A Multichannel Canonical Correlation Analysis Feature Extraction with Application to Buried Underwater Target Classification

Bryan Thompson, Jered Cartmill, Mahmood R. Azimi-Sadjadi, and Steven G. Schock

Abstract—Multichannel Canonical Correlation Analysis (MCCA) is used in this paper for feature extraction from multiple sonar returns off of buried underwater objects using data collected by the new generation Buried Object Scanning Sonar (BOSS) system. Comparisons are made between the classification results of features extracted by the proposed algorithm and those extracted by the two-channel Canonical Correlation Analysis (CCA) algorithm. This study compares different feature extraction and classification algorithms, and the results are presented in terms of confusion matrices. The results show that MCCA yields higher correct classification rates than CCA while reducing the classifier’s structural complexity.

I. INTRODUCTION

Detection and classification of buried underwater objects is a challenging problem due to various factors including: variability of the target signatures and features with respect to aspect angle, presence of natural and man-made clutter, obscured objects and bottom reverberation effects. In 2002, a new generation Buried Object Scanning Sonar (BOSS) system, designed to scan for objects buried in the sea floor, was developed [1]. Using this system, a data set was collected at St. Andrew’s Bay, Panama City, Florida in May 2004 that contains various mine-like and non-mine-like objects buried 5-20 inches below the seabed. This real data set presents a formidable challenge to any target detection and classification system due to: (a) the presence of a wide variety of buried objects and natural clutter, (b) the small number of pings collected for each of the buried objects, and (c) variations in the operating conditions. Thus, it is desirable to devise feature extraction and target detection/classification methodologies that remain robust to such changes in the data and can screen the entire data set for potential mine-like objects.

In the past, many different methods [2], [3] have been developed and applied to the task of correctly detecting and differentiating between mine-like and non-mine-like objects. From these studies, it has been determined that one way to effectively deal with this problem is to utilize multiple looks at an object to make a more informed classification decision. This is because sonar returns for two different objects at certain orientations could potentially be so similar that they may easily be confused. Consequently, a more reliable decision about the presence and type of an object can be made based upon the observations of multiple sonar ping returns. This allows for more information to accumulate about the size, shape, composition, and orientation of the objects, which in turn yields more accurate discrimination. Moreover, when the feature space undergoes changes due to different operating and environmental conditions, multiple ping decision-making is almost a necessity in order to maintain performance [4]. Generally, sequences of three or four pings are used since in actual mine-hunting scenarios an object is declared as a potential mine if a strong indication exists in these sequences of sonar pings.

There are two general approaches to this problem: a multi-ping feature extraction and/or a multi-ping classification fusion. In the first approach, common features among multiple sonar pings are identified and extracted for better object discrimination [3]. The motivation behind this approach is the idea that mine-like objects exhibit different patterns of coherence when compared to non-mine-like objects. The second approach finds its basis in multi-look classification fusion systems. Possible fusion algorithms include feature-level fusion [4], decision-level fusion [2], and a combination of both feature and decision-level fusion. In this paper, a new approach for multi-ping feature extraction is proposed which makes use of the multichannel extension of CCA [5]. A comparison of this method and the decision-level fusion method is also made to evaluate the performance for mine-like versus non-mine-like object discrimination.

The organization of this paper is as follows. Section II reviews the MCCA method and its application as a multi-ping feature extractor on the BOSS data set. Section III describes the BOSS system as well as the data collection process. The procedures used to implement CCA and MCCA-based classification systems as well as a decision-level fusion are discussed in Section IV. Section V compares the results of applying the classifiers for detection and classification purposes on two different runs through the target field. Finally, conclusions and future work are discussed in Section VI.

II. MULTICHANNEL CANONICAL CORRELATION ANALYSIS

A method for exploiting the linear dependence (coherence) between multiple (more than two) sonar returns is presented and implemented in order to generate features used for classification. It is well-known that canonical correlation analysis (CCA) provides an elegant framework [6] for analyzing linear dependence and mutual information between two data channels. The results of [7] showed that MCCA provides a
useful measure used to interpret and summarize the evolution of psoriasis in medical patients. In this study, it is desired to test the hypothesis that mine-like objects yield a similar coherence pattern among multiple sonar returns. This pattern of coherence is different than that found among sonar returns for non-mine-like objects. This has been demonstrated experimentally using two sonar returns [3], [8]. The multichannel extension of CCA [5] is used here for extracting features from the BOSS data set.

