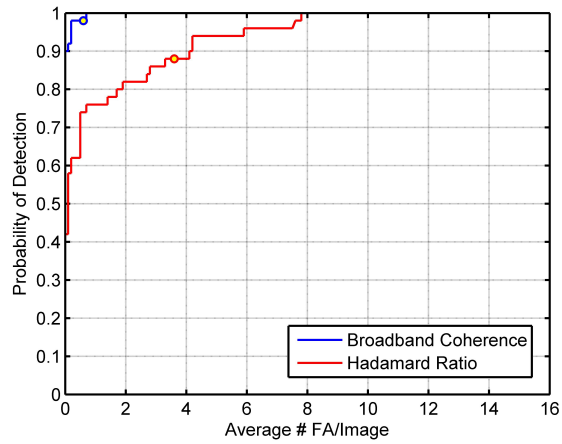
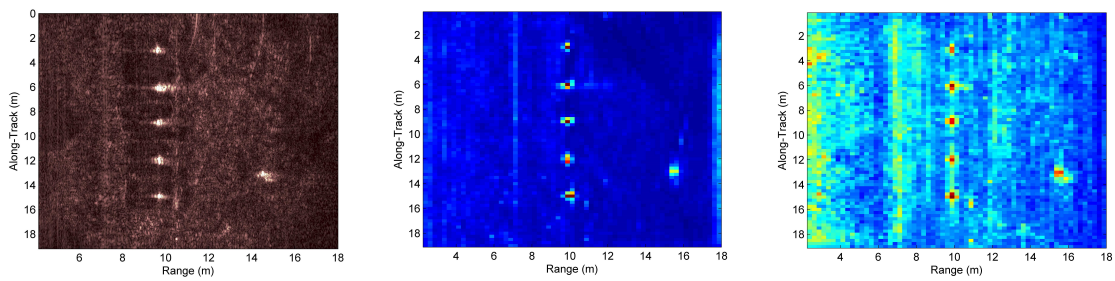


Figure 8: Receiver Operating Characteristic (ROC) curve for the PondEx10 dataset.



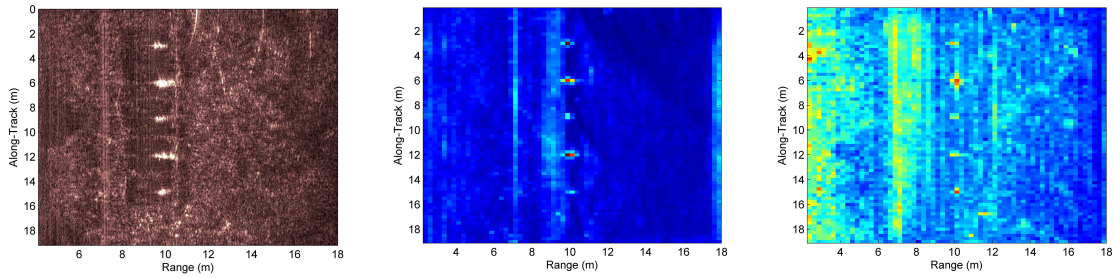
(a) SAS Image (b) Broadband Coherence (c) Hadamard Ratio

Figure 9: Beamformed SAS and likelihood ratio images for both detection methods at 0° object orientation.

5.1.2 Detection Results

The broadband coherence detector given in (4) was applied to all 10 runs of the PondEx10 dataset and compared to the Hadamard ratio given in (3). Using all 50 objects in the dataset (10 runs $\times$ 5 objects per run) as well as a randomly selected set of locations corresponding to background, Figure 8 displays the ROC curve for both detection methods. This curve is formed by varying the threshold of the detector over a range of values and, at each threshold, comparing the percentage of targets that are detected to the average number of background realizations per image that are mis-detected. From this figure it is clear that, by taking advantage of the WSS assumption, the broadband coherence detector outperforms the Hadamard ratio. One can also see that the broadband coherence detector performs well at discriminating UXO from background as the proposed technique detects all of the objects after only about an average of one false alarm per image. The broadband coherence detector also exhibits a knee-point (the point where $P_d + P_{fa} = 1$ which is shown by a small circle in Figure 8) at a probability of detection of $P_d = 99\%$. Compared to the results of the detector from the previous report [20], one can see that this new version of the broadband coherence detector exhibits a 3% improvement in P_d at the knee-point of the ROC curve over that of its previous version.

Figures 9 (a) and 10 (a) show the beamformed SAS images generated using k-omega beamforming for a run with 0° object orientation and one with 80° orientation, respectively. Note that, at 0° orientation the objects are broadside while a 90° orientation corresponds to situation where the nose of the objects are pointed toward the array. Figures 9 (b) and 10 (b) likewise show images of the broadband coherence statistic given in (4) computed at every 0.25×0.25 m location in the



(a) SAS Image (b) Broadband Coherence (c) Hadamard Ratio

Figure 10: Beamformed SAS and likelihood ratio images for both detection methods at 80° object orientation.

target field for this two runs. Likewise, Figures 9 (c) and 10 (c) show images of the likelihood ratio for the Hadamard ratio detector given in (3). From these images one can see that both methods are able to detect the objects with relative ease when at an orientation of 0° but have more difficulty when the objects are oriented at 80°. However, by comparing the likelihood ratios of the broadband coherence and Hadmard ratio detectors, one can once again visually observe that the broadband coherence detector does a better job of discriminating target from background.

5.2 MSC Classifier Results

In this subsection, we provide results of the MSC when tested on the PondEX09, PondEX10 and TREX13 datasets. First, brief reviews of the training and testing datasets will be provided. Following these discussions, we present the performance of the MSC when tested on PondEX09 and PondEX10 datasets using MSC with signal subspace components extracted via the regular SVD, K-SVD, and LP-KSVD methods. This is repeated for the TREX13 objects in two different range groupings.

