

THESIS

OPTIMAL DICTIONARY LEARNING WITH APPLICATION TO UNDERWATER TARGET
DETECTION FROM SYNTHETIC APERTURE SONAR IMAGERY

Submitted by

Justin Kopacz

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2014

Master's Committee:

Advisor: Mahmood R. Azimi-Sadjadi

Ali Pezeshki
F. Jay Breidt

Copyright by Justin Kopacz 2014

All Rights Reserved

ABSTRACT

OPTIMAL DICTIONARY LEARNING WITH APPLICATION TO UNDERWATER TARGET DETECTION FROM SYNTHETIC APERTURE SONAR IMAGERY

K-SVD is a relatively new method used to create a dictionary matrix that best fits a set of training data vectors formed with the intent of using it for sparse representation of a data vector. K-SVD is flexible in that it can be used in conjunction with any preferred pursuit method of sparse coding including the orthogonal matching pursuit (OMP) method considered in this thesis. Using adaptive filter theory, a new fast OMP method has been proposed to reduce the computational time of the sparse pursuit phase of K-SVD as well as during on-line implementation without sacrificing the accuracy of the sparse pursuit method. Due to the matrix inversion required in the standard OMP, the amount of time required to sparsely represent a signal grows quickly as the sparsity restriction is relaxed. The speed up in the proposed method was accomplished by replacing this computationally demanding matrix inversion with a series of recursive “time-order” update equations by using orthogonal projection updating used in adaptive filter theory. The geometric perspective of this new learning is also provided.

Additionally, a recursive method for faster dictionary learning is also discussed which can be used instead of the singular value decomposition (SVD) process in the K-SVD method. A significant bottleneck in K-SVD is the computation of the SVD of the reduced error matrix during the update of each dictionary atom. The SVD operation is replaced with an efficient recursive update which will allow limited in-situ learning to update dictionaries as the system is exposed to new signals. Further, structured data formatting has allowed a multi-channel extension of K-SVD to merge multiple data sources into a single dictionary capable of creating a single sparse vector representing a variety of multi-channel data.

Another contribution of this work is the application of the developed methods to an underwater target detection problem using coregistered dual-channel (namely broadband and high-frequency) side-scan sonar imagery data. Here, K-SVD is used to create a more optimal dictionary in the sense of reconstructing target and non-target image snippets using their respective dictionaries. The ratio of the reconstruction errors is used as a likelihood ratio for target detection. The proposed methods were then applied and benchmarked against other detection methods for detecting mine-like objects from two dual-channel sonar datasets. Comparison of the results in terms of receiver operating characteristic (ROC) curve indicates that the dual-channel K-SVD based detector provides a detection rate of $P_D = 99\%$ and false alarms rate of $P_{FA} = 1\%$ on the first dataset, and $P_D = 95\%$ and $P_{FA} = 5\%$ on the second dataset at the knee point of the ROC. The single-channel K-SVD based detector on the other hand, provides $P_D = 96\%$ and $P_{FA} = 4\%$ on the first dataset, and $P_D = 96\%$ and $P_{FA} = 4\%$ on the second dataset at the knee point of the ROC. The degradation in performance for the second dataset is attributed to the fact that the system was trained on a limited number of samples from the first dataset. The coherence-based detector provides $P_D = 87\%$ and $P_{FA} = 13\%$ on the first dataset and $P_D = 86\%$ and $P_{FA} = 14\%$ on the second dataset. These results show excellent performance of the proposed dictionary learning and sparse coding methods for underwater target detection using both dual-channel sonar imagery datasets.

ACKNOWLEDGEMENTS

I would first like to thank my adviser, Mahmood R. Azimi-Sadjadi, for his support and invaluable guidance throughout the course of my research. He has taught me the value of well-developed research. His time and assistance has been greatly appreciated throughout my graduate education.

I would also like to thank my committee members, Dr. Ali Pezeshki and Dr. F. Jay Breidt, for their time and assistance.

I would like to thank the Office of Naval Research (ONR) and the Naval Surface Warfare Center (NSWC) in Panama City, Florida for providing the funding and the data used in this project. This project has been funded by the Office of Naval Research (ONR-3210E) under contract number N00014-12-C-0017. Without the funding from ONR and the data from NSWC, this work would have never been completed.

I would also like to thank my colleagues in the Signal and Image Processing Lab. They have provided an excellent working environment. Special thanks to Nick Klausner for his much needed help and advice.

Finally, I would like to thank my family for their support and guidance throughout my entire education.

TABLE OF CONTENTS

| | |
|---|------|
| ABSTRACT..... | ii |
| ACKNOWLEDGEMENTS | iv |
| LIST OF TABLES | vii |
| LIST OF FIGURES | viii |
| Chapter 1. Introduction | 1 |
| 1.1. Problem Statement and Motivation..... | 1 |
| 1.2. Literature Review on Single and Multi-Channel Detection | 2 |
| 1.3. Research Objectives | 4 |
| 1.4. Organization of the Thesis..... | 5 |
| Chapter 2. Sparse Coding and Dictionary Creation..... | 6 |
| 2.1. Introduction..... | 6 |
| 2.2. Review of K-SVD Method | 7 |
| 2.3. Conclusion | 11 |
| Chapter 3. New Fast Sparse Coding and Dictionary Learning Methods | 13 |
| 3.1. Introduction..... | 13 |
| 3.2. Fast OMP using Orthogonal Projection Updating..... | 14 |
| 3.3. Fast Dictionary Update Using Recursive Updating Process..... | 18 |
| 3.4. Conclusion | 20 |
| Chapter 4. Application to Underwater Target Detection Using Sonar Imagery | 21 |
| 4.1. Introduction..... | 21 |
| 4.2. Data Description and Multimodal K-SVD | 22 |
| 4.3. Detection Results | 27 |

| | |
|---|----|
| 4.4. Conclusion | 36 |
| Chapter 5. Conclusion and Suggestions for Future Work | 37 |
| 5.1. Conclusion and Discussion | 37 |
| 5.2. Future Work | 38 |
| BIBLIOGRAPHY | 40 |
| LIST OF ABBREVIATIONS | 43 |

LIST OF TABLES

| | | |
|-----|---|----|
| 2.1 | OMP Algorithm..... | 9 |
| 2.2 | K-SVD Algorithm..... | 11 |
| 3.1 | Fast OMP Algorithm..... | 17 |
| 4.1 | Detection Results without Grouping..... | 35 |
| 4.2 | Detection Results with Grouping..... | 35 |

LIST OF FIGURES

| | | |
|------|--|----|
| 3.1 | Geometric Interpretation of Fast OMP Learning. | 18 |
| 3.2 | Timing Analysis of Pursuit Methods. | 18 |
| 4.1 | Difficult High Frequency Image Containing High Density of Bottom Clutter. | 23 |
| 4.2 | Difficult Broadband Image Containing High Density of Bottom Clutter. | 23 |
| 4.3 | Exclusive Dictionary Reconstructions for a Target Sample. | 25 |
| 4.4 | Training Samples and the Corresponding K-SVD Atoms. | 26 |
| 4.5 | Diagram of K-SVD Testing Phase. | 27 |
| 4.6 | Receiver Operating Characteristic Dataset 1. | 28 |
| 4.7 | A HF Sonar Image And Detected Contacts. | 29 |
| 4.8 | A HF Sonar Image And Detected Contacts - Updated Threshold. | 29 |
| 4.9 | 3D Analysis of Likelihood Ratios $\Lambda(\mathbf{z}_k)$ | 30 |
| 4.10 | Receiver Operating Characteristic Dataset 1 using ROI Grouping. | 31 |
| 4.11 | A Difficult HF Sonar Image With Detected Contacts After Grouping. | 32 |
| 4.12 | A Difficult HF Sonar Image With Detected Contacts - Without Clustering. | 32 |
| 4.13 | Receiver Operating Characteristic Dataset 2. | 33 |
| 4.14 | Broadband Image in Dataset 2. | 34 |
| 4.15 | High Frequency Image in Dataset 2 with Detected Contacts. | 34 |

CHAPTER 1

INTRODUCTION

1.1. PROBLEM STATEMENT AND MOTIVATION

Underwater mine detection has attracted considerable attention in recent years attributed to the inherent risks involved with the mines in question. This problem is complicated as the mines are subjected to spatially random bottom clutter both of natural (e.g., coral, rocks, etc.) and man-made sources (e.g., scrap metal, oil barrels, lobster traps, etc.). Other than the widely varying environmental state, the mines can be proud on the surface of the ocean floor or even partially or fully buried. Further, mine shape can vary, and they can even be of varied orientation and size as a result of operational conditions.

These underwater mines can make global waterways unsafe for both military and commercial vessels. Impacted commercial industries include telecommunication, shipping, offshore renewable energy generation, offshore oil exploration and extraction, dredging, historical archeology, and even recreational diving. Mines threaten oceanic development and travel as well as pose a danger to humans and the environment. As a result, they must be detected and located quickly and efficiently.

Side-scan sonar has become increasingly common in the field of underwater mine detection allowing large scale searches. In the past, side-scan sonar was operated by towing a submersible sonar vehicle (tow-fish) which emitted narrow fans of acoustic energy perpendicular to the direction of motion [1]. The acoustic return (echo) is then received in complex-value and stored for post-processing where it can be used for target detection and classification. More recently, autonomous underwater vehicles (AUV) such as the REMUS600 have been equipped with the Small Synthetic Aperture Minehunter (SSAM) platform [2] to automate the sonar image acquisition process. Added complications arise in locating the targets in these sonar generated images as the AUV can be deployed in many different ocean environments, such as sandy, rocky, or one with heavy vegetation.

Due to the wide area this specific AUV can cover (0.28 square nautical miles per hour), any applicable target detection algorithm must be robust with regard to environmental variations.

1.2. LITERATURE REVIEW ON SINGLE AND MULTI-CHANNEL DETECTION

Substantial effort has been devoted to developing both single and multi-channel or multi-sonar (high frequency and broadband) detection methods [1] - [13]. Note that high frequency (HF) sonar yields higher resolution images with better target definition while broadband (BB) sonar allows higher clutter suppression capabilities. An in-depth review of many different single-channel side-scan sonar detection methods can be found in [3]. One of the most straightforward single-channel methods involves employing pure image processing methods and high quality sonar images [1]. It is well-known that targets are characterized by a distinct highlight followed by a shadow. Highlights are first detected using the highest percentiles of the pixel intensity histogram within a local region, and shadows are detected based off the lowest percentiles. Highlight/shadow pairs that fit a set of geometrical relationships are determined as targets. This allows the detector to pickup mines of varying shape, orientation, and size which will change as a function of the elevation and direction of the AUV. For high quality images, this method has proved to be successful; however, heavily cluttered images will lead to a high false alarm rate and targets may be missed as this detector is easily deceived. A similar tactic is used in [4] though instead of using geometrical relationships, a model-based method is used in which templates of known shadows of mines are compared to an extracted shadow. Both of the previous methods rely on high resolution sonar images and well-defined target highlight-shadow structures. Much of the difficulty in this detection problem lies in the fact that partially and fully buried mines will not produce a prominent highlight/shadow pair while pronounced seafloor clutter (e.g. rocks and coral) may have a distinct highlight/shadow pair. If this is the case, both of these methods will fail to detect the target while producing high incident of false alarm.

Neural networks have also been applied to single-channel sonar datasets. In [5], a matched filter was employed based on a template of a target. The result of this filter was fed into a k-nearest neighbor classifier and an additional optimal discriminatory filter to determine if a region of interest (ROI) contains a target. A neural network was also employed in [6]; however, each ROI is first decomposed using a wavelet transform before inputting into the neural network. This wavelet transform allowed a unique multi-resolution analysis of each ROI before the neural network determines if the ROI contains a target. Both of these methods have proven to be very effective at distinguishing targets from non-targets due to their feature extraction processes; however, they only employ a single sonar channel. As a result, when a mine is partially or fully buried and does not have a distinct a highlight/shadow characteristic, the ROI will generally be disregarded after the feature extraction process due to the lack of mine-like features.

More recently, in [7] Canonical Correlation Analysis (CCA) is used to discriminate between mine-like objects and non-mine-like objects from dual-channel sonar imagery by exploiting the differences between coherence patterns of sonar pings that contained mine-like objects compared to those that did not contain mine-like objects. Results in [8] - [11] on the buried object scanning sonar (BOSS) system dataset have also shown high discrimination between targets and non-targets using single-channel sonar. The work in [7] is furthered in [12] with a natural extension through a generalized eigenvalue problem of a covariance matrix of N channels and the corresponding block-diagonal matrix through what is named Multi-Channel Coherence Analysis (MCA) to parallel its single sonar predecessor (CCA). Using the new multi-channel coordinate system, the expression for the log-likelihood ratio used in the Gauss-Gauss detector are reformulated. Substantial detection rate gains are noticed when using two-sonar (HF and one BB) and three-sonar (HF and two BB) detection systems. The point of diminishing return occurred when increasing from three to four-sonar (HF and three BB) images indicating that the fourth sonar image does not add any new information.