The motivation behind extending CCA to the multichannel case is to determine the linear dependence among multiple (greater than two) sets of data, e.g., sonar pings. Unlike CCA, MCCA transforms multivariate multichannel data into new orthogonal coordinates referred to in this paper as “surrogate canonical coordinates” which are used to exploit the differences and similarities in the multichannel data sets. The reason they are called surrogate canonical coordinates will be clarified later. Like CCA, the computed canonical correlations are invariant to non-singular linear transformations [5]. In what follows we adopt the iterative MCCA approach which extracts the surrogate coordinates and canonical correlations one at a time [5]. Different cost functions have been proposed for finding the coordinates/correlations which lead to various generalized eigenvalue problems as will be shown in this section.

Consider a set of \( n \) random vectors \( x_1, x_2, \cdots, x_n \) which comprise the composite random vector \( x = [x_1^T, x_2^T, \cdots, x_n^T]^T \). Each vector \( x_j \in \mathbb{R}^{d_j} \) has a size \( d_j \), where it is assumed that \( d_1 \leq d_2 \leq \cdots \leq d_n \). Moreover, it is assumed that the channels have zero mean, i.e., \( E[x_j] = 0, \forall j \in [1, n] \). The correlation matrix of the defined composite vector \( x \) is

\[
R_{xx} = E[x x^T] = \begin{bmatrix}
R_{11} & R_{12} & \cdots & R_{1n} \\
R_{21} & R_{22} & \cdots & R_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
R_{n1} & R_{n2} & \cdots & R_{nn}
\end{bmatrix}
\]

(1)

where \( R_{jk} = E[x_j x_k^T] \) and \( R_{jk} = R_{kj}^T \).

Similar to CCA, the goal is to search for linear mapping vectors, \( a_{i,j} \), of the data channels \( x_j \), to produce coordinates

\[
v_{ij} = a_{i,j}^T x_j, \quad j \in [1, n]
\]

(2)

that satisfy certain requirements discussed later. The \( i^{th} \) multichannel surrogate canonical coordinate, \( i \in [1, d_i] \), for the \( j^{th} \) channel are given by (2). If this process is repeated for all the channels \( j \in [1, n] \), the following vector is obtained

\[
v_i = \begin{bmatrix}
a_{i,1}^T x_1 \\
a_{i,2}^T x_2 \\
\vdots \\
a_{i,n}^T x_n
\end{bmatrix} = \begin{bmatrix}
v_{i,1} \\
v_{i,2} \\
\vdots \\
v_{i,n}
\end{bmatrix}
\]

(3)

which consists of the \( i^{th} \) multichannel surrogate canonical coordinates for all the channels. The correlation matrix, \( R_{v_i v_i} \), of (3) is

\[
R_{v_i v_i} = E[v_i v_i^T] = \begin{bmatrix}
a_{i,1}^T R_{11} a_{i,1} & a_{i,1}^T R_{12} a_{i,2} & \cdots & a_{i,1}^T R_{1n} a_{i,n} \\
a_{i,2}^T R_{21} a_{i,1} & a_{i,2}^T R_{22} a_{i,2} & \cdots & a_{i,2}^T R_{2n} a_{i,n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{i,n}^T R_{n1} a_{i,1} & a_{i,n}^T R_{n2} a_{i,2} & \cdots & a_{i,n}^T R_{nn} a_{i,n}
\end{bmatrix}
\]

(4)

Clearly, the elements of this correlation matrix are the correlations between different pairs of surrogate canonical coordinates defined in (2).