5.2.1 Description of Training and Testing Datasets

5.2.1.1 Fast-Ray Model Generated Dataset – Training The Fast-Ray model can be utilized to create AC data for a variety of environments and simulated runs. Using the model and procedures described in Sections 4.2 and 4.3, raw sonar runs for 10 different objects used in the PondEX and TREX experiments were modeled. The modeled runs were designed to replicate the condition of the PondEX and TREX data. In particular, these runs were generated for a 21 m path length at horizontal ranges of 10 m and 30 m, with the source/receiver interface elevated 3.8 m above the sediment. These synthetic sonar datasets and their corresponding AC data were generated for two different environments. Both modeled environments simulated water sound speeds matching those that were used in the generation of the scattering coefficients in (8) which were generated via FE method. Similar to the testing data, object rotations from $-80°$ to $+80°$ in 20° increments, were generated to give AC data for 360° of aspect for each object and environment. This was then repeated for the two ranges we observed in TREX13 data collection, 10m and 30m. It must be mentioned that dictionaries were trained for each of the test sets separately, with FRM training data that matched the appropriate range and included only the objects viewed in that experiment.

For each object, the AC data generated was decimated along the frequency dimension to have $N = 310$ frequency bins spanning the $0 - 31$ kHz frequency band corresponding to approximately 100 Hz separation of frequency bins. After generating AC's at stops that one might encounter in a 21 m run, each training set was uniformly sub-sampled to a set containing $1/7^{\text{th}}$ of all possible aspects of our synthetic training data. Using these synthetic sonar datasets, signal subspaces were

Table 2: Fast Ray Model training data – experiment objects for which FRM exists

No.	Class	FRM Name	Object Description
1	non-UXO	alcyl2ft	2 ft Aluminum Cylinder
2	non-UXO	alcyl3ft	3 ft Aluminum Cylinder
3	non-UXO	alpipe	2 ft Aluminum Pipe Section
4	UXO	aluxo	100 mm Aluminum Rocket Round
5	UXO	ssuxo	100 mm Solid Steel Rocket Round
6	UXO	bullet_105mm_air	105 mm Bullet (Air Filled)
7	UXO	bullet_105mm_h2o	105 mm Bullet (H_2O Filled)
8	UXO	howitzer_cap_air	155 mm Howitzer with Cap (Air filled)
9	UXO	howitzer_cap_h2o	155 mm Howitzer with Cap (H_2O filled)
10	UXO	howitzer_nocap	155 mm Howitzer no Cap

trained using the methods described in Sections 4.4.1 and 4.4.2. For the trained dictionaries to be used in Pond testing, the training data was truncated to the frequency bins of the incident LFM signal power, spanning $1 - 30.5$ kHz, hence leading to AC vectors with $N_{Pond} = 295$ dimensions. Similarly, for the TREX dictionaries, since the incident LFM was in the $3 - 30.5$ kHz frequency range, the resulting truncated AC vectors were $N_{TREX} = 275$ dimensional. Dictionaries of sizes $\mathbf{H}_{Pond} \in \mathbb{R}^{295\times K}, \mathbf{H}_{TREX} \in \mathbb{R}^{275\times K}$ (for 10m and 30m ranges) were generated using sparsity factor $\tau = 10$ for both LP-KSVD and KSVD methods. Here $K = 400$ for non-UXO class and $K = 600$ for the UXO class. As was mentioned before, and in contrast to the SVD method, in K-SVD and LP-KSVD testing, the sparsity of reconstruction codes was enforced by applying OMP algorithm. The number of iterations used in training was chosen to be 30 as this was found to be acceptably close to minimum achievable reconstruction error. Table 2 enumerates and briefly describes each of the training data objects used. For the SVD subspace training, the top 30 non-UXO eigenvectors, associated with the 30 largest eigenvalues, of the non-UXO training set were kept whereas for the UXO training set, the top 50 eigenvectors were kept. Objects 1, 3, 4, and 5 were only observed in the PondEX experiments while all 10 objects were present in the TREX13 test set.

5.2.1.2 PondEX Datasets – Testing Figure 11 shows the layout of the PondEX09 and PondEX10 experiments including the relative locations of the rail-mounted sonar system and the placement of objects in the target field. The 21 m rail the sonar system is mounted on is fixed to eliminate platform motion as the sonar interface moves along its track. The sonar transmit signal is a 6 msec LFM over $0 - 31$ kHz with a 10% taper between the leading and trailing edges to minimize ringing in the transmitted signals. Sonar backscatter is received with $L = 6$ hydrophone elements that are arranged in a linear array normal to the seafloor.

As can be seen from Figure 11, the target field in PondEX10 contained five objects at a time, all of which were located approximately 10 m from the rail system and are proud on the sandy bottom. For each run that was made in the PondEX data collection, the orientation of the axes of symmetry for all the observed objects were the same relative to the fixed sonar rail. Nine total object orientations were used, ranging from $-80°$ to $+80°$ in $20°$ increments, where the run with $0°$ object orientation designates a configuration where an object's major axes of symmetry is parallel to the rail system. Each run of data consists of 769 pings in which the sonar platform moved along the fixed rail in increments of 0.025 m, transmitting and receiving once for each fixed position. The data was sampled at 1 MHz and the sonar platform was tilted at a fixed $20°$ grazing angle for all runs (angle of the sonar main response axis with respect to the horizontal plane).