In [13], the detection problem is cast as a test of independence among multiple random vectors and is solved using the Generalized Likelihood Ratio Test (GLRT). The GLRT tests whether or not the covariance matrix composed of N channels is block-diagonal through the use of a generalized Hadamard ratio. Assuming the null hypothesis is true (no target), the covariance matrix is block diagonal and each of the N channels contain only noise (or background). If the alternative is true, the covariance matrix won't be block diagonal and the N channels contain noise plus some unknown signal (e.g. mine-like object) leading to higher levels of correlation among all N channels. Note that due to the generalization of this algorithm it has a wide range of applications in addition to underwater mine detection. This multi-channel coherence framework is used later in this thesis as the baseline results of a two-channel detector.

1.3. RESEARCH OBJECTIVES

As mentioned in the previous section, a wide range of detectors have been developed for side-scan sonar imagery with each method having its own benefits and shortcomings. The goal of this particular work is to develop and present a new detection method based upon sparse coding and dictionary learning. The idea is to sparsely represent each region of interest (ROI) within a sonar image on the basis of some dedicated dictionary to determine if it contains a target. The detection hypothesis is that a dictionary formed exclusively from data samples of one class (target and non-target) will poorly represent data samples from the opposite class. This hypothesis will be tested as a target detector by sparsely representing an ROI on the basis of two exclusive class dictionaries and measuring the dictionaries ability to represent the original ROI. As two sonar channels are available, the effects of expanding the dictionaries to include both channels will also be investigated. As the sonar operates in a wide variety of environments, the detector must be robust to the typical changes in operational and environmental conditions. The sparse detector should operate with either a single sonar or allow multiple (HF and BB) sonar channels to improve target/non-target discrimination capabilities if available.

Results will be presented for both single-channel (HF) dictionaries and a dual-channel (HF and BB) dictionaries for two datasets provided by the Naval Surface Warfare Center, Panama City (NSWC-PC). As target and non-target samples are required to train the dictionaries, they will be hand selected from only the first dataset. Detection results will be in the form of ROC curves for both single-channel and dual-channel detectors as well as the multi-channel frequency-based coherence detector in [13].

1.4. ORGANIZATION OF THE THESIS

This thesis is organized as follows: Chapter 2 gives an in-depth review of the K-SVD method which is one of the most popular dictionary creation methods for sparse coding. A common sparse pursuit method, namely orthogonal matching pursuit (OMP) is first reviewed and its properties are described. This is the first step in the K-SVD algorithm. Then, the dictionary update step of the K-SVD algorithm is shown in detail. For both phases of K-SVD, the limiting factors are discussed. Chapter 3 further expands on OMP by developing a new fast OMP method based upon orthogonal projection updating used in adaptive filter theory. The dictionary update phase of K-SVD is also replaced with a recursive set of coupled equations. Both of these methods are then applied in Chapter 4 to two dual-channel sonar imagery datasets as a way to validate the effectiveness of the algorithms and the detection hypothesis. Detection results and the ROC curves are presented for the single-channel K-SVD detector, dual-channel K-SVD detector and the dual-channel coherence detector in [13]. Finally, Chapter 5 concludes the thesis with a discussion and goals for future work.

CHAPTER 2

SPARSE CODING AND DICTIONARY CREATION

2.1. INTRODUCTION

Sparse coding has recently become popular due to widespread applications in signal and image processing. Sparse coding involves finding a linear combination of a small set of vectors from a typically large and redundant (overcomplete) set of dictionary vectors that best represent a particular data set. Basis pursuit (BP) [14], matching pursuit (MP) [15], orthogonal matching pursuit (OMP) [16] are among the more commonly used methods for finding both the set of vectors and the coefficients associated with each vector. Each method requires a suitable dictionary of atoms to guarantee the output of an accurate sparse signal representation, and each method features a tradeoff between accuracy and speed. Naturally, selection of the dictionary matrix is critical to the accuracy of the sparse signal.

Many methods have been proposed to create such dictionaries based on the multi-variate distribution of an input training data set. For example, Method of Optimal Dictionary (MOD) [17] finds the optimum dictionary by solving a minimum mean squared error (MMSE) problem. A maximum a posteriori (MAP) method [18] has also been proposed which alters MOD by looking at the posteriori distribution rather than the likelihood function as in MOD. The K-SVD algorithm [19], on the other hand, is considered as a generalization of the K-means algorithm [20] where dictionary atoms are selected such that the reconstruction error is minimized. In general, K-SVD involves a sparse coding phase followed by a dictionary update phase which are repeated to monotonically reduce the reconstruction error. Dictionaries created using K-SVD are superior in both compression and signal recovery due to the Gauss-Seidel nature of the learning. Details of the K-SVD method are given in this chapter to provide a foundation for the developed methods in Chapter 3.

This chapter is organized as follows. Section 2.2. gives an in-depth review of the K-SVD and its two phases, namely the sparse pursuit phase and the dictionary update phase . Section 2.3 gives concluding remarks on the standard K-SVD method and its computational problems.

2.2. REVIEW OF K-SVD METHOD

The main purpose of K-SVD is to create an optimal dictionary that can be used to reduce the dimension of a signal vector by representing it as a sparse linear combination of relatively few atoms. K-SVD aims to solve a constrained minimization problem to reduce the reconstruction error in a set of training vectors. Let $Y \in \mathbb{R}^{N \times M}$ be a matrix consisting of the input training data vectors \underline{y}_k for $k \in [1, M]$ as its columns, $D \in \mathbb{R}^{N \times K}$ be the dictionary matrix to be found, and $X \in \mathbb{R}^{K \times M}$ be the sparse representation of Y in terms of the dictionary atoms. Note it is desired that the number of non-zero elements of each \underline{x}_k is substantially less than N as the dimension must be reduced in this process. The constrained optimization problem is given by,

$$\min_{D, X} \|Y - DX\|_F^2 \quad s.t. \quad \|\underline{x}_k\|_0 \leq \tau \quad \forall k \quad (2.1)$$

where $\|\cdot\|_F^2$ is the Frobenius norm of a matrix, and $\|\cdot\|_0$ is the l_0 norm which simply counts the non-zero elements of a vector limited by some fixed value τ .

During the training, the K-SVD algorithm is composed of two-phases: (i) a sparse representation phase where for each \underline{y}_k the corresponding \underline{x}_k is computed based on a given D using some pursuit method, e.g. OMP [16], (ii) and a dictionary update phase where D is updated based on minimizing the reconstruction error using SVD. These two phases are repeated until convergence through monotonic MSE reduction [19] which can be seen in detail in the following subsections.

2.2.1. SPARSE CODING PHASE

In the sparse coding phase, a pursuit algorithm is applied to data matrix Y on the basis of a dictionary matrix D to create a sparse feature matrix X . The goal is to minimize the cost function

in (2.1) or alternatively $\sum_{k=1}^M \left\| \underline{y}_k - D\underline{x}_k \right\|^2$ by finding sparse solution vectors, \underline{x}_k 's, given dictionary matrix D . Any pursuit algorithm [14] - [16], can be employed to represent each \underline{y}_k sparsely based on the dictionary D to create the corresponding \underline{x}_k ; however, OMP [16] is adopted here due to its much faster performance and comparable accuracy.

OMP is a greedy algorithm which iteratively selects the best dictionary atoms for each \underline{y}_k to reduce the reconstruction error. The reconstruction error (known as the residual), is initialized as the entire input signal $\underline{r}_0 = \underline{y}_k$ and the set of selected atoms is initialized as $S_0 = \emptyset$ i.e. the empty set. At iteration t , the residual is projected onto the dictionary atoms, and the atom which yields the largest inner product is selected and the index set is updated. That is, $S_t = S_{t-1} \cup \{k_t\}$. This is done using

$$k_t = \underset{j}{\operatorname{argmax}} \left| \underline{r}_{t-1}^t \underline{d}_j \right|, \forall j \notin S_{t-1} \quad (2.2)$$

An augmented dictionary of selected atoms is then created to reflect the new index in S_t by including the new atom i.e. $D_t = [D_{t-1} \ \underline{d}_{k_t}]$. After updating D_t , the new sparse signal $\hat{\underline{x}}_k(t)$ can be found using the Least Squares (LS) method as,

$$\hat{\underline{x}}_k(t) = (D_t^t D_t)^{-1} D_t^t \underline{y}_k = D_t^\dagger \underline{y}_k \quad (2.3)$$

where D_t^\dagger is the Moore-Penrose matrix inverse of D_t . Note that the sample index k has been dropped from D_t and S_t for simplicity in notation. Using this new sparse vector $\hat{\underline{x}}_k(t)$, the residual must be recomputed as

$$\underline{r}_t = \underline{y}_k - D_t \hat{\underline{x}}_k(t) \quad (2.4)$$

It is also possible to view this algorithm in a geometrical perspective, where $P_{D_t} = D_t (D_t^t D_t)^{-1} D_t^t$ is the projection matrix onto the subspace spanned by the columns of D_t . Using this projection matrix, the residual \underline{r}_t can be viewed simply as the projection of \underline{y}_k onto the orthogonal complement

of subspace $\langle D_t \rangle$. In other words $\underline{r}_t = P_{D_t}^\perp \underline{y}_k$ where $P_{D_t}^\perp = I - P_{D_t}$. This geometrical perspective will be used in Chapter 3 for the Fast OMP method.

This algorithm is very easy to implement due to the iterative nature of the method and the simplicity of the concept. Stopping conditions can be set to either break at a certain number of iterations for a predefined sparsity limit or stop when the magnitude of the residual falls below a predefined error threshold. The steps in the OMP algorithm can be seen in Table 2.1.

TABLE 2.1. OMP Algorithm

| |
|---|
| Task: For a signal vector $\underline{y}_k \in \mathbb{R}^N$, dictionary $D \in \mathbb{R}^{N \times K}$, and sparsity restriction τ , find a coefficient vector $\hat{\underline{x}}_k(t) \in \mathbb{R}^K$ with at most τ non-zero coefficients. |
| Initialization: $\underline{r}_0 = \underline{y}_k$, $S_0 = \emptyset$, $\hat{\underline{x}}_k(0) = \underline{0}$, $D_0 = \mathbf{0}$ |
| Procedure: |
| (a) $k_t = \underset{j}{\operatorname{argmax}} r_{t-1}^t d_j , \forall j \notin S_{t-1}$ |
| (b) Update the index set $S_t = S_{t-1} \cup \{k_t\}$ |
| (c) $D_t = [D_{t-1} \ d_{k_t}]$ |
| (d) $\hat{\underline{x}}_k(t) = (D_t^t D_t)^{-1} D_t^t \underline{y}_k$ |
| (e) $\underline{r}_t = \underline{y}_k - D_t \hat{\underline{x}}_k(t)$ |
| (f) Set $t = t + 1$. Repeat steps (a) through (f) until convergence. |

2.2.2. DICTIONARY UPDATE PHASE

The notation in this subsection is mostly from [19] for consistency. In the dictionary update phase, the cost function in (2.1) can be rewritten as

$$\begin{aligned} \|Y - DX\|_F^2 &= \left\| Y - \sum_{j=1}^K d_j \underline{x}_T^j \right\|_F^2 = \\ &= \left\| \left(Y - \sum_{j \neq k} d_j \underline{x}_T^j \right) - d_k \underline{x}_T^k \right\|_F^2 = \|E_k - d_k \underline{x}_T^k\|_F^2 \end{aligned} \quad (2.5)$$

where \underline{x}_T^k is not the sparse representation of the k^{th} training vector, rather it is the row vector of the coefficients of all the \underline{x}_i 's corresponding to the k^{th} atom, and E_k is the error matrix which represents the reconstruction error when neglecting the k^{th} atom.

At this point, SVD can be used to minimize the error in (2.5) to find \underline{d}_k such that the combination $\underline{d}_k \underline{x}_T^k$ best approximates E_k ; however, this does not address the sparsity requirement. Due to the SVD computation, \underline{x}_T^k will be filled rather than being sparse. Instead a group of indices will be defined as $\omega_k = \{i \mid 1 \leq i \leq K, x_T^k(i) \neq 0\}$ where $x_T^k(i)$ is the i^{th} element of \underline{x}_T^k indicating the training samples in Y that use the dictionary atom \underline{d}_k . The restricted error matrix E_k^R when only using these indicated atoms is then defined as $E_k^R = E_k \Omega_k$. Note that in this equation, Ω_k is defined as a matrix of size $[M \times \omega_k]$ with ones on the $(\omega_k(i), i)$ entries and zeroes everywhere else. When post-multiplying E_k by this matrix such as $E_k^R = E_k \Omega_k$, the selection process chooses columns of E_k associated with the dictionary atom \underline{d}_k . When post-multiplying a vector by Ω_k such as $\underline{x}_T^k \Omega_k$, the zero entries are discarded, and the row vector \underline{x}_R^k is returned with a length of ω_k . The cost function in (2.5) can then be written as

$$\left\| E_k \Omega_k - \underline{d}_k \underline{x}_T^k \Omega_k \right\|_F^2 = \left\| E_k^R - \underline{d}_k \underline{x}_R^k \right\|_F^2 \quad (2.6)$$

When decomposing E_k^R using SVD as $E_k^R = U \Delta V^T$, the solution for the updated dictionary atom $\hat{\underline{d}}_k$ is the eigenvector corresponding to the largest eigenvalue. In other words, $\hat{\underline{d}}_k$ is the first column of U .