In the standard CCA [6], assuming that the previous \((i - 1)^{th}\) canonical coordinates are successfully extracted, the \( i^{th} \) canonical coordinate is obtained by maximizing only one correlation, namely, \( a_{i,j}^T R_{ij} a_{i,2} \), while imposing the unit variance constraint, \( a_{i,j}^T R_{ii} a_{i,1} = a_{i,2}^T R_{22} a_{i,2} = 1 \). The multichannel case [5] extends this process to maximizing the sum of all correlations between pairs of surrogate variables in (2) subject to a similar constraint. The constraints analyzed by [5] included: (a) unit variance (i.e., \( a_{i,j}^T R_{ij} a_{i,j} = 1, \forall j \in [1, n] \)) and (b) unit trace (i.e., \( tr\{R_{v_i v_i}\} = \sum_{j=1}^{n} a_{i,j}^T R_{jj} a_{i,j} = 1 \)), where \( tr\{A\} \) denotes the trace of matrix \( A \). Using constraint (a), it is desired to find the \( i^{th} \) set of \( n \) vectors, \( a_{i,j} \) for \( j \in [1, n] \), by solving

\[
\text{Maximize: } q_i = \frac{1}{2} \sum_{j=1}^{n} \sum_{k \neq j}^{n} a_{i,j}^T R_{jk} a_{i,k}
\]

subject to: \( a_{i,j}^T R_{jj} a_{i,j} = 1, \forall j \in [1, n] \).

Alternatively, using the Lagrange multiplier method, the problem can be posed as finding the desired mapping vectors by maximizing:

\[
b_i = q_i - \sum_{j=1}^{n} \lambda_{i,j} (a_{i,j}^T R_{jj} a_{i,j} - 1)
\]

(6)

where \( \lambda_{i,j} \) is the \((i, j)^{th}\) Lagrange multiplier. The optimal solution of (6) yields an eigenvalue problem which has no unique solution [5]. Certainly, using this constraint is analogous to the procedure adopted by [6] for two-channel CCA. However, the result is not a generalized eigensystem. Hence, a different constraint is utilized. The unit trace constraint (b) can be used which leads to the following maximization problem:

\[
b_i = q_i - \lambda_i (tr\{R_{v_i v_i}\} - 1)
\]

(7)

with \( \lambda_i \) being the \( i^{th} \) Lagrange multiplier. The optimal solution of (7) yields the following generalized eigenvalue
problem:
\[
\begin{bmatrix}
R_{11} & R_{12} & \cdots & R_{1n} \\
R_{21} & R_{22} & \cdots & R_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
R_{n1} & R_{n2} & \cdots & R_{nn}
\end{bmatrix}
\begin{bmatrix}
a_{i,1} \\
a_{i,2} \\
\vdots \\
a_{i,n}
\end{bmatrix}
= \lambda_i
\begin{bmatrix}
a_{i,1} \\
a_{i,2} \\
\vdots \\
a_{i,n}
\end{bmatrix}
\]
which is valid for \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{d_1} \) and \( i \in [1, d_1] \).

In the two-channel CCA case, this iterative procedure together with a deflation process [6] yields valid canonical coordinates. However, in the multichannel case a deflation process is still needed to generate these "canonical coordinates." Thus, the coordinates that are obtained using the above-mentioned procedure are referred to as "surrogate canonical coordinates." It is important to note that for this feature extraction method, only the canonical correlations are utilized, not the coordinates. The generalized eigensystem obtained in (8) is used to find these canonical correlations. These correlations are the same as those obtained using (7) and are indeed canonical.

III. BOSS SYSTEM AND DATA COLLECTION

The new Disk BOSS [1] is designed to scan for buried underwater objects using a spherical acoustic source mounted in the same horizontal plane as the receiver array. In order to overcome the increased scattering noise generated by the omnidirectional source, the receiver array is composed of 252 hydrophones utilizing elements mounted with equal spacing on a 1.5m diameter circular disk. The omnidirectional source generates an FM pulse over the band of 3 to 19 kHz. Target backscattering is measured by the 252 hydrophones (or a subset) and is then processed by a CPU to generate a final composite image. BOSS processing involves a sequence of operations that includes chirp sonar processing, 3-D near field focusing using data from a real or synthetic receiving aperture, transmission beam steering that allows target illumination at a broad range of aspect angles, real-time 3-D image construction of buried objects showing burial depth, shape, and orientation and location of buried objects, and real-time buried object detection [9].