Since the useful spectral information in the collected data has a Nyquist frequency well below

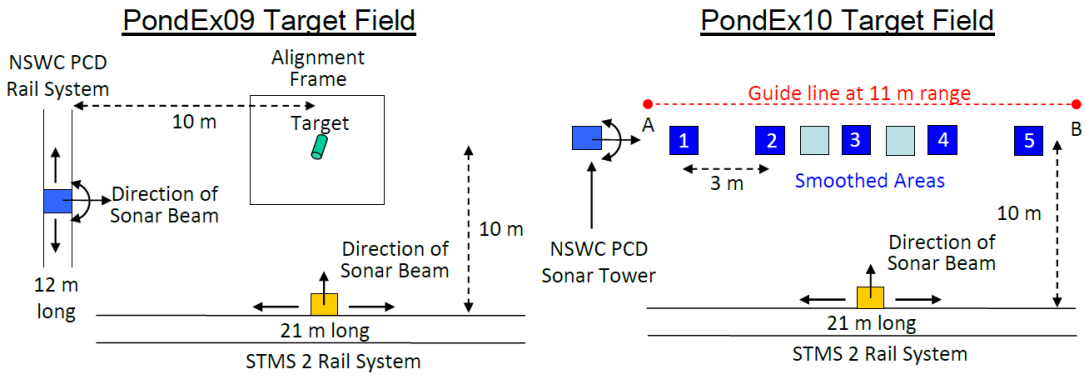


Figure 11: Layout of the target fields for PondEx09 and EX10 data collection [21]

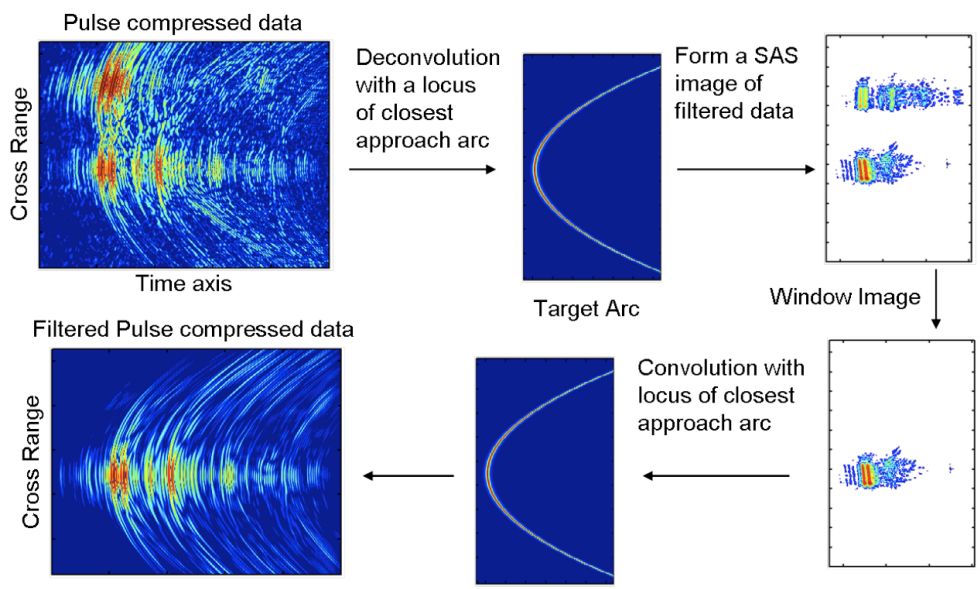


Figure 12: Inverse SAS filtering [14]

the sampled rate, the 1 MHz data was down-sampled by a factor of 10 resulting in effective sampling frequency of 100 kHz. Owing to the fact that the LFM incident signal for PondEX experiments has a range of frequency $1 - 30.5$ kHz, only those frequency bins were retained resulting in AC feature vector dimension of $N_{Pond} = 295$. In the PondEX experiments since the objects are laid out very close together, to generate AC data for each individual object free of the influence of the neighboring ones, a procedure using the Inverse-SAS filtering process [14] is utilized. This procedure is depicted in Figure 12 which shows this inverse-SAS filtering to remove neighboring object returns from the pulse-compressed data.

As this classifier is applied to the output of the previously developed broadband coherence-based detector [22], the detected regions of interest (ROI) that contained objects of interest were tested by the classifier. A hard-limiting power threshold was implemented in gathering test aspects from detected regions of interest. This threshold selects observations from the filtered runs by summing the squared frequency bin components in a single AC aspect and determining the total power of the observation. Those aspects that meet the hard-limiting threshold are kept. For too aggressive threshold, the top 5% of aspects will be chosen. That is, at least 5% of aspects in detected ROI's

Table 3: Objects in the Pond EX09-10 experiment testing data

FRM Name	# Runs/Aspects
alcyl2ft	10/3282
alpipe	10/1580
aluxo	10/315
ssuxo	10/717
Total # Tested:	5894

are picked for testing. One might note the difference in prevalence of class samples. This is also due to this power threshold. It was found that for the Aluminum Cylinder and Aluminum Pipe, many views of the object would return soundings above the set threshold that isolated the best UXO type soundings.

5.2.1.3 TREX13 Dataset – Testing The same rail system described in the PondEx experiments was utilized in the TREX13 data collection process. However, the environmental conditions were vastly different as the TREX13 dataset was collected in the bay area off of Panama City coast. The LFM incident signal used in the TREX13 dataset had power in the frequency range of $3 - 30.5$ kHz. Thus, using the same frequency sampling, the AC feature vector dimension for TREX13 becomes $N_{TREX} = 275$, i.e. truncating the original $N = 310$ frequency bins spanning $0 - 31$ kHz, to those that span $3 - 30.5$ kHz. Additionally, the same Inverse-SAS filtering, AC generation, and hard-limiting threshold processes described in Section 5.2.1.2 were implemented to obtain filtered data AC portions for each object in the TREX test set. Table 4 lists those objects and ranges for which Inverse-SAS filtered runs exist. Ranges 10m and 30m were used in our testing to further validate the performance of the classifier on data from a more realistic environment when compared to PondEX dataset.