Because $\hat{\underline{d}}_k$ has been changed, \underline{x}_R^k must also be updated to reflect these changes otherwise E_k would be computed incorrectly for the next atom. The coefficient vector \underline{x}_R^k is updated as the first column of V multiplied by $\Delta(1,1)$ i.e. $\hat{\underline{x}}_R^k = \underline{v}_1 \Delta(1,1)$ where \underline{v}_1 is the first column of V . After updating the k^{th} dictionary atom, the same procedure is followed for the next atom until all atoms have been individually updated. Using this newly updated dictionary for iteration J , $D^{(J)}$, the sparse matrix X can be recomputed using any pursuit method. The dictionary update and sparse representation phases will be repeated until the stopping conditions are met (either maximum number of iterations or a reconstruction error below some threshold). The steps in the K-SVD algorithm are given in Table 2.2.

TABLE 2.2. K-SVD Algorithm

| |
|--|
| Task: Find the best dictionary $D \in \mathbb{R}^{N \times K}$ to represent the data samples $\{y_k\}_{k=1}^M$ as sparse compositions by solving $\min_{D,X} \ Y - DX\ _F^2$ s.t. $\ x_k\ _0 \leq \tau \forall k$ |
| Initialization: Set the dictionary matrix $D^{(0)} \in \mathbb{R}^{N \times K}$ with random l_2 normalized columns. Set $J = 1$. Repeat the following steps until a stopping rule is met: |
| Procedure: |
| Sparse Coding Stage: Generate X by computing the sparse representation x_i for each y_k based on the dictionary D . This can be accomplished by using a pursuit algorithm such as OMP. |
| Codebook Update Stage: Each column $k \in 1 : K$ in $D^{(J-1)}$ is updated by: <ul style="list-style-type: none"> (a) Compute error matrix: $E_k = Y - \sum_{j \neq k} d_j x_T^j$ (b) Define group of training data that use the atom: $\omega_k = \{i \mid 1 \leq i \leq K, x_T^k(i) \neq 0\}$ (c) Obtain E_k^R by restricting E_k to only columns corresponding to ω_k (d) Apply SVD: $E_k^R = U \Delta V^T$. The updated dictionary column \hat{d}_k is the first column of U, and the updated coefficient vector \hat{x}_R^k is $v_1 \Delta(1, 1)$. |
| Set $J = J + 1$ and repeat until convergence |

2.3. CONCLUSION

K-SVD has become increasingly popular due to the optimal dictionary it creates. The major downside of K-SVD is the speed of training. OMP is the first step in K-SVD due to the sparse coding phase where each training sample must be sparsely represented based on the current dictionary matrix D . As seen in Section 2.2.1, the inverse of matrix $(D_t^t D_t)$ is required to determine the coefficients of \hat{x}_k at each iteration. For low sparsity, this inverse can be computed quickly as the matrix is of low dimension. As the sparsity requirement is relaxed, the amount of time required to compute the inverse grows rapidly. A method that does not require matrix inversion would therefore speed up not only training time, but also the time required for in-situ operation where OMP must be completed many times on a testing dataset. This fast pursuit method can be found in Section 3.2.

Additionally, due to the SVD computation required in the update of each dictionary atom in K-SVD, a substantial amount of time is devoted to train a dictionary. For offline training, this is acceptable as training will not interfere with normal operation. When in-situ dictionary updates are required, a faster method is needed that produces the same optimal dictionary without the costly SVD operation. An alternative dictionary update for K-SVD is presented in Section 3.3, which uses a recursive learning process void of any SVD operation to extract only the dominant eigenvector.

CHAPTER 3

NEW FAST SPARSE CODING AND DICTIONARY LEARNING METHODS

3.1. INTRODUCTION

As stated in Section 2.3, K-SVD is limited by the fact that it can only be used in offline learning situations. As the detector may encounter new contacts when applied to new environments in a real operational setting, it may be necessary to update/expand the dictionaries to retain robustness to environmental variations. The emphasis of this chapter is on replacing the phases of K-SVD with faster operating methods to allow higher operating speeds with the goal of in-situ learning or allowing shorter offline learning sessions. The limitation is due in part to the costly SVD operation involved in the dictionary update phase as well as the matrix inversion in the OMP step. This chapter features a dictionary update that extracts only the dominant eigenvalues as seen in [19]. As before in K-SVD, each atom is updated one at a time; however, the process is much more lightweight without the SVD operations.

As K-SVD is a two phase process, the time required to complete K-SVD can be further reduced by increasing the speed of the sparse pursuit phase. Though OMP is a relatively fast sparse pursuit method to begin with, it still relies on the computationally demanding matrix inversion. For low sparsity, this computation can be computed quickly; however, if sparsity is relaxed this can become problematic. The OMP algorithm used in the sparse coding phase of K-SVD has been improved by replacing this matrix inversion with an efficient recursive update based on orthogonal projection updating used in adaptive filter theory. As an added benefit, this recursive update also prevents costly recomputation of commonly used factors within the algorithm.

This chapter is organized as follows: Section 3.2 introduces the new fast OMP algorithm using a series of lightweight recursive update equations. Section 3.3 contains details of the fast dictionary

update for reducing the time required in the dictionary update phase of K-SVD. Finally, Section 3.4 contains concluding remarks and observations.

3.2. FAST OMP USING ORTHOGONAL PROJECTION UPDATING

Due to the number of occurrences of OMP in training and in the in-situ operation of the target detector, one of the major bottlenecks in the process is the time it takes to sparsely represent an image snippet. It is ideal to enhance this procedure to decrease the amount of time required to locate targets. As before, let $D = [d_1, d_2, \dots, d_K]$ be an overcomplete dictionary matrix designed by using K-SVD. Given some observation vector $\underline{y}_k \in \mathbb{R}^N$, we would like to find a sparse representation $\underline{x}_k \in \mathbb{R}^K$, $K \ll N$ for this vector using a fixed number of non-zero components of \underline{x}_k computed as $\|\underline{x}_k\|_0 \leq \tau$. Instead of using a predefined sparsity limit, it is also possible to set a lower limit on the reconstruction error with a preselected tolerance ϵ by using the smallest possible number of basis vectors in D . That is, the error in this representation $\underline{e}_k = \underline{y}_k - D\underline{x}_k$, satisfies $\|\underline{e}_k\| \leq \epsilon$.

Recall the most time consuming computation from the original OMP algorithm is the computation of D_t^\dagger in the Least Squares solution for $\hat{\underline{x}}_k(t)$ as illustrated in Section 2.2.1. Further, due to the growing size of the matrix inversion in (2.3), the required time rapidly increases with each additional iteration. To remedy this problem, let us start by considering, $Q_{D_t} = (D_t^t D_t)^{-1} D_t^t$ as the LS filter based upon matrix D_t , i.e.

$$\hat{\underline{x}}_k(t) = (D_t^t D_t)^{-1} D_t^t \underline{y}_k = Q_{D_t} \underline{y}_k \quad (3.1)$$

Also, recall that when considering the geometric perspective of OMP, the residual error was written as,

$$\underline{r}_t = \underline{y}_k - D_t \hat{\underline{x}}_k(t) = P_{D_t}^\perp \underline{y}_k \quad (3.2)$$

where $P_{D_t}^\perp = I - P_{D_t}$ and $P_{D_t} = D_t (D_t^t D_t)^{-1} D_t^t$ is the projection matrix onto subspace spanned by columns of D_t . As can be seen in (3.1) and (3.2), the computational effort of the original

OMP algorithm grows very fast as the number of chosen atoms increases. Additionally, to relate this sparse coding to iterative K-SVD an interface between these two processes is needed. To develop our fast OMP, we use the orthogonal projection updating equations for P_{D_t} and Q_{D_t} [21] by first considering the previous iteration's projection matrix $P_{D_{t-1}} = D_{t-1}(D_{t-1}^t D_{t-1})^{-1} D_{t-1}^t$ and expanding it to include the new atom \underline{d}_{k_t} . This can be written as

$$\begin{aligned}
P_{D_t} = P_{[D_{t-1} \ \underline{d}_{k_t}]} &= [D_{t-1} \ \underline{d}_{k_t}] \left\{ \begin{bmatrix} D_{t-1}^t \\ \underline{d}_{k_t}^t \end{bmatrix} [D_{t-1} \ \underline{d}_{k_t}] \right\}^{-1} \begin{bmatrix} D_{t-1}^t \\ \underline{d}_{k_t}^t \end{bmatrix} \\
&= P_{D_{t-1}} + P_{D_{t-1}}^\perp \underline{d}_{k_t} (\underline{d}_{k_t}^t P_{D_{t-1}}^\perp \underline{d}_{k_t})^{-1} \underline{d}_{k_t}^t P_{D_{t-1}}^\perp \\
&= P_{D_{t-1}} + P_{\underline{d}_{k_t}}
\end{aligned} \tag{3.3}$$

If $\hat{\underline{d}}_{k_t} = P_{D_{t-1}} \underline{d}_{k_t}$ is the projection of \underline{d}_{k_t} onto $\langle D_{t-1} \rangle$, let $\tilde{\underline{d}}_{k_t} = P_{D_{t-1}}^\perp \underline{d}_{k_t}$ be the orthogonal component of this projection (or innovation). Using this, the orthogonal projection updating equation can be written as

$$P_{D_t}^\perp = P_{D_{t-1}}^\perp - P_{\tilde{\underline{d}}_{k_t}} \tag{3.4}$$

where $P_{\tilde{\underline{d}}_{k_t}} = \frac{\tilde{\underline{d}}_{k_t} \tilde{\underline{d}}_{k_t}^t}{\|\tilde{\underline{d}}_{k_t}\|^2}$ is the projection matrix for $\tilde{\underline{d}}_{k_t}$. It can also be seen that $Q_{D_t} P_{D_t} = Q_{D_t}$. When premultiplying the update equation for P_{D_t} by Q_{D_t} , the update equation for the LS filter can be found.

$$\begin{aligned}
Q_{D_t} &= \begin{bmatrix} Q_{D_{t-1}} \\ 0 \end{bmatrix} + \begin{bmatrix} -Q_{D_{t-1}}^t \underline{d}_{k_t} \\ 1 \end{bmatrix} (\underline{d}_{k_t}^t P_{D_{t-1}}^\perp \underline{d}_{k_t})^{-1} \underline{d}_{k_t}^t P_{D_{t-1}}^\perp \\
&= \begin{bmatrix} Q_{D_{t-1}} \\ 0 \end{bmatrix} + \begin{bmatrix} -\underline{b}_{t-1} \\ 1 \end{bmatrix} \underline{q}_t^t
\end{aligned} \tag{3.5}$$

where $\underline{b}_{t-1} = Q_{D_{t-1}} \underline{d}_{k_t}$ i.e. filtered version of \underline{d}_{k_t} based upon LS filter $Q_{D_{t-1}}$, and $\underline{q}_t^t = \frac{\tilde{\underline{d}}_{k_t}^t}{\|\underline{d}_{k_t}\|^2}$ is the LS filter operator using $\tilde{\underline{d}}_{k_t}$.

Premultiplying (3.4) and (3.5) by \underline{y}_k yields the recursive update equations for \underline{r}_t and $\hat{\underline{x}}_k(t)$, respectively, i.e.

$$\underline{r}_t = \underline{r}_{t-1} - \alpha_t \tilde{\underline{d}}_{k_t} \quad (3.6)$$

and

$$\hat{\underline{x}}_k(t) = \begin{bmatrix} \hat{\underline{x}}_k(t-1) \\ 0 \end{bmatrix} + \alpha_t \begin{bmatrix} -\underline{b}_{t-1} \\ 1 \end{bmatrix} \quad (3.7)$$

where $\alpha_t = \underline{q}_t^t \underline{y}_k = \frac{\underline{d}_{k_t}^t \underline{r}_{t-1}}{\|\underline{d}_{k_t}\|^2}$ i.e. filtered version of \underline{y}_k based upon LS filter operator \underline{q}_t^t . Thus, the adjustment term in (3.7) corresponding to the previous coefficients is equal to the products of two filtered outputs, namely \underline{b}_{t-1} and α_t ; whereas the coefficient associated with the newly added atom is α_t . These equations allow for a “time-order” update after adding a new dictionary atom.