A test range located at St. Andrew's Bay in Panama City, Florida contains a number of mine-like and non-mine-like objects buried at various depths beneath the ocean surface. The BOSS vehicle collected sonar data at this test range. In order to collect sonar data, the BOSS vehicle is translated through the 105m long and 10m wide target field. This is done for several passes through the target field. For these tests, the BOSS vehicle was towed through the target field at a speed of approximately 3kts with a ping rate of 10 pings per second. Of the 252 available channels, a 48-channel subset was used to emulate the array structure of the Wing BOSS currently being used.

Buried mine-like objects in the target field included two bomb-shaped markers, a bomb-shaped object, and two large cylinders. Non-mine-like objects included two artillery shells, two smaller cylinders, a large bullet-shaped object, a half-buried steel sphere, and a fully buried sphere. A number of naturally occurring objects (such as rocks) are also present in the target field. Additionally, there are several concrete bricks used as position markers located at various points within the field. Although the target field contains a wide variety of buried objects, only a subset of these are considered for the classification problem in this study. These objects are listed in Table I.

<table>
<thead>
<tr>
<th>#</th>
<th>Description</th>
<th>Class</th>
<th>L4 pings</th>
<th>L2 pings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.75&quot; Bomb-shaped Marker</td>
<td>Mine-like</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6&quot; Iron Cylinder</td>
<td>Mine-like</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>10.75&quot; Bomb-shaped Marker</td>
<td>Mine-like</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>14&quot; Stainless Steel Sphere</td>
<td>N-mine-like</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>Bullet-shaped Object</td>
<td>N-mine-like</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>2&quot; Iron Cylinder</td>
<td>N-mine-like</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Clean Bottom Return</td>
<td>N-mine-like</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>8</td>
<td>Natural Clutter</td>
<td>N-mine-like</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>5.5&quot; Bomb-shaped Object</td>
<td>Mine-like</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

This study uses BOSS data from two test runs through the target field, namely, Lines 4 and 2. For Line 2, 991 pings were collected and, for Line 4, 733 pings were collected. Since the trajectory, orientation, and elevation of the BOSS vehicle varies through the target field between Lines 2 and 4, different numbers of pings are available for each object. Table I shows how many pings of data are available for each object in each run.

In this study, the data collected for objects in Line 4 is used for training and validation purposes, and the data collected for Line 2 is used for testing the robustness and generalization ability of the trained classifiers.

IV. FEATURE EXTRACTION AND CLASSIFICATION

Once extracted, the multichannel canonical correlations are then used as features for classification. This idea was proven effective when applied to underwater target classification on different data sets for the two-channel CCA case [3], [8]. For this application, three-channel MCCA is utilized to extract features to represent mine-like and non-mine-like objects. To build the ensemble sets of these three channels, backscattered signals with a certain ping separation are partitioned into overlapping blocks of range cells. The blocks of signals are then viewed as the realizations of the three input channels, namely, \( x_1 \), \( x_2 \), and \( x_3 \). Each realization of the channel is the time series vector associated with a block of range cells in the corresponding sonar return [3]. In order to generate enough realizations for the three channels, the 200 sample time series obtained from each of the 48 receiver channels are concatenated together. The signals are then partitioned into overlapping blocks of 40 samples, each of
which has 50% overlap (this choice of block size and overlap was determined experimentally). Figure 1 provides a visual representation of this Preprocessing procedure. This process yields a statistically rich data set for computing MCCA coordinates and correlations.

After preprocessing, the feature extraction process is implemented via CCA/MCCA using the newly formed data channels created from sonar pings, as depicted by Figure 1. Figures 2(a) and (b) show the ping separations that are used for the two feature extraction methods. For two-channel CCA, Figure 2(a) shows that channel 1 is formed from the first ping and channel 2 is formed from the third ping in a sequence. For three-channel MCCA, the first two channels are formed using the first two pings in a sequence, respectively. While channel 3 is formed by using the fourth ping in the sequence, as shown in Figure 2(b). These channels are used as inputs for generating a feature vector via CCA or three-channel MCCA. These ping separations have been experimentally determined to produce the best classification results for CCA [8] and three-channel MCCA when applied to the given data set. The canonical correlations produced from these groups of two (for CCA) and three (for MCCA) pings are used as the feature vector that represents the first ping in the group. Once a feature vector has been extracted, this process is repeated for the next set of pings until all sonar pings containing the object of interest are exhausted, as shown in Figures 2(a) and (b). This is done as many times as necessary for each object in order to produce feature vectors for each of the various objects in the data set.

