Target Detection in M-Disparate Sonar Platforms using Multichannel Hypothesis Testing

James D. Tucker, Nick Klausner, and Mahmood R. Azimi-Sadjadi
Department of Electrical and Computer Engineering, Colorado State University,
Fort Collins, Colorado 80523, Email: dtucker@engr.colostate.edu

***Abstract*—The use of multiple disparate sonars allows one to exploit a high resolution sonar with good target definition while taking advantage of the clutter suppressing abilities of a low resolution broadband sonar co-registered over the same region to provide potentially much better detection and classification performance comparing to those of the single sonar cases. In this paper the standard Neyman-Pearson detector is extended to the dual disparate sonar case allowing target detection across two sensory platforms simultaneously. For three disparate sonar platforms, two detectors are implemented with the final decisions being fused. Test results of the proposed methods on a data set of underwater side-scan sonar imagery are presented. This database contains data from 3 different side-scan sonars, namely one high frequency sonar and two broadband sonars, operating at three different frequencies and bandwidths. The data was collected in different bottom conditions and contains various mine-like and non-mine-like objects with varying degree of difficultly and bottom clutter. Test results illustrate the effectiveness of the proposed detection system in terms of probability of detection, false alarm rate, and the receiver operating characteristic (ROC) curve.**

I. Introduction

A critical need of the U.S. Navy is the development of a reliable, efficient and robust underwater target detection and classification system that can operate in real-time with multiple disparate sensor systems and in different bottom or environmental conditions. In a surveillance area there could be multiple AUV's each equipped with a wide variety of sensors including different types of sonar, magnetics, or electro-optical systems. Preliminary detection, feature extraction and object classification can be performed based upon the data collected using every sensory system on one or multiple vehicles. A final decision-making usually takes place at the central station using some type of a decision-level or a feature-level fusion. However, decision-making based upon individual sensory data typically leads to incomplete, degraded or biased decisions hence resulting in an unacceptable final detection and classification performance at the fusion center.

To allow collaborative decision-making among several sonar or other sensor platforms, it is essential to detect and further scrutinize the information-bearing parts of the data collected by various sensory systems. This involves detecting, isolating and representing, in terms of some pertinent attributes, the *coherent* information among the multiple data sets. This is an extremely challenging problem due to the disparate nature of the problem and variations in the operating and environmental conditions. Thus, to develop a system-level solution, new methodologies are needed to: (a) collaboratively detect and agree on threats occurring within the field of view of the sensors, (b) perform collaborative feature extraction to capture common target attributes from multiple sensor platforms, (c) perform object classification and identification, (d) and finally develop a single integrated target assessment picture based upon the detected, localized and classified targets from multiple disparate sensors.

The existing work in the area of target detection from sonar imagery has primarily been focused on one sonar platform, with fusion across multiple algorithms. The method in [1], [2] utilizes a nonlinear matched filter to identify mine-size regions in the sonar image that match the target signature. For each detected region, several features are extracted based on the size, shape, and strength of the target signature. A stepwise feature selection process is then used to determine the subset of features that maximizes the probability of detection and classification. A k-nearest neighbor and an optimal discrimination filter classifier are used to classify each feature vector and the decisions of the two classifiers are fused for the final decision. An adaptive clutter filter detector was presented in [3] which exploits the difference in correlation characteristics between clutter and targets. After detection, features are extracted and orthogonalized using a mapping matrix such that the elements of the vector posses approximately statistically independent conditional probability density functions. Classification is then performed on the orthogonalized feature set using an optimal Bayesian classifier. In [4], the algorithm described in [3] is individually applied to three different sonar images varying in frequency and bandwidth. Final classification is done using an optimal set of features using a nonlinear log-likelihood ratio test where the decisions of the individual detector and classifier are fused. The optimal set of features is determined based upon cascading another classifier on the previous classifier during the training stage. This is done as a repeated application during the training stage where at each iteration the threshold and optimal feature set is chosen and updated. Chandran in [5] presented the use of a matched filter designed to capture the target structure. Higher order spectra are extracted from the phase of the Fourier transform as the feature set to classify objects using a k-nearest neighbor classifier, a minimum distance classifier, and a threshold-based classifier where the threshold is determined from the minimum and maximum values of a feature obtained over all classes, and the final decisions are fused.

In our previous work, [6] we developed a new framework for dual-sensor coherence analysis using canonical coordinate decomposition (CCD) method [7] - [9] that can be applied to the data collected using two disparate sonar systems. Using the proposed method allows for simultaneous detection and feature extraction of coherent target information among two sonar images. The CCD data channels consist of ROI's of co-registered sonar images. Our *detection hypothesis* in this framework is that presence of objects in the two-platform sonar data leads to high level of coherence measure (or mutual information) comparing to that of the background clutter. A subset of coherent-based features that capture high discriminatory information can then be applied to a simple classifier to discriminate target and non-target objects. Our results [6] on a data set collected using a side-scan HF and a broadband (BB) sonar systems indicated excellent detection and discrimination ability of the proposed method.