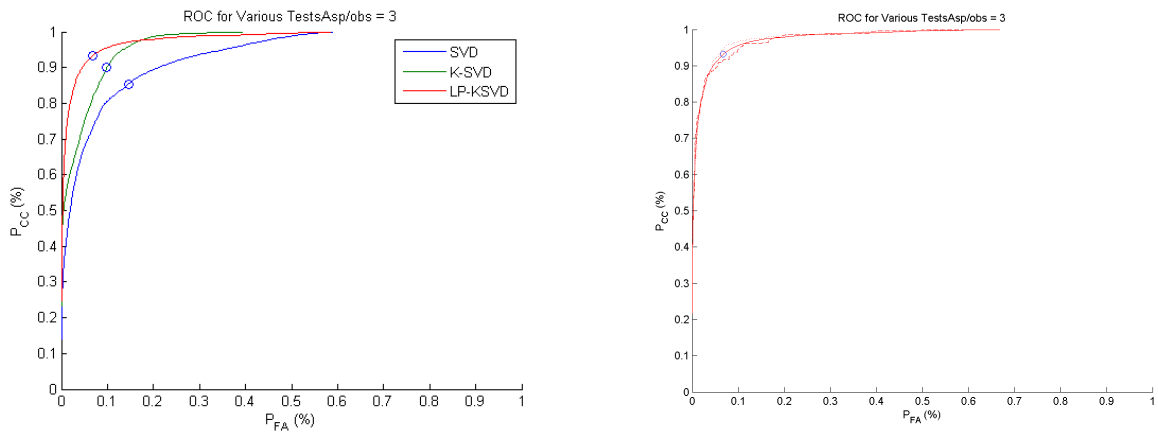
Table 4: TREX13 available target ranges

No.	**TREX #**	**FRM Name**	**5m**	**10m**	**15m**	**20m**	**25m**	**30m**	**35m**	**40m**
1	Target 17	alcyl2ft		X						
2	Target 7	alcyl3ft						X	X	X
3	Target 16	alpipe			X		X	X		
4	Target 20	aluxo		X	X			X		X
5	Target 21	ssuxo		X	X		X	X		
6	Target 25	bullet_105mm_air	X		X	X	X			
7	Target 29	bullet_105mm_h2o			X	X			X	
8	Target 9	howitzer_cap_air		X	X	X			X	
9	Target 28	howitzer_cap_h2o			X		X	X		
10	Target 8	howitzer_nocap	X				X	X		X

5.2.2 Classification Results and Analysis

5.2.2.1 PondEX Testing Results To make a classification decision for a given sonar ping observation, the corresponding AC data vector of dimension $N_{Pond} = 295$ that contains the spectral features of an underwater UXO or non-UXO object at a particular aspect is applied to the classifier. For the MSC, this decision-making is implemented using the classification rule given in

(16). The same process can be carried out for all object aspects in a run in the PondEX09 and PondEX10 testing dataset. However, using the more general decision rule (15) for MSC, one can make decision on an augmented data matrix $\mathbf{Y}$ that contains multiple AC observation vectors at different aspects. This multi-aspect decision-making provides better opportunity to discriminate between UXO and non-UXO object classes. Moreover, multi-aspect classification is more amenable to actual operational situations where several views from an underwater objects are received. Here, we use three aspects of an object to perform classification decision. Although, in the PondEX experiments aspect separation is uniform due to the rail system, to account for platform instability in a realistic data collection scenario aspect separation is modeled by a uniformly distributed random variable $s \sim unif\ \{8, 16\}$ which shuffles data, beginning with the first aspect in a given run that meets the power threshold. Note that the procedure does not shuffle between rotations of objects, only the order in which the aspects of a single linear run are encountered is changed.



(a) ROCs for three methods (3 Asp/dec.)

(b) ROCs for LP-KSVD on Pond data (3 Asp./dec.)

Figure 13: ROCs for Pond data

Figures 13a and 13b give the Receiver Operating Characteristic (ROC) curves of the classifier for a three-aspect decision and using 50 Monte Carlo trials (one for each random three-aspect combination with a fixed initial aspect) for SVD, K-SVD, and LP-KSVD dictionaries. More specifically, Figure 13a gives the average performance over 50 Monte Carlo trials in which each filtered run (single rotation) dataset is grouped into 3 aspects with a uniformly distributed separation, as mentioned before. As can be observed by the 3 blue circles in Figure 13a, the ROC curves for the SVD, K-SVD, and LP-KSVD exhibit knee-point (the point where $P_{CC} + P_{FA} = 1$) probability of correct classification of $P_{CC,SVD} \approx 85.41\%, P_{CC,KSVD} \approx 90.07\%,\ P_{CC,LP-KSVD} \approx 93.20\%$ probability of false alarm $P_{FA,SVD} \approx 14.59\%, P_{FA,KSVD} \approx 9.93\%,\ P_{FA,LP-KSVD} \approx 6.80\%$, respectively. These results show that the MSC performs very well in discriminating 'UXO' vs. 'non-UXO' in the PondEX dataset in spite of the fact that there are indeed obvious discrepancies between the model data used for training and the AC sonar data of actual objects. Additionally, the best overall results for the PondEX datasets are obtained for a classifier trained and tested using data from LP-KSVD dictionary learning method. Figure 13b, on the other hand, shows the best and worst performance over 50 trials for this method. This ROC plot indicates that the classification performance is indeed statistically consistent over 50 trials.

Table 5 displays the confusion matrix for the MSC. It was found that the most common type I error (i.e. 'UXO' mis-classification) occurred for the Stainless Steel UXO object (Table 3, object 4), which was most commonly mis-classified as 2 ft Aluminum Cylinder. All the observed type II errors (i.e. False Alarms) occurred for samples of Aluminum Pipe object, which were almost equally mis-classified as the Aluminum and Stainless Steel UXO classes. These results suggest

that the inclusion of specific misleading aspects in the training of subspaces could potentially have a significant influence on the power of FRM data in representing different classes. Additionally, the results in Figure 13a and Table 5 reveal much improved performance of our approach when compared to the results of previously tested classifiers [14], including: kernel matching pursuit, support vector machine, and relevance vector machine, which at best demonstrated target vs. non-target discrimination with $P_{CC} \approx 90\%$ and $P_{FA} \approx 10\%$. However, in these previous cases the classifiers were trained on real sonar data whereas here training was exclusively done on model-generated datasets via the fast ray model [13].