Note that in this algorithm computing the projection matrix P_{D_t} and LS filter operator Q_{D_t} are completely avoided and hence no matrix inversion is required. The updated atom is $\tilde{\underline{d}}_{k_t} = P_{D_{t-1}}^\perp \underline{d}_{k_t}$; however, we only need the filtered output \underline{b}_{t-1} since,

$$\begin{aligned} \tilde{\underline{d}}_{k_t} &= P_{D_{t-1}}^\perp \underline{d}_{k_t} \\ &= \underline{d}_{k_t} - P_{D_{t-1}} \underline{d}_{k_t} \\ &= \underline{d}_{k_t} - D_{t-1} Q_{D_{t-1}} \underline{d}_{k_t} \\ &= \underline{d}_{k_t} - D_{t-1} \underline{b}_{t-1} \end{aligned} \quad (3.8)$$

The steps of the entire algorithm are given in Table 3.1 in the order of computation. Additionally, the geometric interpretation of this learning and in particular \underline{r}_{t-1} and $\tilde{\underline{d}}_{k_t}$ is illustrated in Figure 3.1. In this figure, the observation vector \underline{y}_k is orthogonally projected onto the subspace

$\langle D_{t-1} \rangle^\perp$ to produce the residual vector \underline{r}_{t-1} . As demonstrated in the figure, the new residual \underline{r}_t can be expressed using (3.9) where \underline{r}_t and $P_{\underline{d}_{k_t}} \underline{y}_k$ are orthogonal, i.e. $\underline{r}_t^T P_{\underline{d}_{k_t}} \underline{y}_k = 0$.

$$\underline{r}_t = \underline{r}_{t-1} - P_{\underline{d}_{k_t}} \underline{y}_k = \underline{r}_{t-1} - \alpha_t \tilde{\underline{d}}_{k_t} \quad (3.9)$$

TABLE 3.1. Fast OMP Algorithm

| | |
|---|--|
| Task: For a signal vector $\underline{y}_k \in \mathbb{R}^N$, dictionary $D \in \mathbb{R}^{N \times K}$, and sparsity restriction τ , find a coefficient vector $\hat{\underline{x}}_k(t) \in \mathbb{R}^K$ with at most τ non-zero coefficients. | |
| Procedure: | |
| Iteration 1 | Iterations $t > 1$ to τ |
| $\underline{r}_0 = \underline{y}_k, D_0 = \mathbf{0}$ | $k_t = \operatorname{argmax}_j \underline{r}_{t-1}^T \underline{d}_j $ |
| $k_1 = \operatorname{argmax}_j \underline{r}_0^T \underline{d}_j $ | $\underline{b}_{t-1} = Q_{D_{t-1}} \underline{d}_{k_t}$ |
| $D_1 = \underline{d}_{k_1}$ | $\tilde{\underline{d}}_{k_t} = \underline{d}_{k_t} - D_{t-1} \underline{b}_{t-1}$ |
| $Q_{D_1} = (D_1^T D_1)^{-1} D_1^T$ | $\underline{q}_t = \frac{\tilde{\underline{d}}_{k_t}^T}{\ \tilde{\underline{d}}_{k_t}\ ^2}$ |
| $\underline{b}_0 = Q_{D_0} \underline{d}_{k_1} = \underline{0}$ | $\alpha_t = \underline{q}_t^T \underline{y}_k$ |
| $\tilde{\underline{d}}_{k_1} = \underline{d}_{k_1}$ | $D_t = [D_{t-1} \underline{d}_{k_t}]$ |
| $\underline{q}_1^t = \frac{\tilde{\underline{d}}_{k_1}^T}{\ \tilde{\underline{d}}_{k_1}\ ^2}$ | $Q_{D_t}^t = [Q_{D_{t-1}}^t - \underline{q}_t \underline{b}_{t-1}^T, \underline{q}_t]$ |
| $\alpha_1 = \underline{q}_1^T \underline{y}_k$ | $\hat{\underline{x}}_k^t(t) = [\hat{\underline{x}}_k^t(t-1) - \alpha_t \underline{b}_{t-1}^T, \alpha_t]$ |
| $\hat{\underline{x}}_k(1) = \alpha_1$ | $\underline{r}_t = \underline{r}_{t-1} - \alpha_t \tilde{\underline{d}}_{k_t}$ |
| $\underline{r}_1 = \underline{r}_0 - \alpha_1 \tilde{\underline{d}}_{k_1}$ | — |

Figure 3.2 shows the plots of computational time of the fast OMP and standard OMP as a function of the number of selected atoms. It can be seen that the computational requirements of the standard OMP method grow substantially quicker due to the matrix inversion and computations of the coefficients at each iteration compared to the recursive update equations in the fast OMP. For a relaxed sparsity and thus a low reconstruction error, this method is very beneficial.

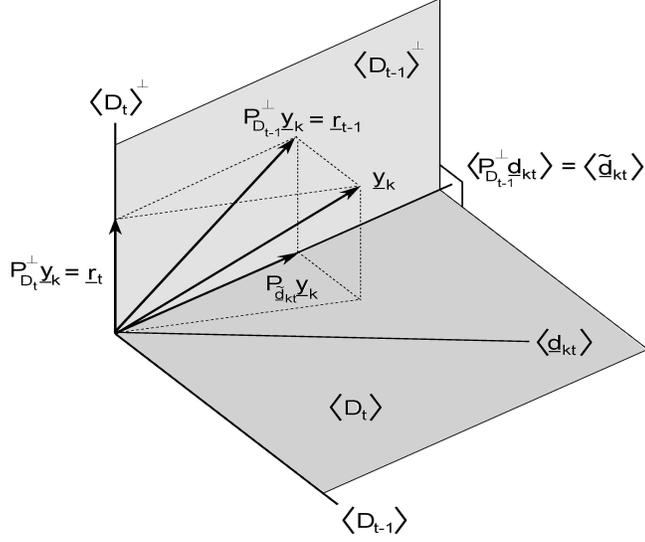


FIGURE 3.1. Geometric Interpretation of Fast OMP Learning.

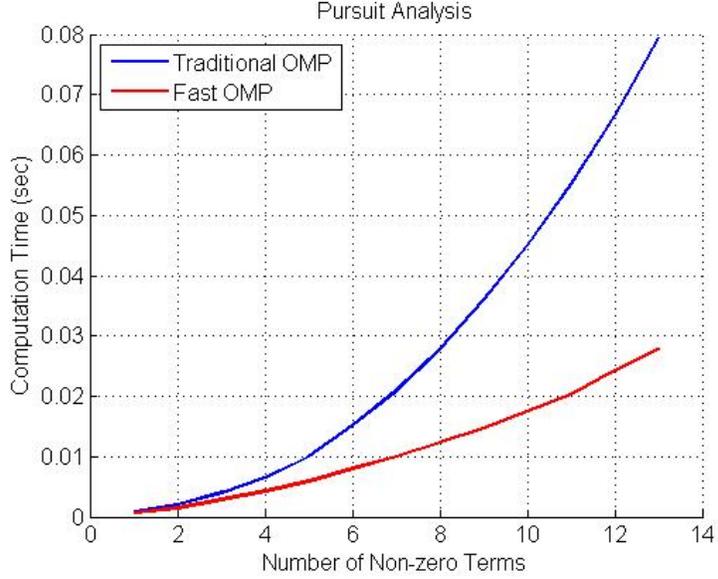


FIGURE 3.2. Timing Analysis of Pursuit Methods.

3.3. FAST DICTIONARY UPDATE USING RECURSIVE UPDATING PROCESS

When referencing the reduced K-SVD cost function $\|E_k^R - \underline{d}_k \underline{x}_R^k\|_F^2$ in (2.6), the objective is to find the combination of \underline{d}_k and \underline{x}_T^k such that $\underline{d}_k \underline{x}_R^k$ approximates E_k^R as closely as possible. This is normally solved by using SVD as explained in Section 2.2.2; however, this is computationally inefficient especially for atoms of large dimension as SVD returns all eigenvectors though we are

only interested in the most dominant. The following process is a common method used to return the most dominant eigenvector. The solution to \underline{d}_k can also be found by taking the partial derivative of the cost function with respect to \underline{d}_k in (2.6) and setting it equal to zero. This results in

$$\frac{\partial}{\partial \underline{d}_k} \left\| E_k^R - \underline{d}_k \underline{x}_R^k \right\|_F^2 = \frac{\partial}{\partial \underline{d}_k} [\text{tr}(E_k^R - \underline{d}_k \underline{x}_R^k)(E_k^R - \underline{d}_k \underline{x}_R^k)^t] = -2(E_k^R - \underline{d}_k \underline{x}_R^k) \underline{x}_R^{kT} = 0 \quad (3.10)$$

Solving for \underline{d}_k yields

$$\underline{d}_k = \frac{E_k^R \underline{x}_R^{kT}}{\underline{x}_R^k \underline{x}_R^{kT}} \quad (3.11)$$

which clearly requires having previously computed \underline{x}_R^k through OMP.

After using (3.11) to update the dictionary atom, the coefficients in \underline{x}_R^k must also be updated to reflect this change. If this does not happen, E_k on the next atom will be incorrect as a dictionary atom has been changed, but the coefficients do not match the change. This update can be done by taking the partial derivative of the cost function with respect to \underline{x}_R^k and setting it equal to zero, i.e.

$$\frac{\partial}{\partial \underline{x}_R^k} \left\| E_k^R - \underline{d}_k \underline{x}_R^k \right\|_F^2 = -2 \underline{d}_k (E_k^R - \underline{d}_k^T \underline{x}_R^k) = 0 \quad (3.12)$$

Solving for \underline{x}_R^k gives the update equation

$$\underline{x}_R^k = \frac{\underline{d}_k^T E_k^R}{\underline{d}_k^T \underline{d}_k} \quad (3.13)$$

The two coupled equations (3.11) and (3.13) are iterated (5 times here) to yield $\hat{\underline{x}}_R^k$ and \underline{d}_k instead of doing SVD. This method is much faster than the original K-SVD. To start the process, \underline{d}_k is initialized with (3.11). Note that it is required that \underline{x}_R^k has been previously calculated by OMP; however, this has already been done when entering the dictionary update phase as the entire sparse matrix X is available after the sparse pursuit phase. After initializing \underline{d}_k , \underline{x}_R^k can be updated for the first time with (3.13). It has been shown in [19] that convergence is assured as the reconstruction

error will monotonically decrease as the dictionary update takes place. When applying only the dictionary update method in this section, the K-SVD training time can be cut in half.

3.4. CONCLUSION

This chapter has shown methods used to replace the two phases in K-SVD: the sparse coding phase and the dictionary update phase. First, a series of light-weight recursive equations were developed to replace the matrix inversion typically used in the LS solution of OMP in the sparse coding phase. Not only do these equations prevent the matrix inversion, but due to the recursive update, they also prevent costly recomputation of commonly used factors within the algorithm. Then, the second phase of K-SVD was addressed in a faster dictionary update phase proposed in [19]. This dictionary update replaces SVD with the coupled update equations from the partial derivatives of the cost function allowing training to be completed in under half of the original time.

With the motivation of reducing overall computation time without sacrificing accuracy, the methods discussed in this chapter have not only speed up computation time but they have also offered a new and interesting perspective on previously existing methods leading to the faster algorithms. Increasing the speed of OMP as a function of sparsity level has two benefits. First of all, it can provide efficient OMP even when sparsity is relaxed due to the efficient “time-order” recursive updates. Second, due to the reduced computational/structural complexity, OMP can be used to sparsely represent many signals very quickly allowing in-situ operation (e.g., on DSP boards) as the target detector is applied to live data. The major limitation of all sparse coding algorithms including OMP, is the accuracy of the resulting sparse signal is heavily determined by the basis dictionary. A dictionary that does not contain the dominant features of the input signals will not be able to accurately represent the signal. As can be seen in Chapter 4, this can be used to the advantage of the detector.

CHAPTER 4

APPLICATION TO UNDERWATER TARGET DETECTION USING SONAR IMAGERY

4.1. INTRODUCTION

With the end goal of testing the effectiveness and applicability of the developed sparse signal-based detector, the learned dictionaries have been applied to the important problem of underwater mine-like detection. While this thesis does not focus on the neutralization and removal of mines, many options have been explored in [22] for this problem including blow-in-place operations and dangerous retrieval operations. Mine neutralization and removal are critical, but this is only half of the battle. Finding mines quickly and efficiently in large scale searches is a key step in the removal process. The focus of this work is therefore on the detection and localization of mine-like objects in sonar imagery gathered from the Small Synthetic Aperture Minehunter (SSAM) [2]. This problem is very difficult as these objects must be discriminated from clutter both of natural (e.g., coral, rocks, etc.) and man-made sources (e.g., scrap metal, oil barrels, lobster traps, etc.). As stated previously, the mines can be proud on the surface of the ocean floor, partially, or fully buried.

This chapter presents the application of a K-SVD-based detector developed in Chapter 2 to this detection problem. First, the dual-channel extension of K-SVD-based detector is shown granting the ability to incorporate high frequency (HF) and broadband (BB) sonar imagery into a single dictionary used as a basis for sparse coding. The dictionaries created through K-SVD are applied to two different sonar datasets as a way to validate the concept. Receiver operating characteristic (ROC) curves are presented in a discussion to highlight benefits and shortcomings of the algorithm. Modification to further reduce false alarms, including ROI grouping, are also discussed and implemented in a series of post-processed sonar images.