The work presented in this paper extends the CCD-based coherent framework in [6] for the hierarchical fusion of multiple-platform sonar imagery for target detection. In this system, the CCD process is applied to a pair of co-registered BB sonar imagery data to determine presence of an object based upon low resolution, but less cluttered images. To improve the confidence in the detection decision, CCD is also applied to a pair of HF and BB sonar imagery to verify that indeed this object also appears in both the high resolution (but more cluttered) HF and in one of the BB images. These two dual-sonar detection decisions are then fused to reduce the number of missed detections and false alarm rate. The data set used in this study was provided by the Naval Surface Warfare Center (NSWC) - Panama City and consists of a high-resolution HF side-looking sonar imagery that contains either no targets, one target, or multiple targets along with two BB side-looking sonar imagery co-registered over the same region.

This paper is organized as follows: Section II reviews the optimum Neyman-Pearson detector for Gaussian random vectors in the CCD framework for both single [10]- [11] and dual-disparate sensor platforms [6]. In Section III, the results of applying the proposed system to the NSWC data-set are presented. Finally, conclusions and observations are made in Section IV.

II. A Review of CCD-Based Detector and Extension to Dual Disparate Sonar Platforms

A. Classical Neyman-Pearson Detector

The classical Neyman-Pearson detection problem [10] is to decide between the two hypotheses of noise only and the other of signal plus noise. Assume there is an observation $\mathbf{y} \in \mathbb{R}^m$ that is an m-dimensional, random normal vector with zero mean and covariance matrix R. We wish to test the hypothesis $H_0 : R = R_0$ (noise alone) versus the hypothesis $H_1 : R = R_1$ (signal plus noise). We will define $R_1 = R_0 + R_s$ where R_0 is the covariance matrix of noise alone and R_s the covariance matrix of the signal. The log-likelihood ratio, $l(\mathbf{y})$, which minimizes the risk involved in deciding between our two hypothesis [10] gives

$$\begin{aligned} \gamma(\mathbf{y}) &= \begin{cases} 1 \sim H_1, & l(\mathbf{y}) > l_0 \\ 0 \sim H_0, & l(\mathbf{y}) \leq l_0 \end{cases} \\ l(\mathbf{y}) &= \mathbf{y}^H Q \mathbf{y} \text{ with } Q = R_0^{-1} - R_1^{-1}, \end{aligned} \tag{1}$$

where l_0 is the decision threshold which is experimentally determined based upon a set of training samples (see Section III). It can be shown [10] that matrix Q can be expressed in terms of the *signal-to-noise ratio matrix*, S, as

$$Q = R_0^{-\frac{H}{2}} (I - S^{-1}) R_0^{-\frac{1}{2}} \tag{2}$$

where $S = R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{H}{2}}$ and $(\cdot)^H$ is the Hermitian operation [12]. If we let $\boldsymbol{\nu} = R_0^{-\frac{1}{2}} \mathbf{y}$, the log-likelihood ratio can then be rewritten as,

$$l(\boldsymbol{\nu}) = \boldsymbol{\nu}^H (I - S^{-1}) \boldsymbol{\nu} \tag{3}$$

The transformed vector $\boldsymbol{\nu}$ is also normal r.v. with zero mean and covariance matrix $R = I$ under the hypothesis H_0 and $R = S$ under H_1.

The J-Divergence [10], which measures the detectability or the distance between the two hypotheses, is

$$\begin{aligned} J &= E_{H_1}[l(\boldsymbol{\nu})] - E_{H_0}[l(\boldsymbol{\nu})] && (4) \\ &= tr(S + S^{-1} - 2I), && (5) \end{aligned}$$

where E_{H_0} and E_{H_1} denote the expectation operator under its respective hypothesis and $tr(\cdot)$ is the trace operation. It is shown in previous work [10] that matrix S can be decomposed using singular value decomposition (SVD) which leads to the generalized eigenvalue problem $R_1 R_0^{-\frac{H}{2}} U = R_0 R_0^{-\frac{H}{2}} U \Lambda$ where the diagonal elements of Λ consist of the eigenvalues of S and U the eigenvectors within its column space. Both (3) and (5) can now be rewritten in terms of the eigenvalues S. Thus, we get

$$\begin{aligned} J &= tr(\Lambda + \Lambda^{-1} - 2I) \\ &= \sum_{i=1}^{m} (\lambda_i + \lambda_i^{-1} - 2) \end{aligned} \tag{6}$$

and

$$l(\boldsymbol{\nu}) = \boldsymbol{\nu}^H (I - \Lambda^{-1}) \boldsymbol{\nu}. \tag{7}$$