Table 5: Confusion matrix of classified signals (LP-KSVD)

Truth-Decision	'non-UXO'	'UXO'
alcyl2ft	1.0000	0
alpipe	0.9681	0.0319
aluxo	0.0297	0.9703
ssuxo	0.2103	0.7897

5.2.2.2 TREX13 Fixed Range Testing Results The same procedure described for the PondEX datasets was repeated for TREX13 datasets for the 10m and 30m runs and for all testing datasets for which FRM training data exist (See Table 4). Note that for the 30m runs, the classifier had to be retrained to include the corresponding model-generated data. Additionally, as mentioned before for the TREX13 datasets the AC feature vector is $N = 275$ dimensional. Again, the aspects of a single rotation run, which met the power threshold, were ordered starting with the first high powered aspect and making aspect separation $s \sim unif\ \{8, 16\}$.

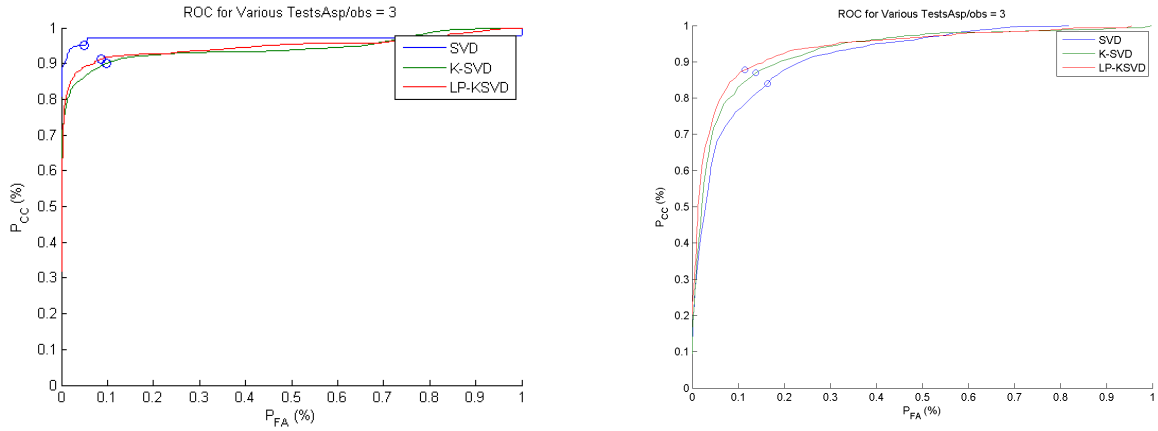
Figure 14b shows the ROC curves for a three-aspect MSC on the 10m run datasets using 50 Monte Carlo trials for the SVD, K-SVD, and LP-KSVD dictionaries. The ROC knee-point performance for these methods are at probability of correct classification of $P_{CC,SVD} \approx 95.1\%, P_{CC,KSVD} \approx 90.1\%$, $P_{CC,LP-KSVD} \approx 91.3\%$, and probability of false alarm $P_{FA,SVD} \approx 4.9\%, P_{FA,KSVD} \approx 9.9\%$, $P_{FA,LP-KSVD} \approx 8.7\%$, respectively. Here, the best overall results for the TREX13 datasets were obtained for a classifier trained and tested using data from the SVD dictionary learning method.

Figure 14a shows the ROC curves for a three-aspect MSC classifier on the 30m run datasets using the SVD, K-SVD, and LP-KSVD dictionaries. The ROC knee-point performance for these methods are at probability of correct classification $P_{CC,SVD} \approx 83\%, P_{CC,KSVD} \approx 86\%$, $P_{CC,LP-KSVD} \approx 88\%$ and probability of false alarm $P_{FA,SVD} \approx 17\%, P_{FA,KSVD} \approx 14\%$, $P_{FA,LP-KSVD} \approx 12\%$, respectively. Again, these results attest to the fact that our three-aspect MSC classifier exclusively trained on model data performs very well in discriminating 'UXO' vs. 'non-UXO' in the TREX13 dataset. Also, as in the PondEX datasets the best overall results for the TREX13 datasets are obtained for a classifier trained and tested using data from LP-KSVD dictionary learning method. Table 6 displays the confusion matrix for the MSC at the decision threshold which corresponds to choosing the minimum statistic between UXO and non-UXO classes in (16).

6 Conclusions and Implications for Future Research/Implementation

6.1 Conclusions and Discussions

The objectives addressed in the current study revolved around the development and testing of UXO detection and classification algorithms from sonar data. A multichannel detection strategy relying on the theory of the GLRT was developed using the broadband coherence method in [1]. For this



(a) ROCs for 10m TREX13 dataset (3 Asp/dec.)

(b) ROCs for 30m TREX13 dataset (3 Asp./dec.)

Figure 14: ROCs for TREX13 dataset

Table 6: Confusion matrix of classified signals (LP-KSVD) at 30m

Truth-Decision	'non-UXO'	'UXO'
alcyl3ft	0.8116	0.1884
alpipe	0.9339	0.0661
aluxo	0.1582	0.8418
ssuxo	0.5556	0.4444
how_cap_h2o	0.0332	0.9668
how_nocap	0.0983	0.9017

problem, our detection hypothesis was that the presence of munitions in the sonar backscatter collected from a hydrophone array over two distinct frequency bands will lead to higher levels of coherence compared to the backscatter from the seafloor alone. This increase in coherence can give one an indication of which areas of the target field may contain potential munitions that will further be analyzed by a UXO vs. non-UXO classifier. Relying on the use of the GLRT and making an assumption that the received time series are wide sense stationary, the amount of coherent information among the channels is measured using the broadband coherence statistic. This multichannel detector was then applied to a dataset collected in a pond facility designed to collect acoustical sonar data from underwater objects in a relatively controlled and clutter-free environment. Results of this preliminary experiment show that the broadband coherence statistic is indeed capable of detecting munitions lying on the seafloor from background as the ROC curve of the detector exhibited a 100% probability of detection with an average of 1 false alarm per image. This suggests that the presence of munitions in the sonar backscatter does indeed lead to a higher level of coherence and that the proposed methodology would be capable of finding regions containing munitions in the target field. This detector could not be tested on the raw TREX13 datasets due to data distribution limitation.

