Undersea Target Classification Using Canonical Correlation Analysis

Ali Pezeshki,* *Member, IEEE,* Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE,* and Louis L. Scharf, *Life Fellow, IEEE*

Abstract

Canonical correlation analysis is employed as a multi-aspect feature extraction method for underwater target classification. The method exploits linear dependence or coherence between two consecutive sonar returns, at different aspect angles. This is accomplished by extracting the dominant canonical correlations between the two sonar returns and using them as features for classifying mine-like objects from non-mine-like objects. The experimental results on a wideband acoustic backscattered data set, which contains sonar returns from several mine-like and non-mine-like objects in two different environmental conditions, show the promise of canonical correlation features for mine-like versus non-mine-like discrimination. The results also reveal that in a fixed bottom condition, canonical correlation features are relatively invariant to changes in aspect angle.

Index Terms

Canonical correlations, linear dependence and coherence, multi-aspect feature extraction, underwater target classification.

This work was supported by the Office of Naval Research (ONR) under contracts N00014-02-1-0006 and N00014-04-1-0084. Preliminary versions of this work were presented at the MTS/IEEE Oceans Conference 2003, San Diego, CA, and the SPIE Defense and Security Symposium 2004, Orlando, FL.

A. Pezeshki (corresponding author) is with The Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA, Phone: (609) 258-6495, Fax: (609) 258-1735 (email: pezeshki@math.princeton.edu).

M. R. Azimi-Sadjadi is with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA, Phone: (970) 491-7956, Fax: (970) 491-2249, (e-mail: azimi@engr.colostate.edu).

L. L. Scharf is with the Department of Electrical and Computer Engineering, and the Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA, Phone: (970) 491-2979, Fax: (970) 491-2249 (e-mail: scharf@engr.colostate.edu).

I. INTRODUCTION

The problem of classifying underwater targets using active sonar has attracted a lot of attention in recent years, e.g. see [1]– [12]. This problem involves discrimination between targets and non-targets. Some of the factors that complicate this process include: non-repeatability and variation of target signatures with aspect angle, range, and grazing angle; diverse sizes, shapes, and scattering properties of the targets; presence of natural and man-made clutter; and a highly variable and reverberant operating environment. The problem is even more complicated when bottom targets are encountered, especially if they are buried or obscured by bottom features.

Due to the above factors, it is often difficult to detect and classify objects of interest based on the measurement from a single object-sensor orientation. There are often orientations at which different objects may look nearly identical. Consequently, in real-life situations, the decision about the presence and type of an object is generally made based upon observations of the received signals at several aspect angles.

In recent years, several multi-aspect-based methods for detection and classification of underwater targets from acoustic backscattered signals have been developed, e.g. see [1]– [12]. A detailed review of these methods is provided in [1]. However, in all these methods, multi-aspect classification is performed using various classification fusion methods, namely *decision-level* fusion [5], [8]- [12], *feature-level* fusion [1], [2], [4], [6], [7], or a combination of decision-level and feature-level fusion [3]. In decision-level fusion, an intermediate decision about the presence and type of the object (target or non-target) is made for every sonar return. The final decision is then made at the fusion center by optimally combining several single-aspect intermediate decisions. In feature-level fusion, feature vectors extracted from multiple sonar returns are simultaneously given to a decision making system to generate the final classification decision.

In [13]- [16] a different approach for multi-aspect detection and classification of underwater objects is reported. In this approach multi-aspect detection and classification is performed by exploring the correlation between two backscattered signals at different aspects, and analyzing the variations in the so-called *angular correlation function*. The angular correlation methods reported in [13]- [16] have been shown to provide better clutter suppression and finer resolution than conventional field imaging methods. The reader is referred to [13]- [16] and the references

therein for more details.

The approach taken in this paper is similar to that in [13]- [16] in the sense that we also explore the correlation or similarities between two-backscattered signals. However, we use a different method for the analysis of correlation. We exploit the linear dependence (or coherence) between two sonar returns, with certain aspect separation, to extract features that capture common target/non-target attributes in these two returns. The idea is that linear dependence between the sonar returns is an indication of the presence of a common signature, whereas linear independence indicates the absence of a common signature. This is the basic idea behind multi-channel tests for linear dependence [17] and the multiple coherence test of [18], [19].

The linear dependence between two data channels is measured by the canonical correlations [17], [20]– [22] of the channels. This implies that canonical correlations can be viewed as features that capture linear dependence or coherence between the two data channels, and hence may be used for detection and classification purposes. We intend to exploit this idea for classifying underwater mine-like objects (targets) from non-mine-like objects (non-targets). In this approach, the channels correspond to acoustic backscattered signals at two aspect angles with aspect/ping separation to be specified shortly.

Using canonical correlations, we exploit the linear dependence (coherence) between two backscattered signals or sonar returns to determine the presence of common signatures associated with targets or non-targets. Owing to the differences in shape, size and composition of the mine-like and non-mine-like objects, we hypothesize that the level of coherence between the sonar returns at different pings/aspects, for a mine-like object is different than that of a non-mine-like object. Therefore, the dominant canonical correlations, which capture most of the coherence between the two sonar returns, may be used as features to classify the objects at the corresponding aspect angles. We test this hypothesis on a subset of a wideband data set that was collected at the Applied Research Lab (ARL), University of Texas (UT)-Austin, and benchmark our results against those in [1] on the same data set.

II. A Review of Canonical Correlation Analysis

In this section, we review canonical coordinates and canonical correlations, and show that how linear dependence and coherence between two data channels may be determined using canonical correlation analysis. The material presented here, and much of the language and terminology,

are drawn from [20].

Consider the composite data vector $\mathbf{z}$ consisting of two random vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$, $m \leq n$, i.e.