B. Neyman-Pearson Detector in CCD Framework.

This detection problem can also be cast in the CCD framework in which the detection test and detection criterion are formed in terms of the canonical coordinates and canonical correlations [11]. In the two-channel CCD framework, channel $\mathbf{x} \in \mathbb{R}^m$ corresponds to noise alone data (i.e. $\mathbf{x} = \mathbf{n}_1$) whereas channel $\mathbf{y} \in \mathbb{R}^m$ could be either noise under H_0 (i.e. $\mathbf{y} = \mathbf{n}_2$) or signal plus noise under H_1 (i.e. $\mathbf{y} = \mathbf{s} + \mathbf{n}_2$). Figure 1 shows how the two-channel CCD can be used as an optimal detector.

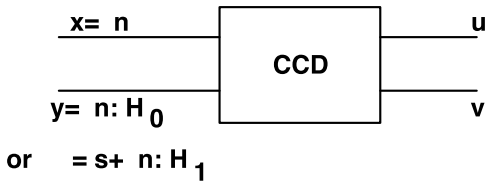


Fig. 1. CCD used as an optimal detector.

Under hypothesis H_1, the composite vector $\mathbf{z} = [\mathbf{x}^H \mathbf{y}^H]^H$ has the composite covariance matrix

$$R_{zz} = E[\mathbf{z}\mathbf{z}^H] = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} = \begin{bmatrix} R_0 & R_0 \\ R_0^H & R_s + R_0 \end{bmatrix} \tag{8}$$

The *coherence* matrix [7] is then formed using $C = R_{xx}^{-1/2} R_{xy} R_{yy}^{-H/2}$, where we have $R_{xx}^{-1/2} R_{xx} R_{xx}^{-H/2} = I$, $R_{xx}^{1/2} R_{xx}^{H/2} = R_{xx}$. The SVD of the coherence matrix C yields

$$C = R_{xx}^{-1/2} R_{xy} R_{yy}^{-H/2} = FKG^H \quad \text{and} \quad F^H C G = K, \tag{9}$$

where $F \in \mathbb{R}^{m \times m}$ and $G \in \mathbb{R}^{m \times m}$ are orthogonal matrices, i.e., $F^H F = FF^H = I$, $G^H G = GG^H = I$, and K is a singular value matrix, with $K = \text{diag}[k_1, k_2, \ldots, k_m]$ and $1 \geq k_1 \geq k_2 \geq \ldots \geq k_m > 0$.

The *canonical coordinates* of $\mathbf{x}$ and $\mathbf{y}$ are then defined as [7]

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} F^H & 0 \\ 0 & G^H \end{bmatrix} \begin{bmatrix} R_{xx}^{-1/2} & 0 \\ 0 & R_{yy}^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \tag{10}$$

Consequently, $W^H = F^H R_{xx}^{-1/2}$ and $D^H = G^H R_{yy}^{-1/2}$ are the matrices that map $\mathbf{x}$ and $\mathbf{y}$ to their corresponding canonical coordinates

$$\mathbf{u} = W^H \mathbf{x} \quad \text{and} \quad \mathbf{v} = D^H \mathbf{y} \tag{11}$$

The elements of $\mathbf{u} = [u_i]_{i=1}^m$ are referred to as the canonical coordinates of $\mathbf{x}$ and the elements of $\mathbf{v} = [v_i]_{i=1}^m$ are referred to as the canonical coordinates of $\mathbf{y}$. The canonical coordinate vectors $\mathbf{u}$ and $\mathbf{v}$ share the cross-correlation matrix $R_{uv} = E[\mathbf{u}\mathbf{v}^H] = K$ while $R_{uu} = E[\mathbf{u}\mathbf{u}^H] = I$ and $R_{vv} = E[\mathbf{v}\mathbf{v}^H] = I$. The cross-correlation matrix K,

$$K = E[\mathbf{u}\mathbf{v}^H] = E[(W^H \mathbf{x})(D^H \mathbf{y})^H] = F^H C G \tag{12}$$

is called the *canonical correlation matrix* of *canonical correlations* $k_i, i = 1, \ldots, m$. Thus, the canonical correlations measure the correlations between pairs of corresponding canonical coordinates. That is, $E[u_i v_j] = k_i \delta_{ij}$; $i \in [1, m]$, $j \in [1, n]$, with δ_{ij} being the Kronecker delta. The canonical correlations k_i are also the singular values of the coherence matrix C.