The method described produces 40 canonical correlations associated with each set of pings. In efforts to reduce the complexity of the problem, only the first half of the canonical correlations are used as features in the ensuing classification problem. This is justified since only the first few canonical correlations yield the greatest contribution to the linear dependence or mutual information between the data channels [6]. Figures 3(a) and (b) show the plots of canonical correlation features obtained for mine-like and non-mine-like objects via CCA and MCCA, respectively. From these figures, it can be visually determined that the mine-like and non-mine-like features generated from MCCA are more distinguishable and separable than those generated from CCA. This increased separation aids in improving detection and classification performance.

The training set for the classifiers is randomly formed from half of the mine-like and non-mine-like feature vectors collected for the objects in Line 4. The remaining Line 4 feature vectors are then used to form the validation set. Using this method ensures that an equal number of mine-like and non-mine-like features are used in the classifier training. Once training and validation are completed, testing the classifiers is performed using the feature vectors generated from Line 2 objects. Since the trained classifiers are blind to the testing set, the classification results for Line 2 will indicate the robustness of the extracted CCA/MCCA features as well as the generalization ability of the classifiers.

Next, classification systems able to classify objects based on individual feature vectors produced via both the CCA and MCCA feature extraction methods are developed. Two classifiers are created; one is trained using individual CCA feature vectors, and the other using feature vectors produced via the MCCA method as described in section II. The single-CCA feature vector classifier can be referenced as a 2-ping classifier because each CCA feature vector is computed using two sonar pings and captures the coherence between the pair of pings. The results are then compared with a similar system, i.e., a single-MCCA feature vector classifier utilizing the three-channel MCCA feature vectors created from groups of three pings (see Figures 2(a) and (b)). The classification results provided by the single-MCCA feature vector classifier are also compared to the results generated from a classification system designed to perform a decision-level fusion on the individual CCA classification decisions. The CCA-based decision-level fusion process is implemented in order to produce a classification system that uses the same number and sequence of sonar pings as the single-MCCA feature vector classifier. This is done in order to illustrate that the classification system using the MCCA features actually performs better and has a much simpler structure than the CCA method involving the additional fusion system.

In order to implement a classifier able to differentiate between mine-like and non-mine-like objects, a simple back-propagation neural network (BPNN) [10] is utilized. Here, a BPNN is used due its ability to solve difficult and diverse classification problems. Determining the best network structure is accomplished by performing a number of random weight initializations for various network structures and then evaluating the results on the training and validation data sets. For each feature extraction method, the neural network that provided the best overall classification results on the training and validation data is retained for further study. Once found, these optimal networks are used in the implementation of the single-CCA and MCCA feature vector classifiers.

Although many possible avenues for performing a classification fusion exist, in this study, a BPNN fusion system designed to implement a decision-level fusion [2] is utilized. In
this system, a second BPNN is used in order to combine the normalized outputs of a single-CCA feature vector classifier in order to make a final decision about the class of an object based on two classifier decisions (see Figure 2(a)). Because two CCA feature vectors are used in this system to render a final decision, the final classification result is essentially a nonlinear function of the information contained in a window spanned by the sonar data contained in four consecutive pings. Because of this, the results of this system can be directly compared to those generated by the single-MCCA feature vector classification system, which also utilizes the information contained in a window spanned by the same four pings, as is illustrated in Figure 2(b).

A two-layer 4-6-2 fusion network utilizing four input nodes corresponding to the two 2-D outputs of the single-CCA feature vector classifiers, six hidden layer neurons, and two outputs is used. The network is trained using groups of two CCA-feature vector classifier outputs generated from the corresponding individual CCA feature vectors contained in the training data set. The inputs to this fusion system are formed by concatenating together groups of two consecutive single-CCA feature vector classifier outputs per object so that each set of two 2-D classifier decisions is transformed into a single 4-D vector. Here, the goal of training is to capture the nonlinear mapping present in the sets of consecutive decisions made by the single-CCA feature vector classifier in the training set. Doing so will allow the fusion system to classify an object of interest based on intermediate decisions made by the single-CCA feature vector classifier for that object, as can be seen in Figure 2(a).