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \in \mathbb{R}^{(m+n)}. \tag{1}$$

We assume that $\mathbf{x}$ and $\mathbf{y}$ have zero means and share the composite covariance matrix

$$\mathbf{R}_{zz} = E[\mathbf{z}\,\mathbf{z}^T] = E\left[\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \begin{pmatrix} \mathbf{x}^T & \mathbf{y}^T \end{pmatrix} \right] = \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix}. \tag{2}$$

This composite covariance matrix may be taken to block tridiagonal form as follows [20]:

$$\begin{bmatrix} \mathbf{F}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_{xx}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{yy}^{-1/2} \end{bmatrix} \mathbf{R}_{zz} \begin{bmatrix} \mathbf{R}_{xx}^{-T/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{yy}^{-T/2} \end{bmatrix} \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{\Sigma} \\ \mathbf{\Sigma}^T & \mathbf{I} \end{bmatrix}. \tag{3}$$

The trick is to choose $\mathbf{F}$, $\mathbf{\Sigma}$, and $\mathbf{G}$ to be the singular value decomposition (SVD) of the *coherence matrix* $\mathbf{C} = E[(\mathbf{R}_{xx}^{-1/2}\mathbf{x})(\mathbf{R}_{yy}^{-1/2}\mathbf{y})^T] = \mathbf{R}_{xx}^{-1/2}\mathbf{R}_{xy}\mathbf{R}_{yy}^{-T/2}$. That is,

$$\mathbf{C} = \mathbf{R}_{xx}^{-1/2}\mathbf{R}_{xy}\mathbf{R}_{yy}^{-T/2} = \mathbf{F}\mathbf{\Sigma}\mathbf{G}^T \tag{4}$$

where $\mathbf{F} \in \mathbb{R}^{m\times m}$ and $\mathbf{G} \in \mathbb{R}^{n\times n}$ are orthogonal matrices, i.e. $\mathbf{F}^T\mathbf{F} = \mathbf{F}\mathbf{F}^T = \mathbf{I}(m)$, $\mathbf{G}^T\mathbf{G} = \mathbf{G}\mathbf{G}^T = \mathbf{I}(n)$, and $\mathbf{\Sigma} = [\ \mathbf{\Sigma}(m) \quad \mathbf{0}\ ] \in \mathbb{R}^{m\times n}$ is a diagonal singular value matrix, with $\mathbf{\Sigma}(m) = \text{diag}[\sigma_1, \sigma_2, \ldots, \sigma_m]$ and $1 \geq \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m > 0$.

Then the transformation

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{F}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_{xx}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{yy}^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \tag{5}$$

resolves $\mathbf{z}^T = [\mathbf{x}^T, \mathbf{y}^T]$ into their canonical coordinates $\mathbf{w}^T = [\mathbf{u}^T, \mathbf{v}^T]$, with the composite covariance matrix

$$\mathbf{R}_{ww} = E[\mathbf{w}\mathbf{w}^T] = \begin{bmatrix} (\mathbf{R}_{uu} = \mathbf{I}) & (\mathbf{R}_{uv} = \mathbf{\Sigma}) \\ (\mathbf{R}_{vu} = \mathbf{\Sigma}^T) & (\mathbf{R}_{vv} = \mathbf{I}) \end{bmatrix}. \tag{6}$$

We refer to the elements of $\mathbf{u} = [u_i]_{i=1}^m \in \mathbb{R}^m$ and $\mathbf{v} = [v_i]_{i=1}^n \in \mathbb{R}^n$ as the *canonical coordinates* of $\mathbf{x}$ and $\mathbf{y}$, respectively. The diagonal cross-correlation matrix $\mathbf{\Sigma}$,

$$\mathbf{\Sigma} = E[\mathbf{u}\mathbf{v}^T] = E[(\mathbf{F}^T\mathbf{R}_{xx}^{-1/2}\mathbf{x})(\mathbf{G}^T\mathbf{R}_{yy}^{-1/2}\mathbf{y})^T] = \mathbf{F}^T\mathbf{C}\mathbf{G} \tag{7}$$

is called the *canonical correlation matrix* of *canonical correlations* σ_i, with $1 \geq \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m > 0$. An important property of canonical correlations is that they are invariant under uncoupled nonsingular transformations of $\mathbf{x}$ and $\mathbf{y}$ [20].

The linear dependence L between $\mathbf{x}$ and $\mathbf{y}$ may be measured as [20],

$$L = \det(\mathbf{I} - \mathbf{\Sigma}\mathbf{\Sigma}^T) = \prod_{i=1}^{m}(1 - \sigma_i^2); \quad 0 \leq L \leq 1, \tag{8}$$

The measure L takes the value 0 iff there is linear dependence between elements of $\mathbf{x}$ and $\mathbf{y}$; it takes the value 1 iff the elements of $\mathbf{x}$ and $\mathbf{y}$ are mutually uncorrelated. Equation (8) shows that the linear dependence or coherence between the two channels (e.g. sonar returns at two consecutive aspect angles) depends only on the canonical correlations of the channels and may easily be computed after the decomposition. They also show that each canonical coordinate pair, e.g. (u_i, v_i), contributes to the linear dependence according to its canonical correlation σ_i. This shows that the linear dependence between $\mathbf{x}$ and $\mathbf{y}$ is decomposed into the linear dependence between their canonical coordinates, each of which is determined by a canonical correlation. Thus, only canonical coordinates with large canonical correlations have significant contribution to the linear dependence and coherence. Therefore, they are the only ones that need to be retained for detection and classification.

Remark 1. A conventional method of canonical coordinate decomposition as in (5) does not offer a simple way to compute a small subset of canonical coordinates and correlations. A full SVD for the coherence matrix, along with square-root inverses of data covariances, has to be computed, regardless of the rank-reduction. This makes the conventional method computationally intractable, especially when the data channels have large dimensions. In [23], simple algorithms called *alternating power methods with deflation* have been reported to recursively compute the canonical coordinates and correlations one-by-one or in groups. Provided that the singular values of the coherence matrix are not close together, alternating power methods can be more efficient in computation than the conventional method. The reader is referred to [23] for details.

Remark 2. It is worth mentioning that canonical coordinates have been used in many other two-channel signal processing problems, including reduced-rank estimation [20], [24], low-rank Gauss-Gauss detection [25], and reduced-rank quantization of noisy sources [26]. In this paper, we explore the use of canonical correlation analysis for feature extraction specifically for underwater target classification.