The dominant (largest) canonical coordinates are the ones that maximize the coherence between the two channels $\mathbf{x}$ and $\mathbf{y}$, under the constraint of maximizing the rate at which the mutual information is captured. In other words, the dominant canonical coordinates of the two channels capture most of the coherent features between them with a minimum dimensional feature set.

As shown in [11], using the CCD framework the log-likelihood and J-Divergence in (3) and (5), can be expressed in terms of the canonical coordinates and canonical correlations. More specifically, using the relationship $S = (CC^H)^{-1} = R_0^{-1/2} R_1 R_0^{-H/2}$ (under hypothesis H_1), the log-likelihood can be rewritten as [11]

$$\begin{aligned} l(\mathbf{y}) &= (R_{xx}^{-1/2}\mathbf{y})^H ([(CC^H)^{-1} - I]^{-1} - CC^H)(R_{xx}^{-1/2}\mathbf{y}) \\ &= (G^H R_{yy}^{-1/2}\mathbf{y})^H ([I - K^2]^{-1} - I)(G^H R_{yy}^{-1/2}\mathbf{y}) \end{aligned} \tag{13}$$

Alternatively, we have

$$l(\mathbf{v}) = \mathbf{v}^H ([I - K^2]^{-1} - I)\mathbf{v} \tag{14}$$

Thus, the log-likelihood ratio is expressed in terms of the canonical coordinate vector $\mathbf{v} = G^H R_{yy}^{-1/2}\mathbf{y}$ for the $\mathbf{y}$ channel and the canonical correlation matrix K. More specifically, $l(\mathbf{v})$ is the weighted sum of the magnitude-squared of the canonical coordinates weighted by canonical correlation-dependent weights, i.e.

$$l(\mathbf{v}) = \sum_{i=1}^{m} |v_i|^2 \left(\frac{k_i^2}{1 - k_i^2} \right) \tag{15}$$

where $v_i = \mathbf{g}_i^H R_{yy}^{-1/2}\mathbf{y}$ and $\mathbf{g}_i$ is the i^{th} column of the matrix G.

It can also be shown [11] that the J-divergence between H_1 and H_0 can be expressed solely in terms of the K matrix or canonical correlations as

$$J = tr([I - K^2]^{-1} - I - K^2) = \sum_{i=1}^{m} \frac{k_i^4}{1 - k_i^2}. \tag{16}$$

The function $\frac{k_i^4}{1-k_i^2}$ is non-increasing in the interval (0, 1]. *Consequently the rank-r detector that maximizes the divergence is the detector that uses the canonical coordinates corresponding to the r-dominant canonical correlations* $k_1, \cdots, k_r$. The J-divergence between the two hypotheses considering r dominant canonical correlations is then,

$$J_r = \sum_{i=1}^{r} \frac{k_i^4}{1 - k_i^2}. \tag{17}$$

Thus, for building low-rank detectors, only the dominant canonical coordinates need to be retained in order to find the coherence between the two data channels $\mathbf{x}$ and $\mathbf{y}$.

C. Dual-Sensor Platform CCD-Based Detector

When considering two disparate sensors observing the same scene at different elevation, grazing angle, beam-width, and frequency, the application of CCD in Section II-B slightly differs. In this case, the hypothesis H_0 corresponds to the case when both sonar image region of interest's (ROI's) contain background noise only while H_1 corresponds to the case where both ROI's contain target (signal) and background noise. Therefore, we have, $H_0 : \mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix}$ and

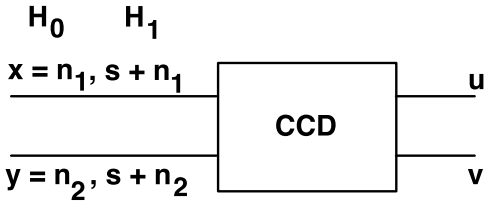


Fig. 2. CCD used as an optimal detector for Dual Disparate Case.

$H_1 : \mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{s}+\mathbf{n}_1 \\ \mathbf{s}+\mathbf{n}_2 \end{bmatrix}$. Figure 2 depicts this dual-sensor detection problem. It is assumed in this analysis that $\mathbf{s}$, $\mathbf{n}_1$, and $\mathbf{n}_2$ are mutually uncorrelated i.e. $E[\mathbf{n}_i\mathbf{n}_j^H] = 0$, and, $E[\mathbf{s}\mathbf{n}_i^H] = 0$, and $E[\mathbf{n}_i\mathbf{n}_i^H] = R_0$ for $i \neq j$ and $i,j \in [1,2]$. In this case, the covariance matrices $\bar{R}_0$ and $\bar{R}_1$ under hypothesis H_0 and H_1, respectively, yield the following Neyman-Pearson detection test for $\mathbf{z}$,