This chapter is organized as follows: Section 4.2 describes the datasets and their properties used in this chapter. Further, it shows how the proposed K-SVD algorithm can be used to generate dictionaries for locating the mine-like objects, and how they are distinguished from competing clutter. Section 4.3 presents the results of the sparse signal based detector. Finally, Section 4.4 gives concluding remarks and discussions.

4.2. DATA DESCRIPTION AND MULTIMODAL K-SVD

4.2.1. SSAM I DATASET 1

The dataset used in this study has been collected by the SSAM I system, developed under the support of the Office of Naval Research (ONR). The SSAM I system collects pairs of coregistered HF and BB sonar images for the purposes of underwater target detection and classification. The first sonar imagery dataset used in this study is composed of 122 pairs of HF and BB images containing 77 targets of varying sizes, shapes, and ranges from the platform. Due to the way in which sonar pings are emitted and received, targets are normally characterized (cross-track) by a highlight followed by a shadow; however, this characteristic can also be shared with common seafloor clutter such as coral and rocks making automatic detection and classification difficult. The difficulty can be seen in the HF image in Figure 4.1. A target is present in this image located at $x = 1400$ (cross-track) and $y = 1000$ (along-track). The target can be detected by an expert sonar operator as the shadow can be seen to the right of the highlight. In this particular case, the BB sonar image, as seen in Figure 4.2, doesn't add any new information to aid the operator. In general, the BB images do not have the sharp details seen in the HF images. Thus, they contain less clutter effects in contrast to the HF images. This study will test if including the coregistered BB channel indeed aids in the detection and localization of the targets. If the target had fallen at approximately $y = 1100$, the shadow would be obscured by the natural bottom clutter and detection would have been even more difficult as is the case with many of the other targets in this dataset.

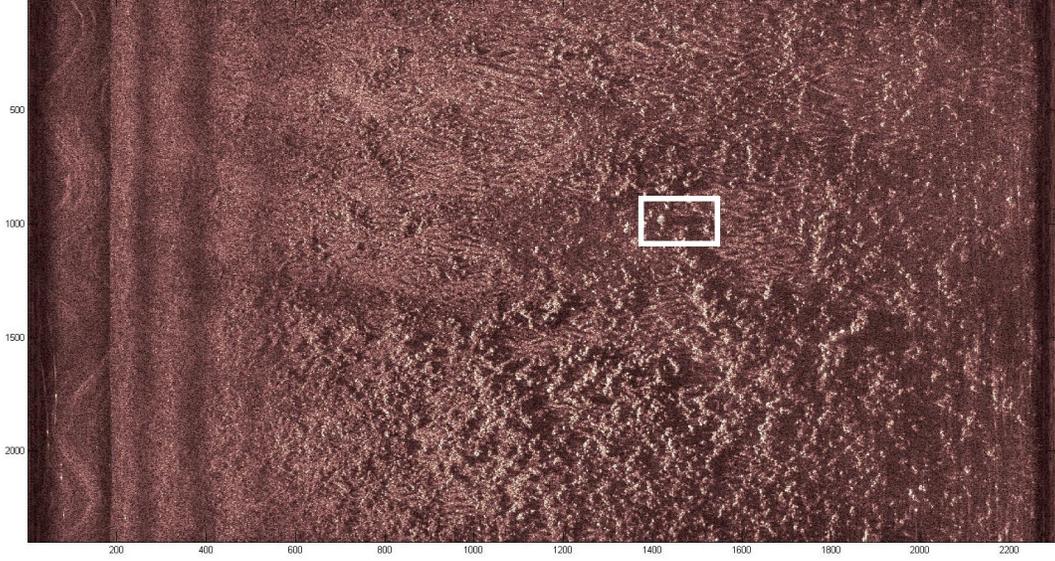


FIGURE 4.1. Difficult High Frequency Image Containing High Density of Bottom Clutter.

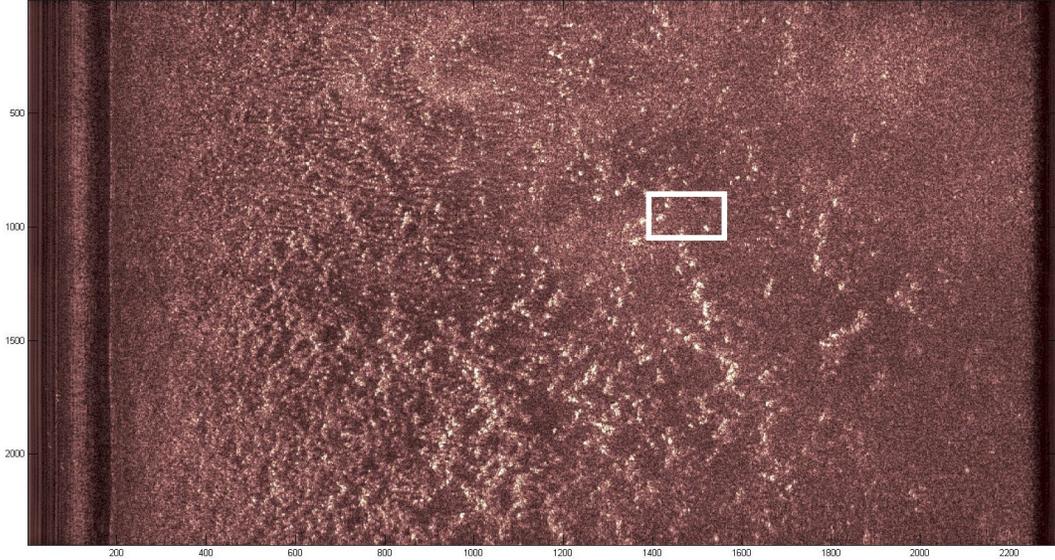


FIGURE 4.2. Difficult Broadband Image Containing High Density of Bottom Clutter.

The detection algorithm is applied to smaller Regions of Interest (ROIs) of size $[80 \times 140]$ pixels by sweeping a window across the entire image. The ROIs overlap by 9 pixels in both the cross-track and along-track directions. The k^{th} pair of coregistered HF and BB ROIs are vectorized and concatenated to form the composite vector $\underline{z}_k = [\underline{y}_k^T(hf) \ \underline{y}_k^T(bb)]^T \in \mathbb{C}^{2N}$ where $\underline{y}_k(hf) \in \mathbb{C}^N$ represents the HF ROI and $\underline{y}_k(bb) \in \mathbb{C}^N$ represents its BB counterpart. After converting to a column vector, N is equal to the number of pixels in the ROI i.e. $N = 80 * 140 = 11200$. The

dataset is complex-valued as it is the output of the k-space or wavenumber beamforming algorithm [23], [24]. When using signals of this format and assuming a properly formatted dictionary, each pair of ROIs can be reconstructed according to the relationship

$$\underline{z}_k = \begin{bmatrix} \underline{y}_k(hf) \\ \underline{y}_k(bb) \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} \underline{x}_k^{(1)} \\ \underline{x}_k^{(2)} \end{bmatrix} \quad (4.1)$$

where D_{11} and D_{22} are the inter-channel dictionaries, D_{12} and D_{21} are the intra-channel dictionaries, and $\underline{x}_k^{(1)}$ and $\underline{x}_k^{(2)}$ are the sparse representations of $\underline{y}_k^T(hf)$ and $\underline{y}_k^T(bb)$; not individually since $\underline{y}_k(hf) = D_{11}\underline{x}_k^{(1)} + D_{12}\underline{x}_k^{(2)}$ and $\underline{y}_k(bb) = D_{21}\underline{x}_k^{(1)} + D_{22}\underline{x}_k^{(2)}$.

This dual-channel (HF and BB) detection strategy is compared to a situation where one employs the HF channel only in which case $y_k(hf) = D_{11}\underline{x}_k^{(1)}$. Note that the dictionary D_{11} will not be the same for these two scenarios, though they are exposed to the same training data. A dual-channel dictionary can be created simply by formatting the K-SVD input data matrix Y such that $Y = [\underline{z}_1 \cdots \underline{z}_M]$ where M is the number of training samples to use.

For this detection problem, two dictionaries (one for target and one for non-target) are created using the K-SVD algorithm discussed in Chapters 2 and 3. The two dictionaries will be created using data matrices Y_1 containing only target samples and Y_2 containing only non-target samples. This setup essentially creates two mutually exclusive dictionaries equipped to represent only a certain type of data, and it exploits the fact that the dictionaries cannot represent data they have not been exposed to via training. It can be expected that a target reconstructed from the non-target dictionary will be very poor, while the reconstruction of a target will be accurate for the target dictionary. Conversely, a non-target reconstructed from a non-target dictionary will be accurate, while the reconstruction of a non-target from the target dictionary will be poor. This relationship can be seen in Figure 4.3. The input target seen on the left side of the figure is sparsely represented as \underline{x}_1 and \underline{x}_2 using the dictionaries D_1 (target) and D_2 (non-target). After

displaying the reconstructed images, $D_1\underline{x}_1$ and $D_2\underline{x}_2$, it can be seen that the input image is most similar to the reconstructed image when sparsely representing the input image using the target dictionary D_1 . This similarity can be represented mathematically by simply looking at the norm of the reconstruction error which is lower for dictionary containing images of the same class as the input, i.e. the dictionary D_1 for this case.

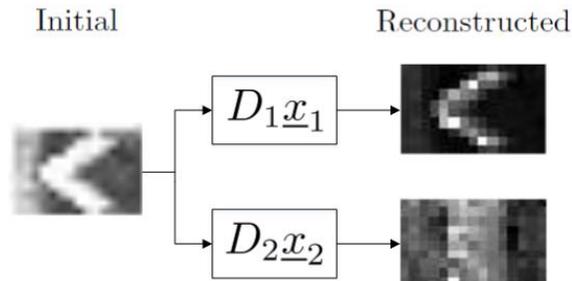


FIGURE 4.3. Exclusive Dictionary Reconstructions for a Target Sample.

Both the single-channel (HF only) and dual-channel dictionaries have been generated using a subset, 16 out of the total 77 targets in the dataset, as well as a selected set of background ROIs as training sets. A set of 30 background ROIs are used in the single-channel dictionary while only 15 background ROIs are used in the dual-channel dictionary. Figure 4.4 shows the selected training samples (HF only) of the K-SVD algorithm and the corresponding atoms after K-SVD for both dictionaries. Only five atoms are shown for these training images to keep the figure simple; however, 13 atoms are generated ($K = 13$) using K-SVD. As expected the target dictionary D_1 contains atoms that capture different distinct features of the selected target samples. Interestingly, the atoms in D_2 appear to be a mixture of random noise and odd patterns of rocks and coral. This is due to the selected non-target training images. The displayed atoms are those that are needed to minimize the reconstruction error within the training set.

After the training and dictionary creation, for each ROI (e.g. k^{th}) extracted from the entire image, the data is represented sparsely based on both dictionaries. The vectors \underline{x}_1 and \underline{x}_2 will be generated sparsely from the dictionaries D_1 and D_2 using fast OMP as discussed in Section

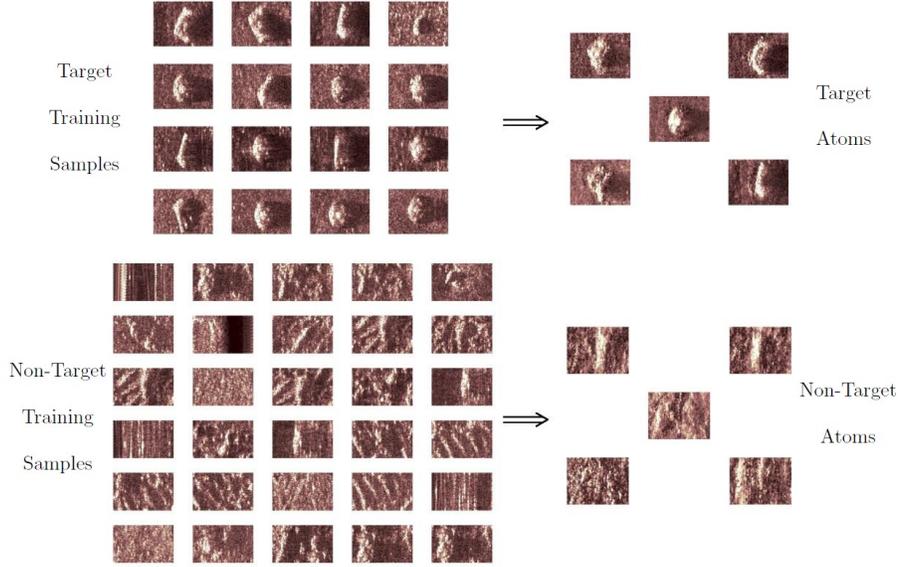


FIGURE 4.4. Training Samples and the Corresponding K-SVD Atoms.