III. WIDEBAND SONAR DATA SET

The sonar data set used in this paper is a subset of a wideband acoustic backscattered data set collected at the ARL-UT, Lake Travis test station (LTTS) [27]. This subset contains acoustic backscattered signals from three mine-like and three non-mine-like objects in two different bottom conditions, namely smooth and rough. For the rough bottom condition the sand is raked, giving it rippled effects. The mine-like objects are two cylindrical steel objects (Targets 6 and 7) of different sizes, and a truncated cone shape plastic object (Target 2). The non-mine-like objects are a water-filled steel drum, a concrete pipe, and a telephone pole.

During the data collection the objects were placed on a rotating seabed, 25-30 ft below the lake surface, with minimal embedding/scouring. The diameter of the seabed was 25 ft. The center of the object was positioned as near to the center of the circular platter (seabed) as possible. The objects were rotated in a horizontal plane while the acoustic panel was set at a depression/elevation (D/E) angle of 13.5 degrees, and range of 105 ft from the center of the seabed. A single rotation was executed while acoustic backscattered signals at nearly uniform 1 degree increments were collected.

The receiver array consisted of 16 channels with 1 inch separation. The transmit signal was a linear frequency modulated (LFM) waveform with a bandwidth of 85 kHz in the range of 15-100 kHz, and was approximately 7 msec long. The backscattered signals were recorded for approximately 16 msec at 500 kHz sampling rate, resulting in 8192 samples. In our experiments, we use the averaged data of four channels to yield a beamwidth that is just wide enough to insonify the object and not much of the surroundings. The beamwidth (in degrees) is computed approximately using $180/\pi\ell$, where ℓ is the aperture length in wavelengths. At 100 kHz (i.e. wavelength=0.6inch) the length of the four-channel aperture is 6.6 wavelengths, hence giving a beamwidth of approximately 8.6 degrees. At a distance of 105 ft, this results in a coverage width of 14 ft, which is wide enough to cover the entire length of the objects and narrow enough to avoid the side edges of the seabed.

An inverse matched filtering algorithm (see [5] for details) is used to remove the artifacts and secondary reflections caused by the straps and supporting barge of the experimented setup. The idea is to apply a window in matched filter domain to extract a clean signal and then perform the inverse matched filtering to obtain a "clean backscattered" signal. The clean backscattered signals (or from hereon backscattered signals) are then used in the feature extraction process

described next.

IV. Feature Extraction Process

In this section, we describe how canonical correlation analysis may be used to extract a set of features that capture common target/non-target attributes among two consecutive sonar returns, with certain aspect separation. Later in Section V, we will apply the feature extraction method presented here to the wideband ARL-UT data set, and use the extracted features for classifying mine-like objects from non-mine-like objects.

To build the ensembles of the two channels ($\mathbf{x}$ and $\mathbf{y}$) for canonical correlation analysis, we partition two backscattered signals from an object, with certain aspect separation, into overlapping blocks, as illustrated in Figure 1. In this figure, sonar return 1 is the backscattered signal from an object at aspect angle, say β_1, and sonar return 2 is the backscattered signal from the same object at aspect angle $\beta_2 = \beta_1 + \Delta\beta$, where $\Delta\beta$ is the aspect separation between these two returns. The blocks of sonar return 1 are taken as the samples of the first channel (the $\mathbf{x}$-channel) and the blocks of sonar return 2 are taken as the samples of the second channel (the $\mathbf{y}$-channel). Referring to Figure 1, the data sample $\mathbf{x}_i$ is the vector of the time series associated with the ith block of range cells in sonar return 1, and $\mathbf{y}_i$ is the vector of the time series associated with the ith block of range cells in sonar return 2. The collections of these data vector samples form the sample data matrices $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_M]$ and $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_M]$. Thus, in the canonical correlation analysis formulation of Section II, the covariance matrix $\mathbf{R}_{xx}$, $\mathbf{R}_{yy}$ and $\mathbf{R}_{xy}$ are replaced by their corresponding sample estimates $\mathbf{S}_{xx} = (1/M)\mathbf{X}\mathbf{X}^T$, $\mathbf{S}_{yy} = (1/M)\mathbf{Y}\mathbf{Y}^T$, and $\mathbf{S}_{xy} = (1/M)\mathbf{X}\mathbf{Y}^T$. The dominant canonical correlations between these sample data matrices $\mathbf{X}$ and $\mathbf{Y}$, which capture most of the coherence, will be used as features to represent the sonar signal at aspect angle β_1.

Remark 3. Taking blocks of the backscattered signals as snapshot vectors to build sample covariance matrices requires that the second-order statistics of the random vector process, which generates the snapshots, is time invariant. As we argue in the Appendix, this occurs either when a short pulse is used to scan a uniform target or when a long pulse is used to scan the entire target.

The aspect separation $\Delta\beta$ should be large enough so that the bottom reverberation effects are almost uncorrelated, but small enough so that the returns from the objects remain coherent.

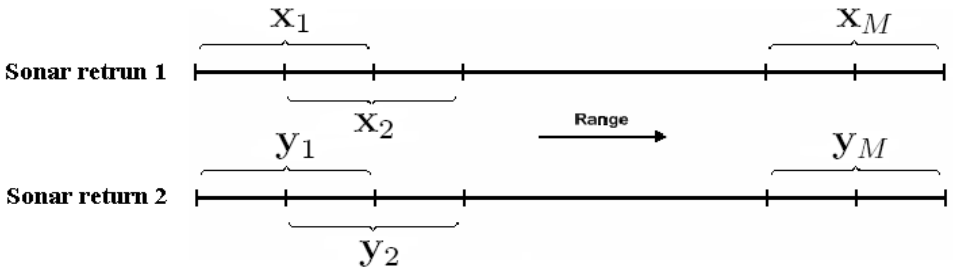

Fig. 1. Building the ensembles of the two channels ($\mathbf{x}$ and $\mathbf{y}$) for canonical correlation analysis from two sonar returns.