$$\begin{aligned} \gamma(\mathbf{z}) &= \begin{cases} 1 \sim H_1, & l(\mathbf{z}) > \bar{l}_0 \\ 0 \sim H_0, & l(\mathbf{z}) \leq \bar{l}_0 \end{cases} \\ l(\mathbf{z}) &= \mathbf{z}^H \bar{Q} \mathbf{z} \text{ with } \bar{Q} = \bar{R}_0^{-1} - \bar{R}_1^{-1}. \end{aligned} \quad (18)$$

where, $\bar{R}_0$ and $\bar{R}_1$ are defined as,

$$\bar{R}_0 = \begin{bmatrix} R_n & 0 \\ 0 & R_n \end{bmatrix} \quad (19)$$

$$\bar{R}_1 = \begin{bmatrix} R_s + R_n & R_s \\ R_s & R_s + R_n \end{bmatrix}. \quad (20)$$

Note that we assumed $E[\mathbf{n}_i\mathbf{n}_i^H] = R_n = R_0$, for $i = 1, 2$ and $E[\mathbf{s}\mathbf{s}^H] = R_s$ where R_s and R_n were defined previously in Section II-A.

The J-divergence measure for this case becomes

$$\begin{aligned} J &= E_{H_1}[l(\mathbf{z})] - E_{H_0}[l(\mathbf{z})] \\ &= tr(-2I + \bar{R}_0^{-1}\bar{R}_1 + \bar{R}_1^{-1}\bar{R}_0). \end{aligned} \quad (21)$$

In [6], we have shown that the J-divergence in (21) can be related to the eigenvalues (λ_i's) of the signal-to-noise ratio matrix S, defined in Section II-A as

$$\begin{aligned} J &= tr\left\{-2I + 2\Lambda(I - \frac{1}{2}\Lambda^{-1})^{-1}\right\} \\ &= \sum_{i=1}^{n} -2 + \frac{4\lambda_i^2}{2\lambda_i - 1}. \end{aligned} \quad (22)$$

However, as shown in [11], the eigenvalues of S (λ_i's) are indeed related to the squared canonical correlations (k_i^2's) using $\lambda_i = \frac{1}{1-k_i^2}$. Thus, the J-divergence in (22) can alternatively be expressed as

$$J = \sum_{i=1}^{n} -2 + \frac{4}{1 - k_i^4} \quad (23)$$

Additionally, in [6] we have also shown that the log-likelihood ratio in (18), can be represented in terms of canonical correlations k_i's and coordinates $\mathbf{u}$ and $\mathbf{v}$ of $\mathbf{x}$ and $\mathbf{y}$ channels (see Section II-B) as

$$\begin{aligned} l(\mathbf{w}) &= \mathbf{u}^H \left[\left(\frac{1}{2}\right) I - \left(\frac{1}{4}\right)\left[K^{-2} - \left(\frac{1}{2}\right) I\right]^{-1}\right] \mathbf{u} \\ &+ 2\mathbf{u}^H \left[\left(\frac{1}{2}\right) K^{-1} - \left(\frac{1}{4}\right)\left[K^{-1} - \left(\frac{1}{2}\right) K\right]^{-1}\right] \mathbf{v} \\ &+ \mathbf{v}^H \left[\left(\frac{1}{2}\right) K^{-2} - \left(\frac{1}{4}\right)\left[I - \left(\frac{1}{2}\right) K^2\right]^{-1}\right] \mathbf{v} \end{aligned} \quad (24)$$

Using the above formulations we have related the dual sensor version of the Neyman-Pearson back to the standard case (one sonar platform) for which expressions are in terms of the corresponding canonical coordinates and correlations. Note that this formulation assumes that all blocks within ROI's of the sonar image are of the same dimensions, $m \times m$.

III. Test Results

The CCA-based detector is applied to a three-platform sonar data set consisting of a HF high-resolution side-looking sonar images as well as two BB sonar images. More information on high-resolution side-looking and broadband sonar can be found in [13] and [14], respectively. As mentioned before, HF sonar provides higher spatial resolution and better ability to capture target details and characteristics while BB sonar offers much better clutter suppression ability but lower resolution. Here we provide a summary of our most recent results using two coherent-based detectors as shown in Figure 3. As can be seen in this figure, the first detector uses one HF image and one BB image with a different spatial resolution, while the second detector uses that same BB image and a second BB image with different frequency characteristics. Both detectors are implemented using our two-channel coherence-based method in [6]. The outputs of each detector is fused to generate a final decision (soft decision). While in the first detector, coherence information in the two BB sonar images is exploited to further suppress clutter ROI's and detect potential targets, the coherence information in the second detector solidifies the decision in the first detector by verifying the joint presence in the HF and BB images. The fusion of the decisions using the additions of the soft decisions scores leads to reduced number of false alarms and misdetections.