3.2. The reconstruction errors are then computed for each sparse composition and used to form an analog of likelihood ratio $\Lambda(z_k)$ for each ROI z_k in the sonar image, i.e.

$$\Lambda(z_k) = \frac{\|z_k - D_2 x_{k,1}\|}{\|z_k - D_1 x_{k,2}\|} \underset{\text{NT}}{\overset{\text{T}}{\geq}} \gamma \quad (4.2)$$

The value of $\Lambda(z_k)$ is then compared against a preselected threshold γ to determine if the k^{th} ROI contains a target or a non-target. The complete testing phase can be seen in Figure 4.5. Note that the selection of γ is purely experimental and is heavily dependent upon the environment and how well the dictionaries are able to represent the data in each ROI. In a heavily cluttered image, γ may be much lower to prevent many false alarms, while in an image with low clutter γ can be larger. In general γ is chosen to provide a desired P_{FA} [25], [26].

4.2.2. SSAM I DATASET 2

The second dataset used in this study has also been collected by the SSAM I System. This dataset is composed of 458 pairs of HF and BB images containing only 68 targets of varying sizes, shapes, and ranges. In this dataset, the amount of targets are vastly outweighed by the amount of

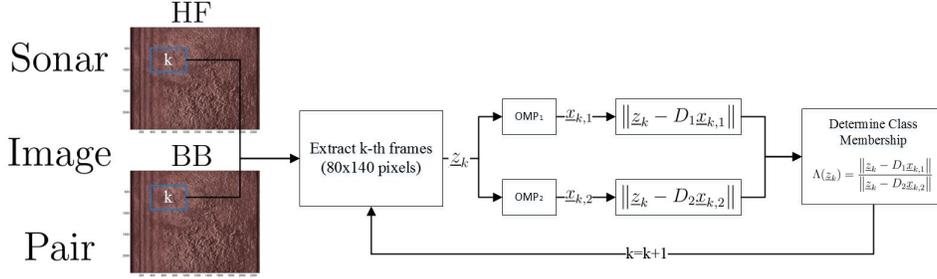


FIGURE 4.5. Diagram of K-SVD Testing Phase.

clutter giving the added benefit of simulating a wide scale search. As an added test, the detector discussed previously will not be trained on this dataset. This gives the added insight as to their robustness against environment versatility.

Compared to dataset 1, this particular dataset has been collected over three different environments adding an extra difficulty. Two of the environments are relatively flat with features similar to dataset 1. The third environment is very complex as it contains a mixture of seagrass and sand formations. The particular species of seagrass is unique to this region, and it is an element foreign to dataset 1. It is expected that the dictionaries generated by dataset 1 will have difficulty representing these elements.

4.3. DETECTION RESULTS

4.3.1. RESULTS ON DATASET 1

Figure 4.6 compares the ROC curves of the detection strategies i.e. single-channel (HF only) and dual-channel (HF and BB). In addition, they are also compared to the multichannel frequency-based detector [13] which looks for high levels of coherence between each pair of HF and BB ROIs. In this detection strategy it was shown that coherence patterns are different when ROI contained mine-like objects compared to those that did not contain mine-like objects. From the ROC curve, one can see that the dual-channel K-SVD algorithm exhibits an improvement in detection performance over its single-channel counterpart as well as the multi-channel coherence detector for this dataset with a knee point of $P_D = 0.92$ and $P_{FA} = 0.08$. In this dataset, the multi-channel coherence based

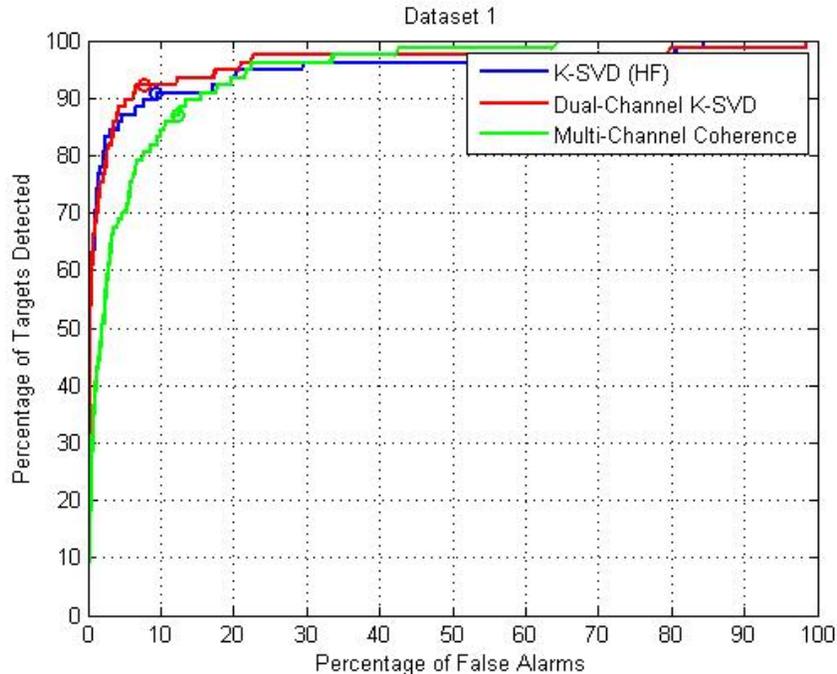


FIGURE 4.6. Receiver Operating Characteristic Dataset 1.

detector is slightly below the K-SVD based detectors. The knee point of the ROC for this detector is at $P_D = 0.87$ and $P_{FA} = 0.13$. Compared to the multi-channel frequency domain coherence based detector [13], the single-channel K-SVD detector offers an improvement in the detection rate while decreasing the false alarm rate with the knee point at $P_D = 0.91$ and $P_{FA} = 0.09$. Even though the BB images offer poor visual quality, the dual-channel K-SVD detector offers the best performance of the three detectors.

Next, we show how we can improve on these results using the same dictionaries by reducing the number of false alarms. Further investigation of the results of K-SVD-based detection leads to a procedure to accomplishing this goal. The first observation to make is the previous ROC curves were generated using single ROI analysis where every ROI is independently analyzed to determine if it contains a target or not. In other words, neighboring ROIs do not impact the decision-making process. Figure 4.7 shows a low cluttered HF image with five prominent targets, four of which were successfully detected by the system (as indicated by the dark blue boxes). The green boxes and

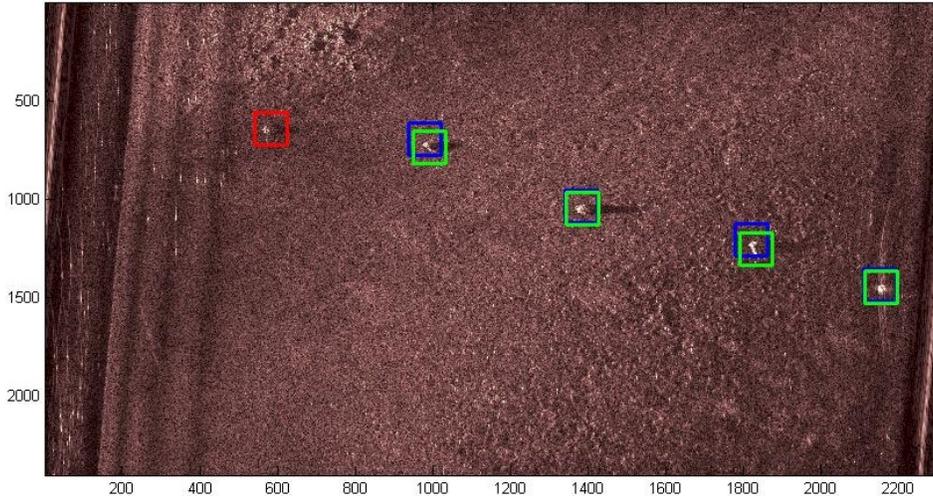


FIGURE 4.7. A HF Sonar Image And Detected Contacts.

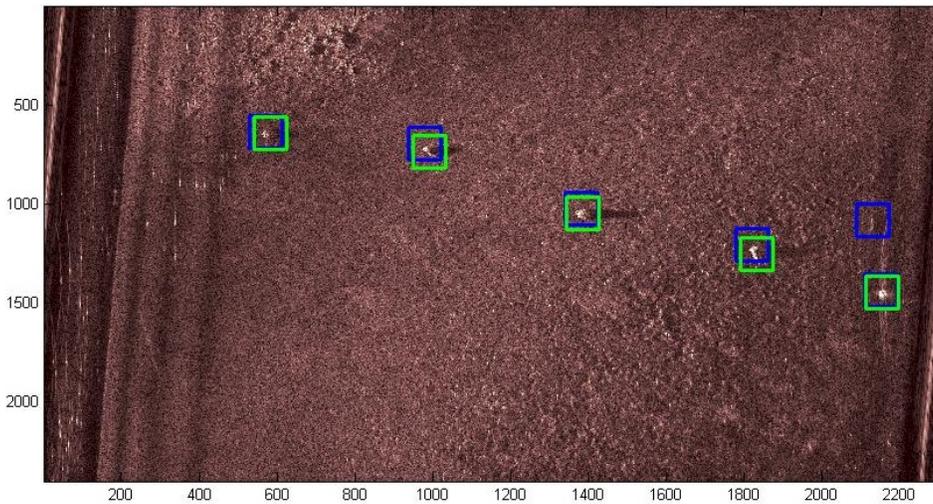


FIGURE 4.8. A HF Sonar Image And Detected Contacts - Updated Threshold

red boxes were drawn on afterward to show the actual locations of the targets based upon divers' markings. A green box indicates a target was found by the system at that location, and a red box indicates a target was missed at that location. An obvious solution for detecting the missed target is to modify the threshold; however, this will clearly increase the number of false alarms as seen in Figure 4.8. False alarms in this application are critical as any alarm must be treated as an actual threat. In some cases, an alarm may involve a team of divers inspecting an underwater site, or it may just take validation by an expert operator. As in all applications, it can be time consuming and costly if there are a large number of false alarms.

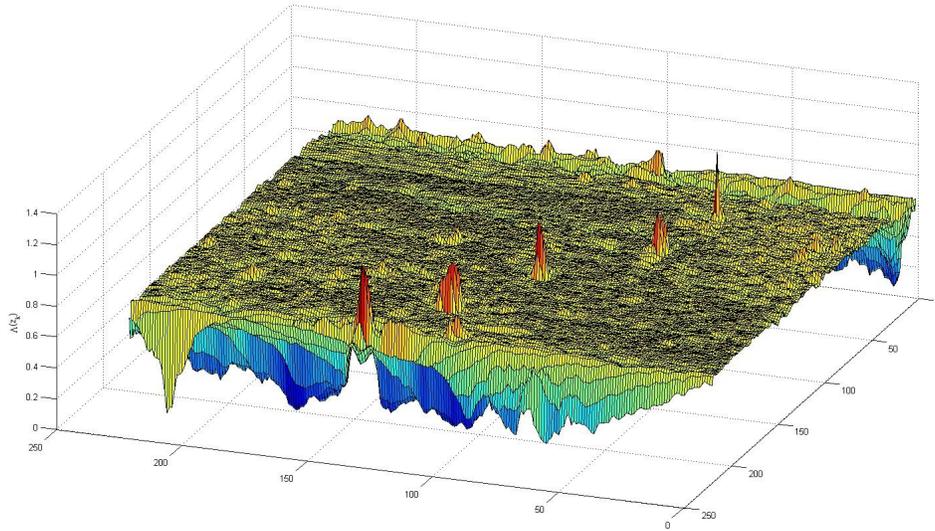


FIGURE 4.9. 3D Analysis of Likelihood Ratios $\Lambda(\underline{z}_k)$.

Instead independently analyzing each ROI to determine the presence of a target, it is possible to employ a grouping requirement where multiple nearby ROIs must be above a threshold within a certain area. Neighboring ROIs can be used to determine if a specific point on the seafloor contains a target. This strategy relies on the assumption that a target will be picked up by multiple ROIs within a small area. This is often the case as seen in Figure 4.9. When the window is centered around a target, the likelihood ratio experiences a sharp peak. As the window moved away from the target, the likelihood ratio decays.

It is possible to use these peaks to locate the targets. If a simple grouping technique is employed (such as requiring a minimum number of hits within an area to declare a target), the ROC curve seen in Figure 4.10 can be generated. The dual-channel K-SVD dictionaries are able to detect 100% of the targets with only 1.2% false alarms (a knee point at $P_D = 0.986$ and $P_{FA} = 0.014$). Interestingly, this is achieved when enforcing the minimum group size of 2 elements. In other words, at least two ROIs must agree on a target within some small area (approximately half the size of an ROI) for a target to be declared. Compared to the results when not enforcing this condition, the dual-channel K-SVD detection method experiences an increase in detection rate of 6.43% while

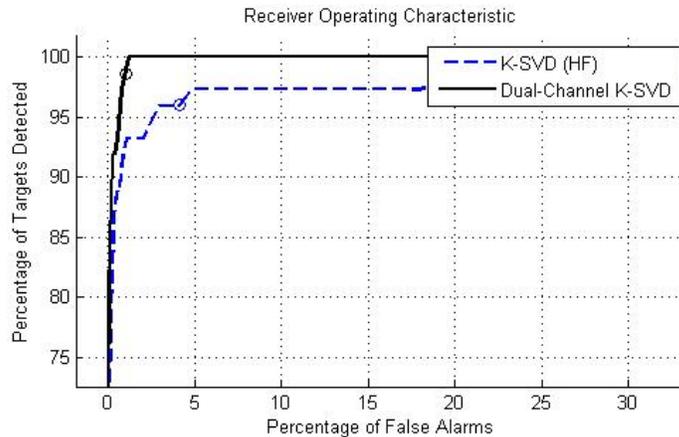


FIGURE 4.10. Receiver Operating Characteristic Dataset 1 using ROI Grouping.

simultaneously reducing the false alarm by 6.74%. In this same study, the single-channel K-SVD is also able to achieve very high detection results (knee point at 95.9% detected targets with only 4.1% false alarms) using the same minimum group size. The single-channel K-SVD experiences an increase of 4.99% in detection rate with a reduction of false alarm rate by 4.27%. Other grouping techniques are possible including other minimum group sizes, using Gaussian shapes, or integrating the volume above the threshold. These techniques may further reduce the false alarm rate but they were not pursued due to time constraints. Nevertheless, as only the addition of the BB channel has been included in the dual-channel K-SVD, it is clear that the simple grouping method played a major role in the success of this algorithm.