Clearly, this choice depends on many factors including specific properties of the sonar such as speed of movement, range and elevation of the vehicle, and size and orientation of the object. For the ARL-UT data set, we have experimentally determined that an aspect separation of $\Delta\beta = 16$ degrees is a reasonable choice. Therefore, in the simulations preformed in Section V, the sonar returns are paired according to aspect angles as follows: $\{1, 17\},\ \{2, 18\} \cdots, \{344, 360\}, \{345, 1\}, \cdots,$ $\{360, 16\}$. The dominant canonical correlations extracted from each pair of aspects $\{\beta, \beta + 16\}$ will be used as features to represent the return at aspect angle β.

V. Canonical Correlation Features and Classification Results

In this section, the feature extraction process described in Section IV is applied to the sonar returns in the ARL-UT data set, and the extracted features are used to classify mine-like objects from non-mine-like objects. In the experiments performed here, the backscattered signals are partitioned into blocks of size 50 samples, with 50% (25 samples) overlap, leading to 50-dimensional $\mathbf{x}$ and $\mathbf{y}$ channels. Although, the block size is determined experimentally here, there are several factors that must be considered. These include duration of the transmit signal, sampling rate, and the shape and size of the objects. Since the length of the backscattered signals is 8192 samples, dividing each backscattered signal into blocks of size 50 samples with 25 samples overlap results in $M = 327$ vector samples for $\mathbf{x}-$ and $\mathbf{y}-$ channels. As a result the sample data matrices $\mathbf{X}$ and $\mathbf{Y}$ will each have size 50×327. The 50×327 data matrices $\mathbf{X}$ and $\mathbf{Y}$ share 50 canonical correlations. We use the first (largest) 15 out 50 canonical correlations between data matrices $\mathbf{X}$ and $\mathbf{Y}$, which are associated with a pair of angles $(\beta, \beta + 16)$, as features to represent the aspect angle β. The first 15 canonical correlations capture a major portion of the coherence between the sonar returns.

Experiment 1: The objective here is to demonstrate the usefulness of canonical correlation features for classifying mine-like from non-mine-like objects in both smooth and rough bottom conditions. The training data set for classification is formed from the feature vectors extracted at 90 aspect angles of the *smooth bottom* data, at aspect increments of 4°. The feature vectors extracted from the rest of the aspect angles of the smooth bottom data (270 aspect angles) are kept to validate the trained classifier. This validation data set is primarily used to select the best trained classifier. To see how well the trained classifier generalizes, the feature vectors extracted from the backscattered signals in the *rough bottom* condition are used as a testing data set.

Figures 2(a)-(c) show the scatter plots of the first two canonical correlation features for the training, validation, and testing data sets, respectively. As can be seen, for the training data set (Fig. 2(a) and the validation data set (Fig. 2(b)), the features of mine-like objects (Targets 2, 6, and 7) are packed together and almost completely separated from those of the non-mine-like objects (steel drum, concrete pipe, and telephone pole). Additionally, the extracted features for the training and validation data sets, for the objects in the smooth bottom condition, are consistent (occupy the same regions in the scatter plots).

In the rough bottom test condition (Fig. 2(c)), features of Target 2, Target 6, and the telephone pole stay in the same regions as in the smooth bottom condition, while those of the drum and concrete pipe become more compact and move slightly towards the right side of the plot. Nonetheless, they still occupy almost the same regions as in training and validation data sets. Features of Target 7, however, move from the upper right corner and spread out to the left side and mix with those of the steel drum and concrete pipe. The reason for changes in features of Target 7 may be attributed to the secondary reflections and the rather large size of this cylindrical target, which is comparable to that of the drum. Clearly, the movement of Target 7 features leads to some degradation in classification performance in the rough bottom condition. These scatter plots show that for five out of six objects, canonical correlations are fairly robust (only slightly change) to the changes in the bottom condition. The same observation is also valid for the third and fourth canonical correlation features.

Subsequently, the extracted canonical correlation features are used to train a back-propagation neural network (BPNN) [28] to classify the mine-like objects from non-mine-like objects. The goal is to determine whether an object is mine-like or non-mine-like, rather than classifying all the objects. In other words, the problem is a two-class classification problem. To find a