The image database used contains 59 co-registered images containing 53 targets with some of the images containing more than one target. The optimum ROI size was experimentally determined to be 72×112 pixels for the HF sonar images and 24×224 for both of the BB sensors. This is due to the disparity in the spatial resolution of these sonar systems. Each sonar image is then partitioned into ROI's with a 50% overlap along both the range and cross-range directions. Each ROI is then channelized using a rectangular blocking scheme of which the dimensions are 6×4 and 2×8 for the HF and two BB's, respectively. Using the CCA method outlined in Section II-C, canonical correlations and coordinates are extracted for each ROI pair of the HF-BB_1 and BB_1-BB_2.

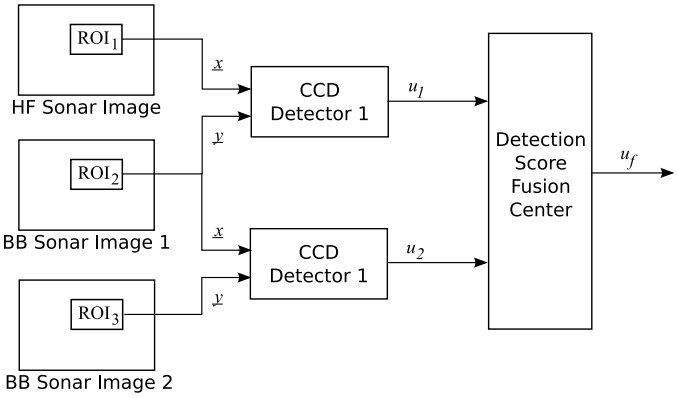
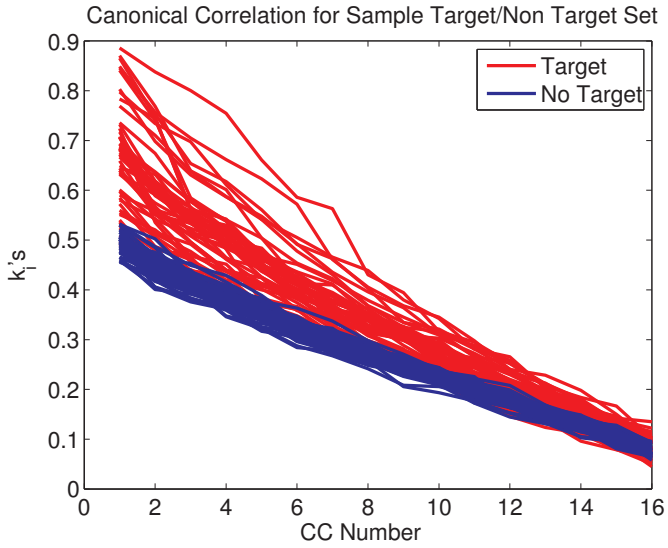
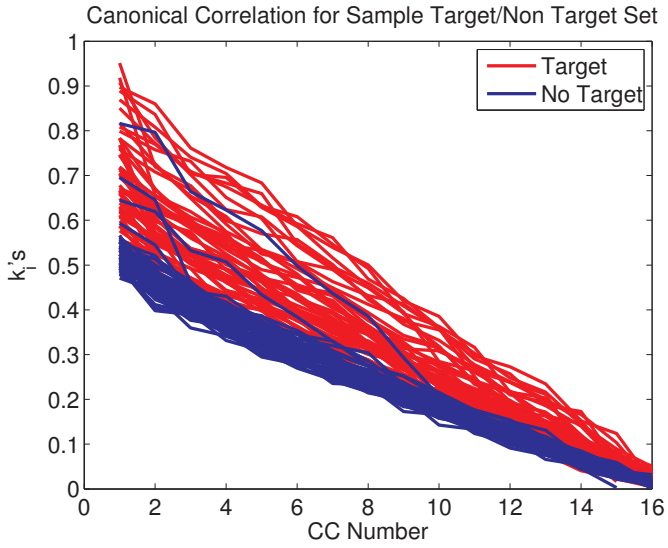


Fig. 3. Block diagram of the Multi-platform Detection System.