The second objective considered in this seed research was the use of scattering models from various UXO and non-UXO objects for training purposes. The scattering model that has been developed by the APL-UW subcontractors allowed for monostatic SAS data sets to be simulated via a fast ray model that combines an acoustic ray approximation for propagation in a fluid-filled halfspace with scattering from a target (UXO or non-UXO) in free-space. This fast modeling provides a large database for feature extraction and classifier training during the classifier development. The hypothesis tested here is whether or not the classifier trained exclusively on model-generated sonar

data of various UXO-like objects provides enough generalization ability to discriminate munitions with sufficient accuracy in real sonar datasets. This is critical as real sonar data for a wide range of UXOs and in various operational and environmental conditions is currently unavailable. Thus, the development of classification systems that are originally trained on model-based data would provide us with clues on what additional information extracted during the actual field operation can be used to augment the training for optimal performance.

Finally, the third task in this study was the development of a MSC-based UXO vs. non-UXO classifier using several dictionary learning methods. The model-generated data via the fast ray model provides an over-complete dictionary for creation of signal subspaces used in the MSC classifier. In this work, we used three different methods namely the regular SVD, K-SVD [3] and LP-KSVD methods [4]. AC data was used in conjunction with these methods to produce the training and testing datasets for the classifier. Once MSC was properly trained on the model-generated data, it was tested on real PondEX09, PondEX10, and TREX13 datasets. Our binary classification results for UXO vs. non-UXO reveal significant performance improvements when compared to the existing results [11],[14]. What is particularly notable about these results is that the utility of the fast ray model in representing real sonar data for classification purposes has been confirmed. Among the three dictionary learning methods, LP-KSVD was found to consistently provide better results on both datasets. For the PondEX experiments, the knee point performance for LP-KSVD exhibits $P_{CC} \approx 93.2\%, P_{FA} \approx 6.8\%$; while for the TREX13 experiment at the 30m range, the knee point performance for LP-KSVD was at $P_{CC} \approx 87.9\%, P_{FA} \approx 12.79\%$. Comparing the two datasets, the degradation in the classification performance for the TREX13 datasets is mostly attributed to more realistic and challenging bottom conditions and larger number of objects with similar characteristics.

6.2 Possible Future Research/Implementation

This seed research surfaced many critical problems that can be pursued in future research opportunities. Below is a list of only a few items that we believe are worthy of further research.

- ***Task 1: Broadband Coherence Detector with Adaptive Thresholding:***

  Although the broadband coherence detector is fully capable of discriminating time series corresponding to UXO objects from those corresponding to seafloor background, one of the most difficult parts of implementing it in any realistic scenario will be determining a detection threshold that is robustly capable of achieving a desired false alarm rate over the wide range of environments that may be encountered. To achieve this, we propose to adapt this parameter of the algorithm as data is collected within the environment by fitting a parametric distribution to the broadband coherence statistics produced from that data. The parameters of this distribution may then be recursively updated as new data is collected to give the detector the ability to adapt to changing statistical behavior in the likelihood ratio. Note that this unsupervised approach relies on an assumption that it is very unlikely to observe a UXO object so that the statistics of the likelihood ratio are a good indication of that for background alone.

- ***Task 2: Incremental Training for Better Robustness:*** Our results presented in this report clearly indicate the promise and effectiveness of the developed methods for detection and classification of munitions from sonar data. However, it is obvious that the spectral features in the AC data for a specific target vary significantly depending on the object's burial condition, seafloor properties and roughness, actual orientation of the object, range and grazing angles with respect to the sonar, etc. Although, it is unrealistic to expect model data will capture all such variations for target characterization, it can provide us with clues on how to augment the training datasets using perhaps a limited training samples from sonar returns of actual objects to improve the robustness in different environmental conditions.

The importance of this data augmentation beyond the Fast Ray Model data and incremental training using real data to improve classification performance is demonstrated here. Several PondEX object samples (only 2 % of training from real PondEX test data) that correspond to the same objects in the TREX13 dataset were added to the model-generated training set to form the dictionary training set and build the MSC classifier. The trained system was then tested on TREX 10m run datasets. It was found that the addition of real Pond samples from the relevant objects improved the overall results noticeably. The knee-point performances of the SVD, K-SVD, and LP-KSVD for this incremental training test were $P_{CC,SVD} \approx 96.70\%$, $P_{CC,KSVD} \approx 90.43\%$, $P_{CC,LP-KSVD} \approx 94.1\%$ and probability of false alarm $P_{FA,SVD} \approx 3.30\%$, $P_{FA,KSVD} \approx 9.57\%$, $P_{FA,LP-KSVD} \approx 5.9\%$ respectively. Clearly, we didn't attempt to optimize this data selection process based upon the information content of the selected sample for classification purposes. Additionally, we had to retrain (in batch) all the dictionaries and consequently the MSC classifier, which is not obviously an efficient way of doing this training. Thus, we propose to accomplish this by developing new methods for optimal sample selection and incremental training that can update the parameters of the MSC classifier iteratively in order to learn the new and informative samples taken in the new environment while guaranteeing the stability of the previously learned samples.