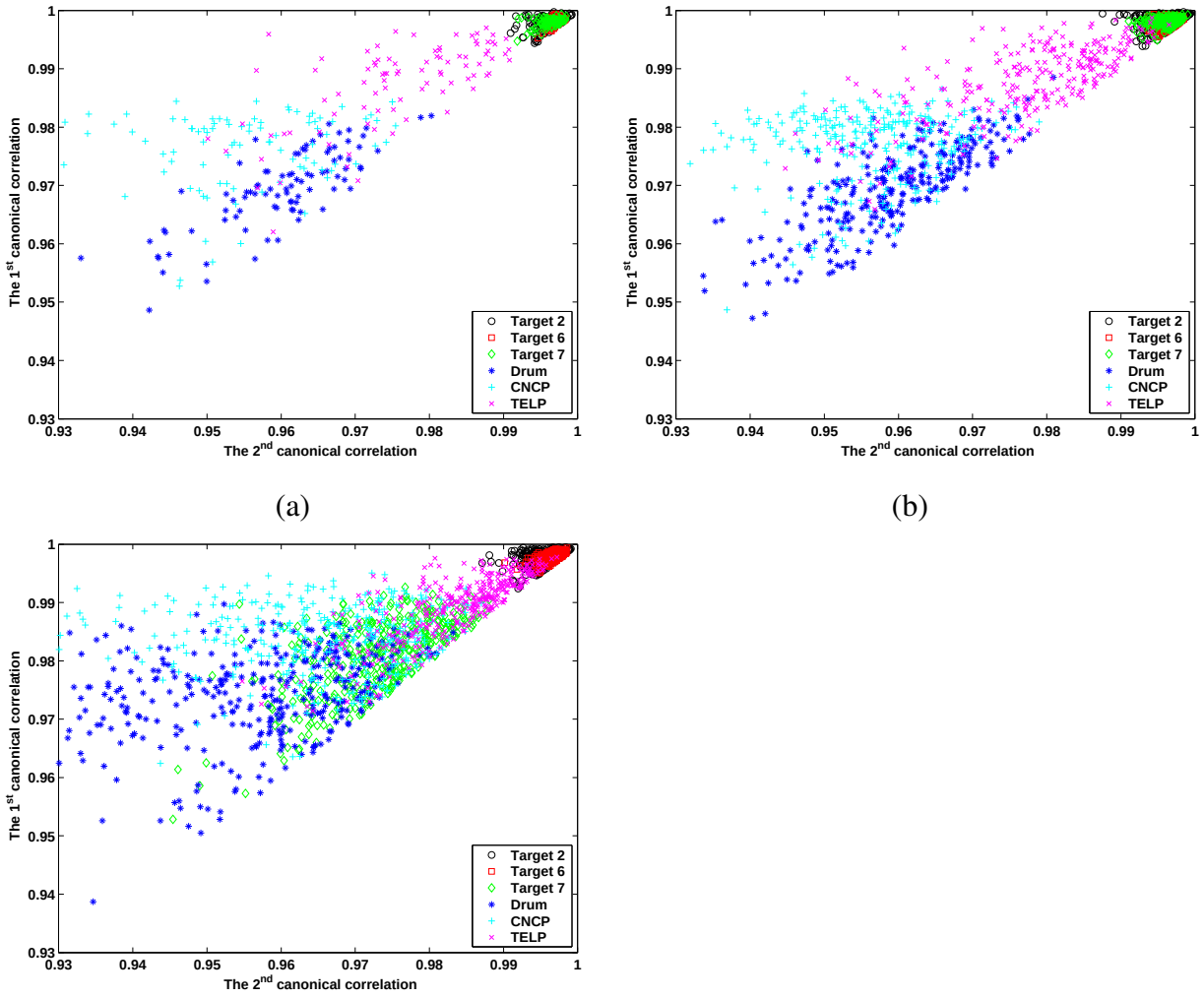

Fig. 2. Scatter plots of the first two canonical correlation features for (a) training, (b) validation, and (c) testing data sets. The scatter plots show that, for five out of six objects canonical correlations are fairly robust (only slightly change) to the changes in the bottom condition.

good network structure, eight different two-layer BPNN structures were tried. The number of hidden layer neurons was varied from 26 to 46. Each network was trained for ten different weight initializations. The training was performed for 10000 epochs, where an epoch was a complete sweep over the entire training data set. The best BPNN classifier, which was selected based on the average classification rates on training and validation data sets, gave an overall correct classification rate of 99.1% on the training data set, 98.6% on the validation data set,

TABLE I

CONFUSION MATRICES OF THE BPNN CLASSIFIER TRAINED WITH CANONICAL CORRELATION FEATURES.

	Validation Data Set (Smooth Bottom Data)		Testing Data Set (Rough Bottom Data)	
Object	Target	Non-Target	Target	Non-Target
Target 6	270	0	360	0
Target 7	270	0	2	358
Target 2	270	0	360	0
Drum	0	270	0	360
Concrete Pipe	0	270	0	360
Telephone Pole	22	248	52	308

and 81.0% on the testing data set. These percentages are obtained based on a hard-limiting decision threshold. We compare these results with those in [1] obtained by using linear predictive coding (LPC) subband features and decision-level fusion method. As reported in [1], the correct classification rates on the training, validation, and testing data sets are 99.6%, 82.5% and 75.2%, respectively. Comparing these results indicate almost 16% and 6% improvements on the validation and testing data sets when canonical correlation features are used. We note that these improvements are achieved without requiring any multi-aspect classification fusion, like the one used in [1]. The confusion matrices obtained for the classifiers, trained using these two feature types, are shown in Table I (for the canonical correlation features) and Table II (for the LPC subband features). These results demonstrate the promise of the canonical correlation features for classifying targets from non-targets in different bottom conditions.

It is interesting to note that the classifier trained using canonical correlation features has correctly classified Targets 2 and 6 at all aspect angles in both smooth and rough bottom conditions, while the classifier trained using the LPC subband features has poor performance on these targets. Additionally the canonical correlation features provide substantially lower false alarm rates (2.7% for validation and 4.8% for testing) compared to the LPC subband features (16.1% for validation and 22.4% for testing). However, the classifier trained with the LPC subband features provides better performance for Target 7 in the rough bottom condition compared to the canonical correlation-based classifier.

TABLE II

CONFUSION MATRICES OF THE BPNN CLASSIFIER TRAINED WITH LPC SUBBAND FEATURES. THE CLASSIFICATION PROBLEM IS A TWO-CLASS (MINE-LIKE OBJECT VERSUS NON-MINE-LIKE OBJECT) PROBLEM.

	Validation Data Set (Smooth Bottom Data)		Testing Data Set (Rough Bottom Data)	
Object	Target	Non-Target	Target	Non-Target
Target 6	215	55	224	136
Target 7	177	93	214	146
Target 2	265	5	349	11
Drum	57	213	128	232
Concrete Pipe	56	214	92	268
Telephone Pole	17	253	22	338

Experiment 2: Our goal in this experiment is to investigate the robustness of the canonical correlation features with respect to aspect angle variation in a fixed bottom condition, namely the smooth bottom. The training data set for each object is formed from the feature vectors extracted for $1/4$ of the aspect angles that correspond to one side of the objects (aspect angles 0 to 179 degrees) only. The feature vectors extracted from the rest of the aspect angles between 0 to 179 degrees in the smooth bottom condition are kept to validate the trained classifier. The generalization and robustness of the trained classifier is tested, in the same bottom condition, using the features extracted from sonar returns from the other side of the objects, i.e. aspect angles 180 to 359 degrees. Clearly, in this experiment the classifier is not exposed to the information on the other side of the objects during the training and validation processes.