When analyzing a more difficult image, such as the HF image seen in Figure 4.1, a large number of false alarms are generated to detect the target (shown in green). Figure 4.11 shows the false alarms generated when detecting the single target in this image (shown in green) as a result of employing the grouping method discussed previously. While this is far fewer false alarms than without the grouping method, as seen in Figure 4.12, better results may be achieved using other advanced grouping methods as mentioned before.

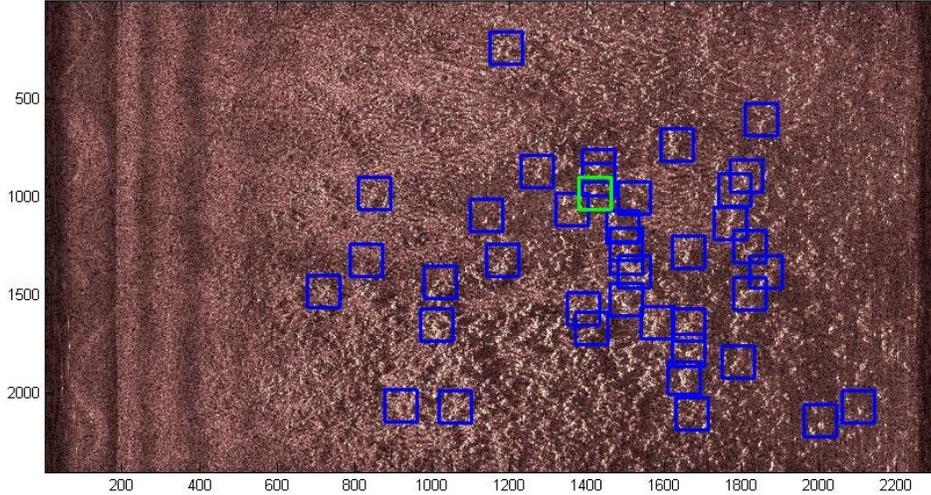


FIGURE 4.11. A Difficult HF Sonar Image With Detected Contacts After Grouping.

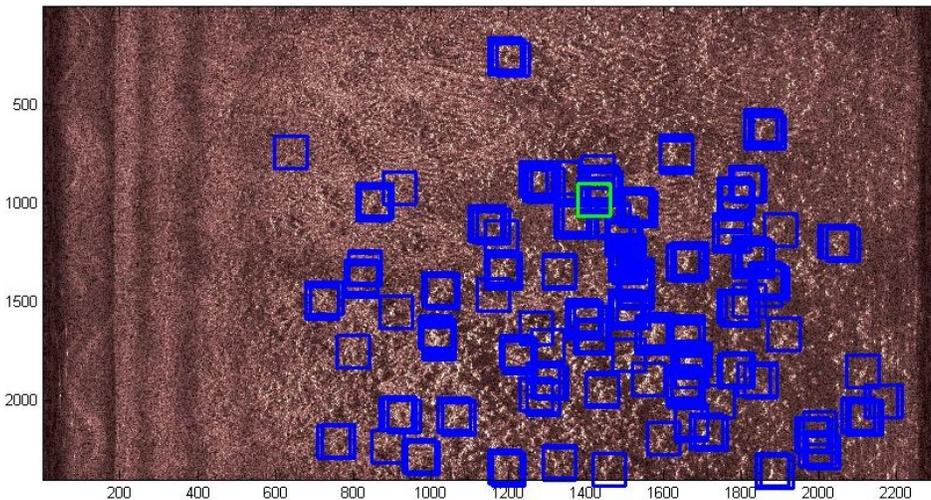


FIGURE 4.12. A Difficult HF Sonar Image With Detected Contacts - Without Clustering.

4.3.2. SSAM DATASET 2

The K-SVD dictionaries generated using dataset 1 will be used here to test the robustness of the designed detector by applying them to dataset 2. The system will then be benchmarked against the multichannel frequency-based detector [13]. Both single-channel (HF only) and the dual-channel detection strategies were applied to dataset 2. As can be seen in the ROC curve in Figure 4.13, the dual-channel K-SVD algorithm did not perform as well as the single-channel K-SVD and multi-channel coherence based detectors on this dataset. The knee point for the multi-channel coherence detector is at $P_D = 0.86$ and $P_{FA} = 0.14$, while the single-channel K-SVD detector yields the knee

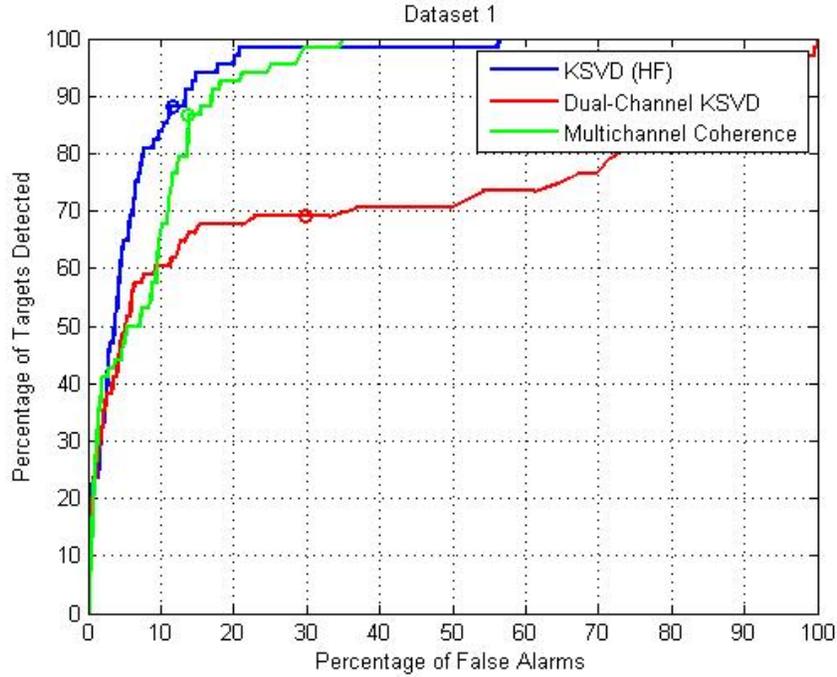


FIGURE 4.13. Receiver Operating Characteristic Dataset 2

point performance of $P_D = 0.88$ and $P_{FA} = 0.12$. The dual-channel K-SVD detector performs very poorly in this dataset with a knee point at only $P_D = 0.69$ and $P_{FA} = 0.31$.

The main reason the dual-channel K-SVD performed poorly on this dataset is the addition of the BB channel makes discrimination much more difficult. Figure 4.14 depicts the BB image in a HF BB sonar pair. Within this image, very distinct highlights can be observed on many of the clutter features. In the first dataset, this was a common feature of a target; however, this characteristic is common among clutter in this dataset thus leading to a high false alarm rate as seen in the HF version of the same image Figure 4.15. Notice only those ROIs with prominent highlights in both BB and HF images have been declared as targets. If the broadband channel is neglected as in the single-channel K-SVD detector, the detection rate is much higher as observed in the ROC curve.

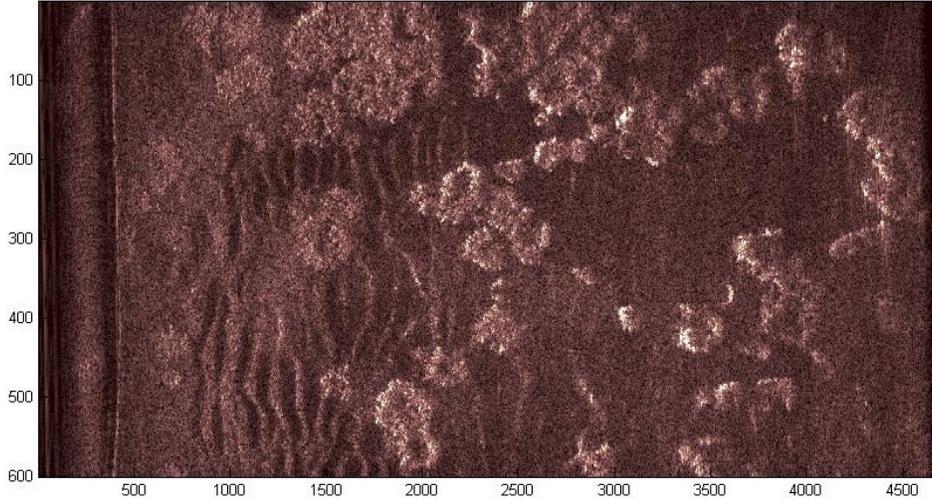


FIGURE 4.14. Broadband Image in Dataset 2.

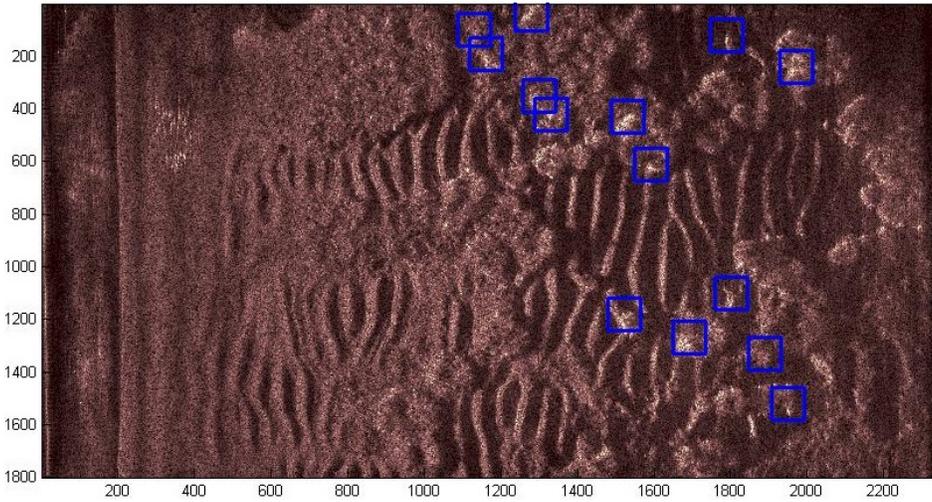


FIGURE 4.15. High Frequency Image in Dataset 2 with Detected Contacts.

4.3.3. TABULATED RESULTS FOR DATASETS 1 AND 2

On the first dataset, the dual-channel K-SVD algorithm exhibited a 1.3% detection rate improvement over the single-channel K-SVD strategy with a false alarm reduction of 1.5%. Compared to the multi-channel coherence detector, the dual-channel K-SVD detector exhibited a 5.2% detection rate improvement with a false alarm reduction of 4.77% as seen in Section 4.3.1. As for the second dataset, the dual-channel K-SVD was unable to accurately represent both HF and BB images and therefore performed poorly. The single-channel K-SVD algorithm, on the other hand,

exhibits a 1.48% detection rate increase over the multi-channel coherence detector while simultaneously decreasing the false alarm rate by 2.05% when not employing any clustering methods. Final results can be seen in Table 4.1.