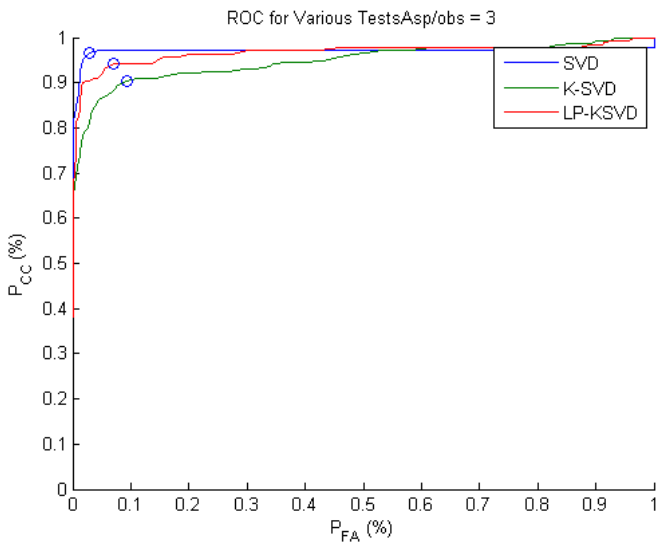
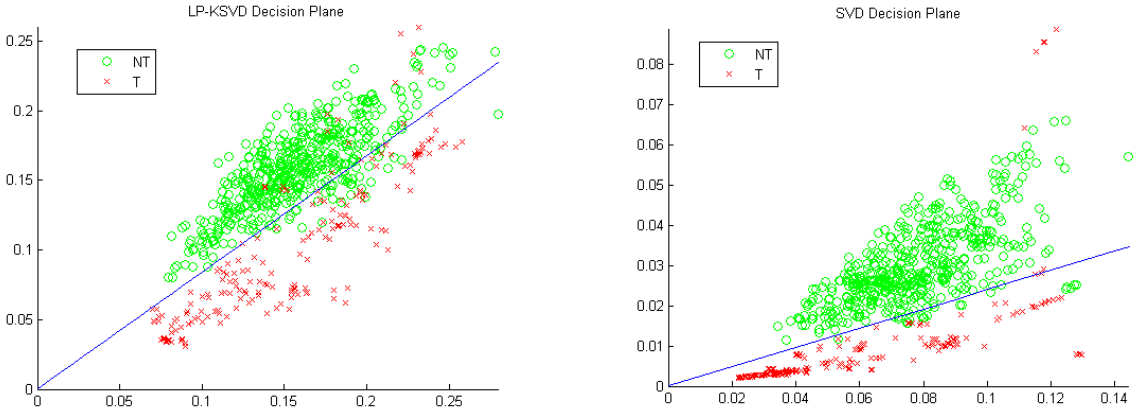


Figure 15: Incremental Learning Result, 10m TREX13 (3 Asp/dec.)

- ***Task 3: Nonlinear Extension of MSC:***

  Figures 16 (a) and (b) shows the decision boundary and the distribution of the MSC's test statistics for UXO and Non-UXO objects in the TREX13 10m runs and LP-KSVD and SVD dictionaries, respectively. As can be seen, there are several objects that are misclassified for both object types. One of the reasons for this imperfect performance has to do with the fact that the MSC is a linear classifier designed based upon the subspaces associated with the UXO and Non-UXO classes. It is well-known [23] that certain non-linearly separable classification problems can easily be converted to linearly separable problems by mapping the data to a high dimensional feature space using kernel-producing nonlinear mapping functions. In this particular application, there is no guarantee that under all conditions the solution of the optimization problem for the model using the LP-KSVD or SVD will lead to features that are sparse and confined to the subspace associated with the class of the object. In other words, the resultant features could "leak" into subspaces associated with the opposite classes, hence leading to possible misclassifications and false alarm as can be seen in Figures

16. To overcome such issues, we propose to investigate the kernel-based extension of the MSC classification framework. It is also interesting to investigate what types of kernels can provide the best results for the corresponding structured learning. We propose to investigate this issue thoroughly in this task.



(a) LP-KSVD dictionary.

(b) SVD dictionary.

Figure 16: Decision Boundary, 10m TREX13 (3 Asp/dec.)

- ***Task 4: Multi-Aspect Classification Fusion:***

  In a real operating environment, the decision about the presence and type of an object is typically made based upon the observation of the properties of the sonar returns at several aspect angles or pings. This is due to the fact that the multi-aspect processing yields substantial improvements in performance, resolution, and sensing of the 3-D properties of the object in a non-isotropic environment. Inspired by these desired properties, we have developed two different general frameworks [9], [11], [8], [10], [24] for performing multiple-aspect or multiple ping processing of sonar data. The framework in [9], [8] is based upon the idea of *multi-aspect feature extraction* that uses the two-channel canonical coordinate decomposition method to extract robust features with maximum coherence (or mutual information) from pairs of sonar pings with certain separation. The idea is that the coherence pattern extracted from the UXO objects differ from those of the Non-UXO objects, hence aiding the overall classification process. The theme of the other framework is *multi-aspect classification* using either (a) a decision-level multi-aspect fusion [24], which linearly or non-linearly combines the individual classification decisions, generated at several aspects, or (b) a feature-level multi-aspect fusion [10] using HMM to generate one decision based upon observing a sequence of AC feature vectors at various aspects with ceratin separations, or (c) a collaborative decision-making process [11], which uses sort of a combination of the feature-level and decision-level fusion methods. In this work, we propose to carefully study and test these different methods and compare their results with the multi-aspect MSC classier in (15). Moreover, we intend to revisit those weighting matrix methods mentioned in the original MSC method [2] which can limit the influence of large error amplitudes for certain measurements or de-emphasize less important AC features by assigning small corresponding weights.