Figures 3(a)-(c) show the scatter plots of the first two canonical correlation features for the training, validation, and testing data sets. As can be seen, the canonical correlation features for targets are similar and almost completely separated from those of the non-targets. Additionally, the extracted features for the training, validation, and testing data sets for the objects are fairly consistent, implying that the canonical correlation features are relatively invariant to aspect angle variation.

For this experiment, the best two-layer BPNN classifier, trained with the extracted canonical correlation features, gave an overall correct classification rate of 99.6% on the training data set,

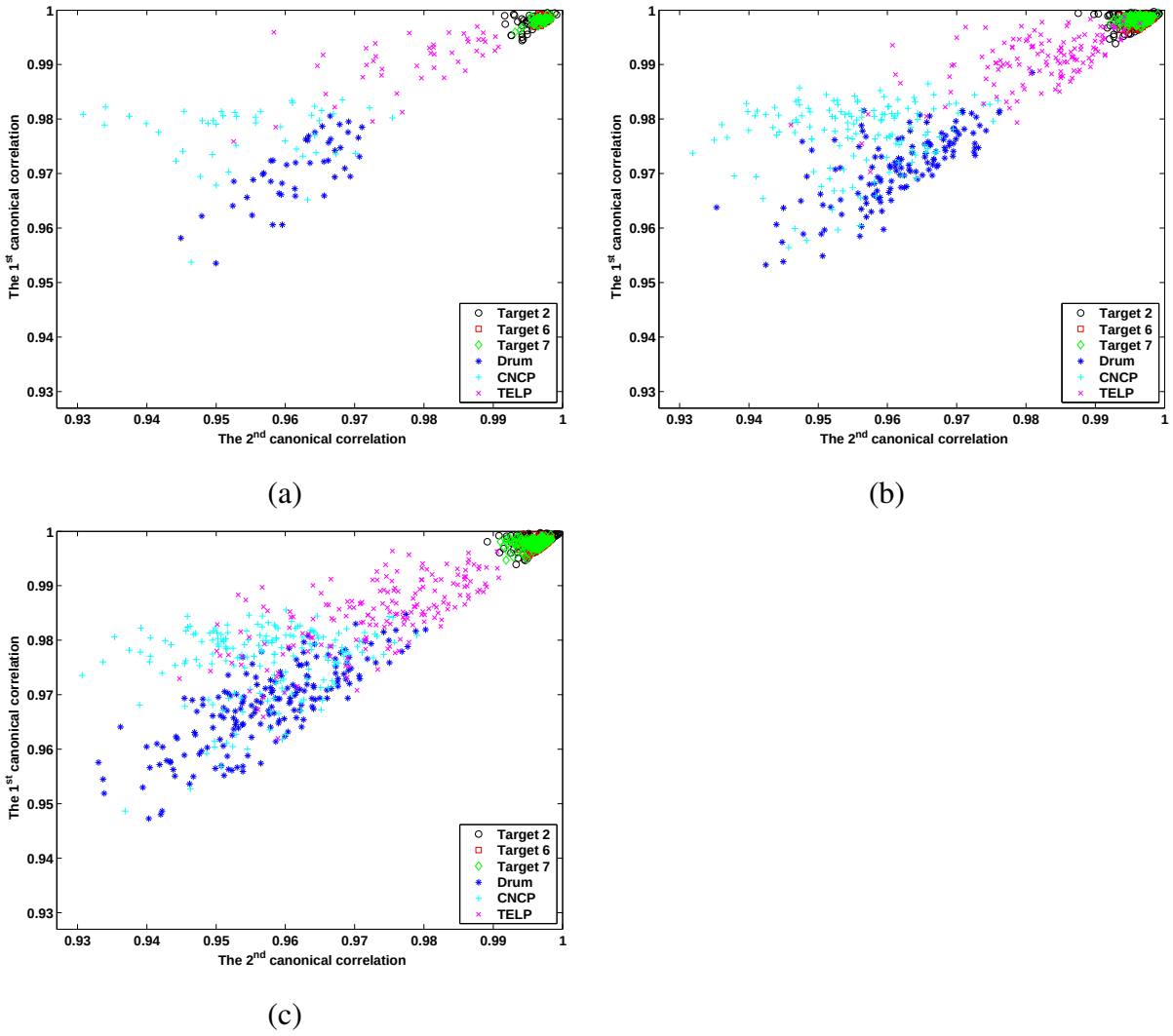

Fig. 3. Scatter plots of the first two canonical correlations for (a) training, (b) validation, and (c) testing data sets. The plots show that canonical correlation features are relatively invariant to aspect angle variation.

98.1% on the validation data set, and 99.8% on the testing data set. The confusion matrices of this classifier for validation and testing data sets are shown in Table III. It is seen that only at a few aspect angles in the validation data set the telephone pole is misclassified as a mine-like-object. Similarly, in the testing data set, there are only four misclassifications. These results demonstrate that canonical correlation features of an object are relatively invariant to changes in aspect angle, provided that the environmental condition remains unchanged.

TABLE III

CONFUSION MATRICES OF THE BPNN CLASSIFIER TRAINED WITH THE CANONICAL CORRELATION FEATURES THAT ARE EXTRACTED FROM ONE SIDE OF THE OBJECTS. THE CLASSIFICATION PROBLEM IS A TWO-CLASS (MINE-LIKE OBJECT VERSUS NON-MINE-LIKE OBJECT) PROBLEM.

	Validation Data Set (Smooth Bottom Data)		Testing Data Set (Smooth Bottom Data)	
Object	Target	Non-Target	Target	Non-Target
Target 6	135	0	180	0
Target 7	135	0	179	1
Target 2	135	0	178	2
Drum	0	135	0	180
Concrete Pipe	0	135	0	180
Telephone Pole	15	120	1	179

VI. CONCLUSIONS

In this paper, canonical correlation analysis was exploited to develop a multi-aspect feature extraction method for underwater target classification from a wideband sonar data set. The basic idea is that when an object (target or non-target) is present in the environment, consecutive sonar returns exhibit a change in the coherence compared to the case when there is no object. Further, we hypothesized that the amount of coherence captured by the dominant canonical correlations of the two sonar returns from a mine-like object is different than that of a non-mine-like object. Our experiments on the wideband ARL-UT data set demonstrated that canonical correlation features offer good separation between mine-like and non-mine-like objects. The results show that except for one of the objects, namely Target 7 which is of comparable size to the drum, the canonical correlation features are robust to changes in the bottom condition. Moreover, we showed that in a fixed bottom condition, canonical correlation features are relatively invariant to changes in aspect angle. Recent results [29] on Buried Object Scanning Sonar (BOSS) data collected in St. Andrews Bay, Panama City, FL, further validate the potential of the proposed canonical correlation-based feature extraction method for detection and classification of buried underwater objects.