(a) High Frequency Broadband Detector



(b) Broadband Broadband Detector

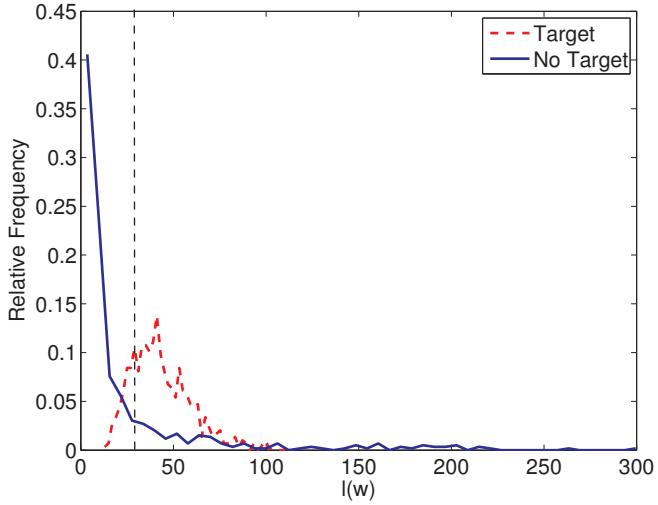
Fig. 4. Plot of Canonical Correlations for Target and Background for both Detectors.

To show the separability of the principal canonical correlations for ROI's that contain targets over background and those that contain only background, a test was conducted on the entire target set ROI's and a same size randomly selected set of ROI's containing only background clutter. The plots of 16 canonical correlations of ROI's containing targets and those containing background only are shown in Figures 4(a) and (b) for the HF-BB_1 and BB_1-BB_2 detectors, respectively. As can be seen, for both pairs of sonar images there is suitable separation among the canonical correlations formed from target ROI's and those from background ROI's, especially for the principal canonical correlations.

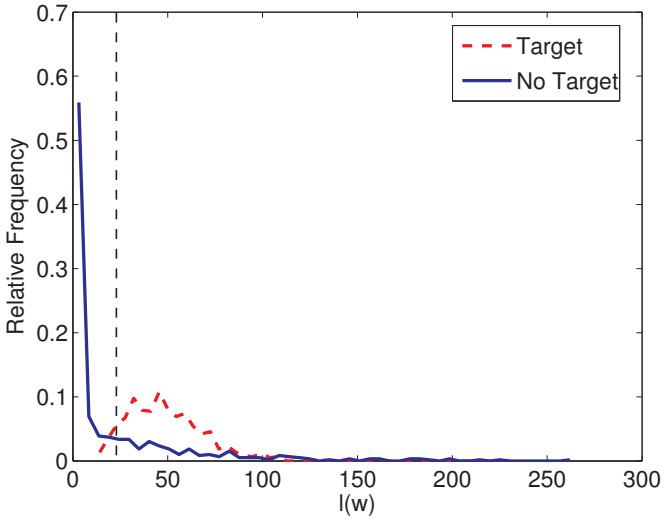
Next, for the detection process, the log-likelihood ratio expression in (24) was found for each block within ROI pair. A detection score was then created based on the percent of the log-likelihood measurements within ROI par that fall above the detection threshold. Detection score $\geq 50\%$ signifies presence of a target within that pair of ROI's. The detection scores of the two detectors are then sent to the fusion center. At the fusion center the scores are added together and if the score is higher than a pre-selected fusion threshold a final decision is declared.

Figures 5(a) and (b) show the histograms of the log-likelihood ratio values of one target ROI and one random background ROI in the HF-BB_1 and BB_1-BB_2 cases, respectively. Using the entire set of target and background ROI's, an optimum threshold was determined to be 26.4 for the HF-BB_1 and 23.9 for the BB_1-BB_2, respectively. These thresholds are shown by dotted vertical lines in Figures 5(a) and (b). These values were then used as detection thresholds for all the ROI's. That is, any block within an ROI pair whose log-likelihood ratio falls above the relevant threshold is flagged as a target block in either HF-BB_1 or BB_1-BB_2 pairs. To determine the threshold for the fusion process, the same test set was used and "soft" decision scores were found for each ROI in the test set. Figure 6 presents the histogram of these scores. As can be seen, there is a good separability in the sum of the decision scores. A threshold of 1.02 was experimentally determined to be adequate for the final decision making.

The system is then implemented on the entire NSWC multi-platform sonar imagery data. The individual coherence-based detectors performed marginally well. More specifically, the HF-BB_1 detector detected 51 of the 53 targets with an average of 10.13 false detections per image. While the BB_1-BB_2 detector performed a little worse with detection of 49 of the 53 targets and an average of 9.86 false detections per image. However, when the decisions of these two detectors are fused, the performance was greatly improved. After the fusion process, the system detected 52 of the 53 targets with an average of only 8.4 false detections per image. The ROC curves for the individual detectors and the fusion of the two



(a) High Frequency Broadband Detector



(b) Broadband Broadband Detector

Fig. 5. Histogram of Example Log-Likelihood Values for Target and Background for both Detectors.