V. TEST RESULTS ON BOSS DATA

To find a good structure for the CCA and MCCA networks, five different two-layer BPNN structures were evaluated. During each test, the number of neurons in the hidden layer was varied from 40 to 48. Each structure was allowed to train for 1000 training epochs using five random weight initializations. The best BPNN classifier structure (20-50-2)
was then selected based on the highest classification rates on both the training and validation data sets. Because correct classification of mine-like objects is more important than misclassifications of non-mine-like objects, i.e., false alarms, the classifier that yielded the best results on the mine-like objects group was given precedence over other classifiers with comparable aggregate classification rates.

The three classifiers discussed produced classification rates presented in confusion matrices, as shown in Table II. When applied to the training set, it was found that the single-CCA feature vector classifier was more likely to misclassify feature vectors representing the small cylinder, the large bullet-shaped object, the steel sphere, and the natural clutter when compared to the other objects in the set. This is mainly due to the fact that the classifier was not trained using any of the feature vectors associated with these objects. Thus, correct classification is difficult because the object may appear mine-like at certain orientations.

For the CCA-based decision-level fusion classifier, misclassifications that occurred in the testing set still tend to be caused by the steel sphere and the natural clutter. However, the decision-level fusion was able to substantially improve the classification performance on the large-bullet shaped object and on the small cylinder. In fact, a significant overall improvement is obtained by using the CCA-based decision-level fusion classifier when compared to the single-CCA feature vector classifier for all objects being classified, as shown in Table II. The reason for this improvement is the supplementary discriminatory power introduced through the use of additional information gained by using an additional feature vector in the decision-making process. One downside to the BPNN fusion scheme is that it requires an additional feature vector classifier for all objects being classified, as shown in Table II. The reason for this improvement is the supplementary discriminatory power introduced through the CCA or MCCA features for every ping in a given test run and then applying the appropriate classification system to every feature vector in order to determine the location of possible mine-like objects.

In this portion of the experiment, both CCA and MCCA features were extracted for all available sonar pings for both Lines 2 and 4. Next, all of the resulting CCA and MCCA features were respectively classified using the trained single-CCA feature vector classifier and the trained single-MCCA feature vector classifier. Also tested was the CCA-based decision-level fusion classification system.

The results of this experiment for Lines 4 and 2 are presented in Figures 4(a) and 5(a), respectively. Shown in each of these images are “detection strips” that clearly show which sonar pings were classified as mine-like by a given classification system. In these detection strips, the black-colored portions represent the pings at which the classification system has declared the presence of a mine-like object. Figures 4(b) and 5(b) show the matched filtered images generated for the receiver channel #1 on Lines 4 and 2, respectively. From these images, one can visually verify the locations of the various mine-like and non-mine-like objects in the target field. By a visual comparison, the detection strips and matched filtered images show a spatial relationship indicating where mine-like and non-mine-like objects are located. Figure 4(c) shows the synthetic aperture sonar (SAS) image of Line 4 generated by the BOSS vehicle as it traversed the target field. As can be seen from Figure 4(c), the relative locations of the objects in the SAS image correspond to the same spatial locations in the matched filtered image and in the detection strips. Thus, a spatial relationship has been determined and is shown between Figures 4(a), (b), and (c).

When the detection strips in Figure 4(a) are examined, it is apparent that although the single-CCA feature vector classifier is able to successfully detect all of the mine-like

<table>
<thead>
<tr>
<th>CLASSIFICATION CONFUSION MATRICES.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single-CCA Feature Vector Classifier</strong></td>
</tr>
<tr>
<td><strong>Line 4 Validation Set</strong></td>
</tr>
<tr>
<td><strong>Chosen Class</strong></td>
</tr>
<tr>
<td><strong>Mine-like</strong></td>
</tr>
<tr>
<td><strong>N-Mine-like</strong></td>
</tr>
</tbody>
</table>