APPENDIX

The backscattered signal at aspect angle θ, denoted by $r(t;\theta)$, may be written as a linear convolution of the form

$$r(t;\theta) = \int_{-\infty}^{\infty} h(\tau;\theta)s(t-\tau)d\tau, \tag{A.1}$$

where $s(t)$ is the transmit pulse and $h(t;\theta)$ is the target random impulse response at time t and aspect angle θ.

The instantaneous cross-correlation function between complex amplitudes $r(t_1;\theta_1)$ and $r(t_2;\theta_2)$, corresponding to two sonar returns at angles θ_1 and θ_2, is given by

$$r(t_1;\theta_1)r^*(t_2;\theta_2) = \int\int s(t_1-\tau)h(\tau;\theta_1)h^*(\tau';\theta_2)s^*(t_2-\tau')d\tau d\tau', \tag{A.2}$$

where * denote complex conjugate.

Assume that the random impulse response $h(t;\theta)$ is wide sense stationary and uncorrelated over time. Then, the cross-correlation $R_{rr}(t_1,t_2;\theta_1,\theta_2)$ between $r(t_1;\theta_1)$ and $r(t_2;\theta_2)$ is

$$\begin{aligned} R_{rr}(t_1,t_2;\theta_1,\theta_2) &= E\left[r(t_1;\theta_1)r^*(t_2;\theta_2)\right] \\ &= \int\int R_{hh}(\tau;\theta_1,\theta_2)\delta(\tau-\tau')s(t_1-\tau)s^*(t_2-\tau')d\tau d\tau' \\ &= \int R_{hh}(\tau;\theta_1,\theta_2)s(t_1-\tau)s^*(t_2-\tau)d\tau, \end{aligned} \tag{A.3}$$

where $R_{hh}(\tau;\theta_1,\theta_2)$ is the target scattering function. In the feature extraction procedure in Section IV, t_1 and t_2 correspond to two different time samples (range cells) within a block, and θ_1 and θ_2 correspond to aspect angles β and $\beta+\Delta\beta$.

The transmit pulse $s(t)$ in this paper is an LFM signal. The baseband representation of an LFM signal with time duration T and bandwidth B is

$$s(t) = \begin{cases} \frac{1}{\sqrt{T}}e^{j\frac{\pi B}{T}t^2}, & |t| \le \frac{T}{2} \\ 0, & |t| > \frac{T}{2} \end{cases} \tag{A.4}$$

Inserting (A.4) in (A.3) yields

$$R_{rr}(t_1,t_2;\theta_1,\theta_2) = \frac{1}{T}\int_{\max[t_1,t_2]-T/2}^{\min[t_1,t_2]+T/2} R_{hh}(\tau;\theta_1,\theta_2)e^{j\frac{\pi B}{T}(t_1^2-t_2^2-2\tau(t_1-t_2))}d\tau. \tag{A.5}$$

Let $t_1 = t + \Delta t$ and $t_2 = t$. Then, we can write (A.5) as

$$R_{rr}(t+\Delta t, t; \theta_1, \theta_2) = \frac{1}{T} \int_{\max[t_1,t_2]-T/2}^{\min[t_1,t_2]+T/2} R_{hh}(\tau; \theta_1, \theta_2) e^{j\frac{\pi B}{T}(2t\Delta t + \Delta t^2 - 2\tau\Delta t)} d\tau$$

$$= \frac{1}{T} e^{j\frac{\pi B}{T}(2t\Delta t + \Delta t^2)} \int_{\max[t_1,t_2]-T/2}^{\min[t_1,t_2]+T/2} R_{hh}(\tau; \theta_1, \theta_2) e^{-j\frac{\pi B}{T}2\tau\Delta t} d\tau. \quad \text{(A.6)}$$

Note that the integral in (A.6) is equal to zero for $\Delta t > T$.

If the pulse length T is short compared to the variation length of $\mathbf{R}_{hh}$ is then we can write

$$R_{rr}(t+\Delta t, t; \theta_1, \theta_2) \approx R_{rr}(\Delta t; \theta_1, \theta_2)$$

$$= \frac{1}{T} e^{j\frac{\pi B}{T}(2t\Delta t + \Delta t^2)} R_{hh}(\tau; \theta_1, \theta_2) \int_{\max[t_1,t_2]-T/2}^{\min[t_1,t_2]+T/2} e^{-j\frac{\pi B}{T}2\tau\Delta t} d\tau$$

$$= \left(\frac{T-\Delta t}{T}\right) sinc\left[\frac{\pi B}{T}\Delta t(T-\Delta t)\right] R_{hh}(t; \theta_1, \theta_2). \quad \text{(A.7)}$$

If R_{hh} is time-invariant then

$$R_{yy}(\Delta t; \theta_1, \theta_2) = \frac{T-\Delta t}{T} sin\left[\frac{\pi B}{T}\Delta t(T-\Delta t)\right] R_{hh}(\theta_1, \theta_2), \quad \text{(A.8)}$$

and the correlation function $R_{rr}(\Delta t; \theta_1, \theta_2)$ becomes time-invariant. Alternatively, if the pulse is much longer compared to the target length so that the pulse scans the entire target, then R_{hh} can be assumed to be time-invariant during the illumination period.