TABLE 4.1. Detection Results without Grouping

| | Dataset 1 | | Dataset 2 | |
|----------------------|-----------|----------|-----------|----------|
| | P_D | P_{FA} | P_D | P_{FA} |
| Single-Channel K-SVD | 0.91 | 0.09 | 0.88 | 0.12 |
| Dual-Channel K-SVD | 0.92 | 0.08 | 0.69 | 0.31 |
| Coherence | 0.87 | 0.13 | 0.86 | 0.14 |

Grouping methods have been developed and implemented to increase the detection rate of these targets while retaining a low false alarm rate. After employing these techniques on dataset 1, the dual-channel K-SVD dictionaries were able to achieve a detection rate of 98.64% (an increase of 6.43% compared to dual-channel K-SVD when not using these clustering methods) while simultaneously reducing the false alarm to 1.03% (reduction of 6.74%). Additionally, single-channel K-SVD achieved a detection rate of 95.9% detection with only 4.1% false alarms (an increase of 4.99% in detection rate and a reduction of false alarm rate by 4.27%). Interestingly, when applying these grouping methods to dataset 2 substantial gains are observed indicating that even the dual-channel K-SVD detector can be generalized to outside datasets if multiple ROIs are required to vote on a target. These results can be seen in Table 4.2

TABLE 4.2. Detection Results with Grouping

| | Dataset 1 | | Dataset 2 | |
|----------------------|-----------|----------|-----------|----------|
| | P_D | P_{FA} | P_D | P_{FA} |
| Single-Channel K-SVD | 0.96 | 0.04 | 0.96 | 0.04 |
| Dual-Channel K-SVD | 0.99 | 0.01 | 0.95 | 0.05 |

4.4. CONCLUSION

Three different detectors were tested against two sonar imagery datasets. The detectors include: single-channel K-SVD, dual-channel K-SVD, and multi-channel frequency-based coherence detectors. Both K-SVD-based methods were trained with ROIs selected from the first dataset. The dual-channel K-SVD algorithm outperformed the other two detection strategies on dataset 1. The single-channel K-SVD algorithm also exhibited a detection rate improvement over the multi-channel coherence detector while simultaneously decreasing the false alarm rate. As for the second dataset, the dual-channel K-SVD method falls short due to the inability to accurately represent both HF and BB images, and the high false alarm rate caused by the distinct highlights in the BB images. The single-channel K-SVD algorithm, on the other hand, outperformed both of the other methods showing better generalization to outside datasets compared to the dual-channel K-SVD detector.

In general, all three detectors face difficulty when targets are occluded by natural underwater clutter (partially/fully buried, etc.). Further, targets that are very close in range to the sonar emitter are corrupted in the beamforming process and become very difficult to detect as they are not characterized by the distinct highlight/shadow pair. Detection of these difficult targets normally come at the cost of many false alarms. An ROI grouping method has been developed and implemented to increase the detection rate of these targets while retaining a low false alarm rate when using the same dictionaries with significant results.

This chapter has illustrated that though the K-SVD dictionaries were trained in one dataset, they can be generalized to other datasets to achieve comparable detection results. When employing the grouping methods, both the single-channel and the dual-channel K-SVD-based detectors were found to performed very well. These results indeed indicate that the K-SVD-based detection is a powerful tool for difficult detection problems. The new fast OMP has sped up the detection rate allowing online operation. Numerous additional improvements have been proposed in Chapter 5.

CHAPTER 5

CONCLUSION AND SUGGESTIONS FOR FUTURE WORK

5.1. CONCLUSION AND DISCUSSION

The problem of underwater mine detection from dual-channel sonar images is complicated as the mines are subjected to spatially random bottom clutter, widely varying environmental conditions, and the mines can be of varying size, shape, and orientation. This work led to the improvement of the existing K-SVD algorithm by utilizing preexisting tools in the adaptive filter area and showed its application to the important problem of underwater mine detection. First, a new fast OMP method was presented based on orthogonal projection updating used in fast transversal filters [27]. As the fast OMP method is void of matrix inversions, this algorithm benefits as sparsity is relaxed compared to the standard OMP method. Then, the dictionary update phase was replaced by a faster update method which replaces SVD with pair of coupled updating coupled equations which extracts only the dominant eigenvector. These two methods have increased the speed of training and allow in-situ operation of this detector on DSP boards or other dedicated hardware.

The success of the K-SVD-based detection algorithm is attributed to the fact that the training data essentially informs the detector of the types of signals associated with targets and non-targets. Compared to the multi-channel frequency domain coherence detector in [13], which simply looks for ROIs exhibiting certain cross spectral structure rather than matching a known model, the K-SVD-based detector tends to yield not only lower false alarm rate, but also a higher target detection rate. This benefit is also one of the most significant downfalls of the K-SVD-based detector as it requires the availability of enough training data to extract relevant atoms. In this application, K-SVD is possible due to the readily available data; however, this may not always be the case. While it is possible to generalize to other datasets after training, if training data is not available to create the initial dictionaries this method will fail while other GLRT-based detectors such as the

multi-channel frequency domain coherence detector method may still succeed as it does not depend on training data.

In the first dataset, the dual-channel K-SVD detector outperformed the single-channel K-SVD detector due to the added information in the broadband channel. Further, the single-channel K-SVD detector achieved superior results compared to the multi-channel coherence detector due to the presence of information about the targets and clutter in the dictionaries. While the dual-channel K-SVD detector performed poorly on the second dataset due to the inability to accurately represent both HF and BB channels of the sonar image, this deficiency was corrected by employing a grouping method where multiple neighboring ROIs are used make a decision.

5.2. FUTURE WORK

Though the K-SVD-based detector featured in this thesis has shown excellent results in the application of detecting underwater mines from dual-channel sonar images, this thesis has laid the foundation for several improvements that can be pursued in the future. These include:

- **In-situ dictionary update:** Due to the slow learning speed of K-SVD it has previously been limited to offline learning scenarios. As a result of fast OMP and the faster dictionary update presented in this thesis, in-situ dictionary updates are now possible. In general, dictionaries will be retrained and the system will learn as it is exposed to unfamiliar environments. A set of conditions must be set to determine a number of actions. The first task is given some ROI, autonomously determine if the sample should be neglected or learned as a target or a nontarget. This involves updating the dictionary atoms to better represent the data. The ROI would simply be added to either Y_1 or Y_2 and fast K-SVD would be used to retrain either D_1 or D_2 respectively. It may be necessary to increment the number of allowable atoms in the dictionary as the dictionary may be unable to accurately hold additional information present in the new training sample. If this is the case, to incorporate a larger dictionary, it may be

required to slightly relax the sparsity restriction of OMP in the target detector. Naturally, when the detector encounters a new environment, the detector should rapidly adapt and incorporate the presence of new data while retaining the previously learned information.

- **Kernel K-SVD:** One limitation of K-SVD is it assumes linearly separable data; however, in many difficult scenarios this is not the case. Kernelization is a common tactic for combating non-linearly separable data because when mapping into hyperspace, data often becomes linearly separable. Kernel K-SVD was discussed in [28], with the aim to pose the modified cost function in the higher dimensional feature space while solving the minimization problem in the original space with the kernel trick. As X becomes sparse on the basis of a kernelized dictionary, a kernelized OMP method must also be employed. After training the kernel dictionaries, the same procedure employed in this thesis would be used to detect the targets in the sonar images where each ROI is represented based on two dictionaries and the likelihood ratio is compared against a threshold.
- **Extension to target classification:** The present framework can be extended to include two-class target classification assuming the presence of labeled data. Instead of grouping training data into two sets Y_1 and Y_2 for mine-like and non-mine-like respectively, matrices will be constructed based upon those samples that pass the detector to assign final class labels with high confidence and reject all the false detections as much as possible.
- **Other applications:** This algorithm is not limited to sonar data. K-SVD-based detection can also be applied to target detection or classification from radar data. It can also be used in other areas including facial recognition, iris recognition, or other multi-sensor biometric classifiers.

BIBLIOGRAPHY

- [1] P. Chapple, *Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations*. Edinburgh, Australia: Maritime Operations Division, 2008.
- [2] D. Brown, D. Cook, and J. Fernandez, “Results from a small synthetic aperture sonar,” in *OCEANS 2006*, Sept 2006, pp. 1–6.
- [3] S. Perry, *Applications of Image Processing to Mine Warfare Sonar*. Melbourne, Australia: Aeronautical and Maritime Research Laboratory, 2000.
- [4] S. Reed, Y. Petillot, and J. Bell, “Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information,” vol. 151. Proc of IEEE Radar Sonar Navigation, Feb 2004, pp. 48–56.
- [5] G. Dobeck, J. Hyland, and L. Smedley, “Automated detection/classification of sea mines in sonar imagery,” vol. 3079. Proc SPIE, 1997, pp. 90–11.
- [6] W. Guo and W. Szymczak, “Multiresolution neural networks for mine detection in side scan sonar images,” vol. 3392. Proc SPIE, 1998, pp. 297–305.
- [7] J. Tucker, “Coherence-based underwater target detection for side-scan sonar imagery,” Master’s thesis, Colorado State University, 1999.
- [8] A. Pezeshki, M. Azimi-Sadjadi, L. Scharf, and M. Robinson, “Underwater target classification using canonical correlations,” vol. 4. Proc of MTS/IEEE Oceans 2003 Conference, Sept 2003, pp. 1906–1911.
- [9] A. Pezeshki, M. R. Azimi-Sadjadi, L. L. Scharf, and M. Robinson, “Underwater target classification using canonical correlations,” vol. 4. Proc of MTS/IEEE Oceans 2003 Conference, September 2003, pp. 1906–1911.
- [10] A. Pezeshki, M. R. Azimi-Sadjadi, and L. L. Scharf, “Undersea target classification using canonical correlation analysis,” vol. 32, no. 4. Proc of IEEE of Journal of Oceanic Engr., Oct

- 2007, pp. 948–955.
- [11] J. Cartmill, N. Wachowski, and M. R. Azimi-Sadjadi, “Buried underwater object classification using a collaborative multi-aspect classifier,” vol. 34, no. 1. Proc of IEEE Journal of Oceanic Engr., 2009, pp. 32–44.
 - [12] N. Klausner, “Underwater target detection using multiple disparate sonar platforms,” Master’s thesis, Colorado State University, 2010.
 - [13] N. Klausner, M. Azimi-Sadjadi, and L. Scharf, “Detection of spatially correlated time series from a network of sensor arrays,” *To appear in IEEE Transactions on Signal Processing*, 2014.
 - [14] S. S. Chen, D. D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” in *SIAM Rev.*, vol. 43, no. 1, 2001, pp. 129–159.
 - [15] S. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” in *IEEE Trans. Signal Processing*, vol. 41, no. 12, 1993, pp. 3397–3415.
 - [16] G. Davis, S. Mallat, and Z. Zhang, “Adaptive time-frequency decompositions,” in *Opt. Eng.*, vol. 33, no. 7, 1994, pp. 2183–2191.
 - [17] K. Kreutz-Degado, J. F. Murray, B. D. Rao, K. Engan, T. Lee, and T. J. Sejnowski, “Dictionary learning algorithms for sparse representation,” in *Neural Computation*, vol. 15, no. 2, 2003, pp. 349–396.
 - [18] K. Kreutz-Delgado, J. Murray, B. Rao, K. Engan, T. Lee, and T. Sejnowski, “From sparse solutions of systems of equations to sparse modelling of signals and images,” in *Neural Computation*, vol. 15, no. 2, 2003, pp. 349–396.
 - [19] M. Aharon, M. Elad, and A. Bruckstein, “K-svd: An algorithm for designing overcomplete dictionaries for sparse representation,” in *IEEE Trans. Signal Processing*, vol. 54, no. 11, 2006, pp. 3431–4322.
 - [20] R. Duda, P. Hart, and D. Stork, “Pattern classification.” John Wiley, 2001.

- [21] S. T. Alexander, “Adaptive signal processing: Theory and applications.” Springer-Verlag, 1984.
- [22] S. Koschinski, “Underwater noise pollution from munitions clearance and disposal, possible effects on marine vertebrates, and its mitigation,” vol. 45, no. 6. Marine Technology Society Journal, 2011, pp. 80–88.
- [23] G. J. Dobeck, J. Hyland, and L. Smedley, “Automated detection/classification of sea mines in sonar imagery,” vol. 3079, pp. 90–110, 1997.
- [24] G. J. Dobeck, “Fusing sonar images for mine detection and classification,” *Proc SPIE*, vol. 3710, pp. 602–614, 1999.
- [25] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.
- [26] H. L. Van Trees, *Detection, Estimation, and Modulation Theory Part I*. John Wiley and Sons, 1968.
- [27] J. Cioffi and T. Kailath, “Fast, recursive least squares transversal filters for adaptive filtering,” in *IEEE Trans. on Acoustic, Speech, and Signal Processing*, vol. 32, 1984, pp. 304–337.
- [28] H. Nguyen, V. Patel, N. Nasrabadi, and R. Chellappa, “Kernel dictionary learning,” 2012.

LIST OF ABBREVIATIONS

- **AUV** - Autonomous Underwater Vehicle
- **BB** - Broadband
- **BOSS** - Buried Object Scanning Sonar
- **BP** - Basis Pursuit
- **CCA** - Canonical Correlation Analysis
- **GLRT** - Generalized Likelihood Ratio Test
- **HF** - High Frequency
- **LS** - Least Squares
- **MAP** - Maximum a Posteriori
- **MCA** - Multi-Channel Coherence Analysis
- **MMSE** - Minimum Mean Squared Error
- **MOD** - Method of Optimal Dictionary
- **MP** - Matching Pursuit
- **NSWC** - Naval Surface Warfare Center
- **ONR** - Office of Naval Research
- **OMP** - Orthogonal Matching Pursuit
- **ROC** - Receiver Operating Characteristic
- **ROI** - Region of Interest
- **SSAM** - Small Synthetic Aperture Minehunter
- **SVD** - Singular Value Decomposition