| **CCA Decision-level Fusion Classifier** |
| **Line 4 Validation Set** | **Line 2 Testing Set** |
| **Chosen Class** | **Mine-like** | **N-Mine-like** | **Mine-like** | **N-Mine-like** |
| **Mine-like** | 100% | 0% | 94% | 6% |
| **N-Mine-like** | 9% | 91% | 15% | 85% |

| **Single-MCCA Feature Vector Classifier** |
| **Line 4 Validation Set** | **Line 2 Testing Set** |
| **Chosen Class** | **Mine-like** | **N-Mine-like** | **Mine-like** | **N-Mine-like** |
| **Mine-like** | 100% | 0% | 100% | 0% |
| **N-Mine-like** | 4% | 96% | 16% | 84% |
objects, it also produces a large number of false alarms. When the decision-level fusion classifier utilizing CCA features is applied to the Line 4 run, all of the mine-like objects are detected, and the number of false alarms drastically reduced when compared to the results generated by the single-CCA feature vector classifier. This is not surprising given the fact that the decision-level fusion system is utilizing information from a wider range of sonar pings than the single-CCA feature vector classifier. The third detection strip represents the results obtained from applying the single-MCCA feature vector classifier to every sonar ping in the Line 4 run. As can be seen from the image, all mine-like objects are again successfully detected, and the numbers of false alarms have been greatly reduced when compared with the single-CCA feature vector classifier case. In fact, the results obtained using this method tend to indicate slightly better detection/classification performance on the entire run than even the CCA-based decision-level fusion classification system, which utilizes the same number of raw sonar pings (four) to render a final decision about the identity of an object. This is an important result, as it indicates that the feature vectors obtained via MCCA allow for a high degree of correct classification without the need for an additional fusion network as in the CCA-based decision-level fusion system.

The detection strips for Line 2 shown in Figure 5(a) exhibit a much higher frequency of false alarms than the Line 4 case, mostly due to the fact that all of the classification systems being tested were blind to data from this run. Despite this fact, all mine-like objects are correctly detected and classified regardless of the decision-making method used. As in Line 4, the single-CCA feature vector classifier produced the most false alarms. Although a portion of these false alarms are eliminated using the CCA-based decision-level fusion, the best resulting detection run was obtained using the single-MCCA feature vector classifier, once again demonstrating the potential for this system to produce better results using a simpler structure than the CCA-based decision-level fusion system.

VI. CONCLUSIONS AND FUTURE WORK

In this study, canonical correlation analysis is extended for multi-ping feature extraction in real sonar data. The extracted MCCA features are used to differentiate between mine-like and non-mine-like buried underwater objects in the BOSS data set. These features were then used in conjunction with a single-MCCA feature vector classifier to perform a multi-ping classification. The CCA-based decision-level fusion classifier consisted of a BPNN fusion system that utilized multiple intermediate decisions made by the single-CCA feature vector classifier to perform a final decision-level fusion. The results generated by the single-CCA feature vector and CCA-based decision-level fusion classifiers were thoroughly compared to the single-MCCA feature vector classifier results.

As expected, classification results based on feature extraction using the MCCA procedure provided significantly better results than the CCA procedure. This conclusion is arrived at based upon the classification on the validation and testing data sets as well as the overall detection/classification results on the entire runs. The single-MCCA feature vector classifier improved the classification rates for all objects in the Line 4 validation and Line 2 testing data sets over the single-CCA feature vector and the CCA-based decision-level classifier.

The CCA-based decision-level fusion scheme provided comparable results at the expense of additional structural complexity due to adding another BPNN that must be trained and tested in order to perform decision-level fusion. This implies that by extracting features via MCCA, rather than CCA, the additional expense and complexity contained in a decision-level fusion scheme can be avoided without sacrificing performance.

Future work includes performing MCCA by optimizing different cost functions and constraints. In addition, research is needed to develop an iterative approach for extracting multichannel canonical coordinates.

ACKNOWLEDGMENT

This research was supported by the Office of Naval Research (ONR) under contract #N00014-05-1-0014.

REFERENCES

Fig. 4. Line 4 (a) Detection Strips, (b) Matched-filtered Domain Image and (c) SAS Image (Courtesy of FAU Dept. of Ocean Engineering).

Fig. 5. Line 2 (a) Detection Strips and (b) Matched-filtered Domain Image.