ACKNOWLEDGEMENTS

Special thanks are due to the Office of Naval Research (ONR) for funding this project. The data used in this work were supplied by The Applied Research Laboratories at the University of Texas at Austin under the auspices of the Office of Naval Research, Code 321 Undersea Signal Processing. We would like to thank Drs. Pat Pitt and Ken Scarbrough from the ARL-UT for providing this data set and their technical assistance. We would also like to thank Dr. Gerald J. Dobeck at the NSWC-Coastal Systems Station for many useful discussions and comments on the results of this paper.

REFERENCES

[1] M. Robinson, M. R. Azimi-Sadjadi, and J. Salazar, "Multi-aspect target discrimination using hidden Markov models and neural networks," *IEEE Trans. Neural Networks*, vol. 16, pp. 447–459, March 2005.

[2] M. Robinson, M. R. Azimi-Sadjadi, D. D. Sternlicht, and D. Lemonds, "Multi-aspect acoustic classification of buried objects," in *Proc. MTS/IEEE Oceans'03*, (San Diego, CA), pp. 478–484, Sept. 22-26 2003.

[3] M. R. Azimi-Sadjadi, D. Yao, A. A. Jamshidi, and G. Dobeck, "Underwater target classification in changing environments using adaptive feature mapping," *IEEE Trans. Neural Networks*, vol. 13, pp. 1099–1111, May 2002.

[4] N. Dasgupta, P. Runkle, L. Couchman, and L. Carin, "Dual hidden Markov model for characterizing wavelet coefficients from multi-aspect scattering data," *Signal Processing*, vol. 81, pp. 1303–1316, June 2001.

[5] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, and G. J. Dobeck, "Underwater target classification using wavelet packets and neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 784–794, May 2000.

[6] P. Runkle, L. Carin, L. Couchman, T. Yoder, and J. Bucaro, "Multi-aspect identification of submerged elastic targets via wave-based matching pursuits and hidden Markov models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, pp. 1371–1378, Dec. 1999.

[7] P. Runkle, P. Bharadwaj, and L. Carin, "Hidden Markov model multi-aspect target classification," *IEEE Trans. Signal Processing*, vol. 47, pp. 2035–2040, July 1999.

[8] G. Okimoto and D. Lemonds, "Principal component analysis in the wavelet domain: new features for underwater object recognition," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, vol. 3710, pp. 697–708, April 1999.

[9] C. F. Barnes, "Acoustic backscatter classification for mine detection using multiple fused aspects and novel database classification rules," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, vol. 3392, pp. 357–368, April 1998.

[10] D. Casasent and N. Kuljanyavivat, "Mine detection from multiple acoustic backscatter data," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, vol. 3392, pp. 370–381, April 1998.

[11] A. Pezeshki, M. R. Azimi-Sadjadi, and L. L. Scharf, "Classification of underwater mine-like and non-mine-like objects using canonical correlations," in *Detection and Remediation Technologies for Mines and Minelike Targets IX*, vol. 5415, pp. 336–341, April 2004.

[12] A. Pezeshki, M. R. Azimi-Sadjadi, L. L. Scharf, and M. Robinson, "Underwater target classification using canonical correlations," in *Proc. MTS/IEEE Oceans'03*, (San Diego, CA), pp. 1906–1911, Sept. 22-26 2003.

[13] G. Zhang, L. Tsang, and Y. Kuga, "The angular correlation function of wave scattering by a buried object embedded in random discrete scatterers under a random rough surface," *Microwave and Optical Technology Letters*, vol. 14, pp. 144–151, Feb. 1997.

[14] G. Zhang, L. Tsang, and Y. Kuga, "Numerical studies of the detection of targets embedded in clutter by using angular correlation functions and angular correlation imaging," *Microwave and Optical Technology Letters*, vol. 17, pp. 82–86, Feb. 1998.

[15] G. Zhang and L. Tsang, "Application of angular correlation function of clutter scattering and correlation imaging in target detection," *IEEE Trans. Geoscience and Remote Sensing*, vol. 36, pp. 1485–1493, Sept. 1998.

[16] G. Zhang, L. Tsang, and K. Pak, "Angular correlation function and scattering coefficients of electromagnetic waves scattered by a buried object under a two-dimensional rough surface," *Journal of Optical Society of America*, vol. 15, pp. 2995–3002, Dec. 1998.

[17] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*, ch. 8-10. Academic Press, 1979.

[18] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple channel signal detection," *IEEE Trans. Signal Processing*, vol. 43, pp. 2049–2057, Sept. 1995.

[19] H. Gish and D. Cochran, "Generalized coherence," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, pp. 2745–2748, April 1987.

[20] L. L. Scharf and C. T. Mullis, "Canonical coordinates and the geometry of inference, rate and capacity," *IEEE Trans. Signal Processing*, vol. 48, pp. 824–831, Mar. 2000.

[21] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[22] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1958.

[23] A. Pezeshki, L. L. Scharf, M. R. Azimi-Sadjadi, and Y. Hua, "Two-channel constrained least squares problems: Solutions using power methods and connections with canonical coordinates," *IEEE Trans. Signal Processing*, vol. 53, pp. 121–135, Jan. 2005.

[24] L. L. Scharf and J. T. Thomas, "Wiener filters in canonical coordinates for transform coding, filtering, and quantizing," *IEEE Trans. Signal Processing*, vol. 46, pp. 647–654, Mar. 1998.

[25] A. Pezeshki, L. L. Scharf, J. K. Thomas, and B. D. V. Veen, "canonical coordinates are the right coordinates for low-rank Gauss-Gauss detection and estimation," *IEEE Trans. Signal Processing*, to appear 2006.

[26] P. J. Schreier and L. L. Scharf, "Canonical coordinates for transform coding of noisy sources," *IEEE Trans. Signal Processing*, to appear 2006.

[27] ARL-UT, Austin, TX, "Target data acquisition setup." Manual.

[28] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ: Printice Hall, 1996.

[29] M. Yamada, J. Cartmill, and M. R. Azimi-Sadjadi, "Burried underwater target classification using the new BOSS and canonical correlation decomposition feature extraction," in *Proc. MTS/IEEE Oceans'05*, (Washington DC), Sept. 2005.