- ***Task 5: Comprehensive Testing & Evaluation:***

  We shall test and evaluate the performance of the developed detection and classification methods developed in this phase of research on several real (e.g., PondEX, TREX13, and BayX) sonar data sets collected in different environmental and background conditions. The specific issues that will be thoroughly studied include:

1. Design and test the modified version of the broadband coherence detector using the proposed adaptive thresholding in many different environmental and operating conditions.
2. Using real and synthesized sonar data determine the effectiveness of the developed adaptive dictionary learning and sparse representation methods in different situations. We will investigate how the model and dictionaries can be designed to guarantee robustness to partial burial and occlusion of the targets. Use these dictionaries to build matched subspace classifiers and test their target versus non-target classification performance using the ROC curve and confusion matrix.
3. The developed kernel matched subspace classifier will be thoroughly tested on the same sonar datasets. The classification performance of the kernel-based MSC will be benchmarked against the linear weighted MSC using the commonly used performance metrics. In addition, robustness to the environmental variations will be evaluated on both PondEX and TREX13 datasets.
4. Benchmark the performance of one of the above-mentioned multi-aspect classifier against the matched subspace classifier in (15) which uses multiple AC feature vectors in the data matrix.
5. Prepare interim and final reports to document all the developments and results of this research after the first and second years. We plan to publish our results in various IEEE Transactions. Our results, publications, reports, and algorithms will be made available, as part of our technology transfer plans.

References

[1] N. Klausner, M. Azimi-Sadjadi, and L. Scharf, "Detection of spatially correlated time series from a network of sensor arrays," *IEEE Transactions on Signal Processing*, vol. 62, no. 6, pp. 1396–1407, March 2014.

[2] A. Salberg, A. Hanssen, and L. L. Scharf, "Robust multidimensional matched subspace classifiers based on weighted least-squares," *IEEE Transactions Signal Processing*, vol. 55, pp. 873–880, March 2007.

[3] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, November 2006.

[4] Y. Zhou, J. Gao, and K. Barner, "Locality preserving KSVD for nonlinear manifold learning," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 3372–3376.

[5] E. T. S. Kargl, K. Williams, "Synthetic aperture sonar imaging of simple finite targets," *IEEE Journal of Oceanic Engineering*, vol. 37, no. 3, pp. 516–532, July 2012.

[6] S. Kargl, K. Williams, T. Marston, and J. Kennedy, "Acoustic response of unexploded ordnance (uxo) and cylindrical targets," *Proceedings of MTS/IEEE Oceans Conference 2010*, pp. 1–5, September 2010.

[7] R. Lim, "Sonar detection and classification of underwater uxo and environmental parameters," *Interim Report, SERDP Project MR-1666*, pp. 1–19, October 2010.

[8] M. A.-S. A. Pezeshki and L. Scharf, "Undersea target classification using canonical correlation analysis," *IEEE Journal of Oceanic Engineering*, vol. 32, pp. 948 – 955, October 2007.

[9] N. Wachowski and M. Azimi-Sadjadi, "A new synthetic aperture sonar processing method using coherence analysis," *IEEE Journal of Oceanic Engineering*, vol. 34, pp. 665–678, October 2011.

[10] M. A.-S. M. Robinson and J. Salazar, "Multi-aspect classification using hidden markov models and neural networks," *IEEE Transactions on Neural Networks*, vol. 16, pp. 447–459, March 2005.

[11] J. Cartmill, M. R. Azimi-Sadjadi, and N. Wachowski, "Buried underwater object classification using a collaborative multi-aspect classifier," *IEEE Journal of Oceanic Engineering*, vol. 34, pp. 32–44, January 2009.

[12] D. Ramirez, J. Via, I. Santamaria, and L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5006–5015, 2010.

[13] S. Kargl, A. Espana, K. Williams, J. Kennedy, , and J. Lopes, "Scattering from objects at a water-sediment interface: Experiment, high-speed and high-fidelity models, and physical insight," *IEEE Journal of Oceanic Engineering*, vol. PP, no. 99, pp. 1–11, August 2014.

[14] S. Kargl and K. Williams, "Full scale measurement and modeling of the acoustic response of proud and buried munitions at frequencies from 1-30khz," *Final Report, SERDP Project MR-1665*, pp. 18–19, May 2012.

[15] A. L. Espana, K. L. Williams, D. S. Plotnick, and P. L. Marston, "Acoustic scattering from a water-filled cylindrical shell: Measurements, modeling, and interpretation," *Acoustical Society of America*, vol. 136, no. 1, pp. 109–121, July 2014.

[16] T. Marston, P. Marston, and K. Williams, "Scattering resonances, filtering with reversible sas processing, and applications of quantitative ray theory," *Proc. 2010 Oceans MTS/IEEE Seattle*, pp. 1–9, September 2010.

[17] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.

[18] L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2146–2157, August 1994.

[19] M. R. Azimi-Sadjadi, J. Kopacz, and N. Klausner, "K-svd dictionary learning using a fast omp with applications," *Proc. of IEEE Intern. Conference on Image Processing (ICIP)*, October 2014.

[20] M. R. Azimi-Sadjadi and S. Kargl, "Multichannel detection and acoustic color-based classification of underwater UXO in sonar," *First SERDP Interim Report*, July 2014.

[21] S. G. Kargl, K. L. Williams, and A. L. Espaa, "Synthetic aperture sonar (sas) and acoustic templates for the detection and classification of underwater munitions." Presented at SERDP & ESTCP Annual Symposium, 30 Nov - 3 Dec 2010, Washington D.C., November 2010.

[22] M. R. Azimi-Sadjadi and S. Kargl, "Multichannel detection and acoustic color-based classification of underwater UXO in sonar," *Second SERDP Interim Report*, October 2014.

[23] B. Scholkopf and S. Alexander, *Learning with Kernels*. The MIT press, 2002.

[24] M. Azimi-Sadjadi, Q. Yao, D. Huang, , and G. Dobeck, "Underwater target classification using wavelet packets and neural networks," *IEEE Transactions on Neural Networks*, vol. 11, no. 3, pp. 784–794, May 2000.