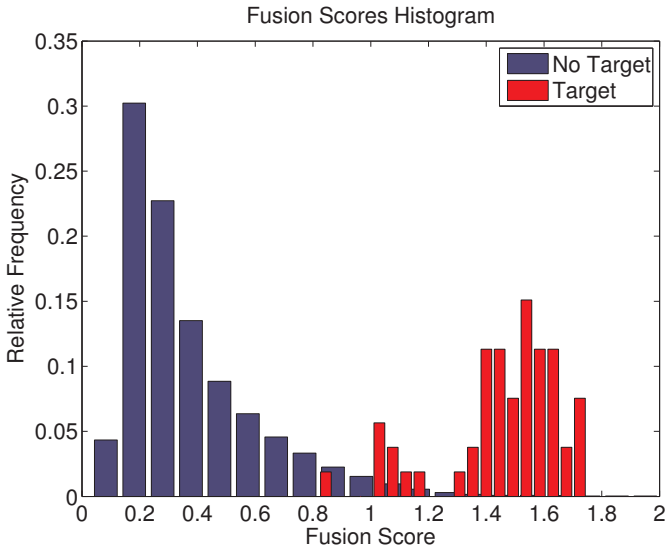


Fig. 6. Histogram of the Fused Decision Scores for the Test Set.

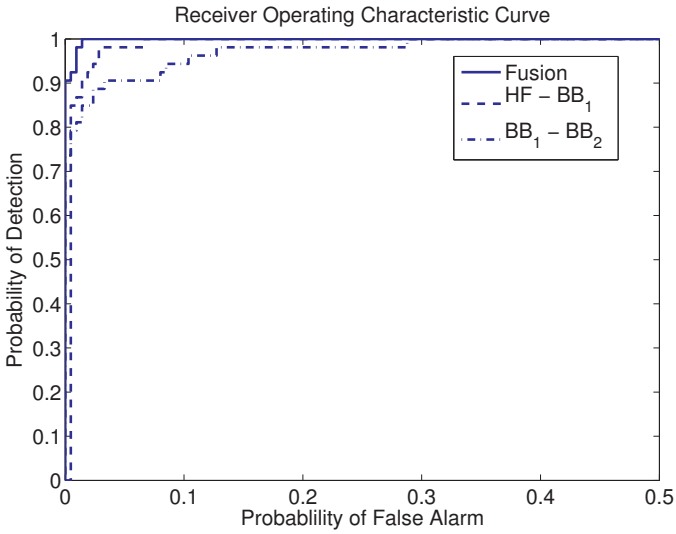


Fig. 7. ROC Curves for the Individual Detectors and the Fusion

are presented in Figure 7. At the knee point for the HF-BB_1 detector we have $P_d = 92\%$ and $P_{fa} = 8\%$. While for the BB_1-BB_2 detector the knee point exhibits $P_d = 93\%$ and $P_{fa} = 7\%$. After the fusion of the two detectors, the knee point of the fused ROC gives $P_d = 99\%$ and $P_{fa} = 1\%$, which demonstrates excellent overall detection performance of the proposed system. The only target that was missed is at very close range and has faint signatures in all the three images. This is due to the fact that this target was close to the tract of the AUV. Consequently, the signature was extremely dark and hard to visually discern in all three images hence leading to overall low coherence and subsequent misdetection. Overall, the detector performed extremely well given the small number of targets and non-targets used to form the detection threshold.

IV. Conclusion

In this work, the CCD method was used to develop a dual-sonar platform coherence-based detector that utilizes the advantages of using side-scan sonar imagery with disparate frequency and bandwidth characteristics. The CCD-based detector was first formulated for the standard (single-sensor) optimum Neyman-Pearson detector and then was extended to the case where dual disparate sonar platforms are used. The detector was then applied to a three-platform sonar imagery dataset which contains one HF and two BB sonar images co-registered over the same region. The coherence-based detector was implemented on the combination of two BB sonar images and then was implemented on a pair of HF and BB sonar imagery. The soft decisions of these two coherence-based

detectors were fused to yield a final decision. Using a BB with a BB sonar, one can determine the presence of an object based upon less cluttered, low resolution imagery. This decision is further solidified by using another detector which exploits the coherence across the HF and BB by verifying that the object appears in both the high resolution (more cluttered) HF and in one of the BB images. The two individual detectors performed fairly well. However, the problem occurs in some sonar images where, due to disparity and frequency in the sonar characteristics, the HF sonar falsely detects a target that does not appear in the co-registered BB sonar images. Thus, a fusion decision was formed by combining the "soft" decisions of each individual detector and applying a threshold. The fusion process significantly improved the probability of correct detection while reducing the incidence of a false alarm. At the knee point of the ROC, the system provided a $P_d = 99\%$ and $P_{fa} = 1\%$, or alternatively only one target misdetection and an average of 8.40 false detections per image on the entire NSWC database.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research, Code 321OE under contract #N00014-08-1-0142

REFERENCES

[1] G. J. Dobeck, J. Hyland, and L. Smedley, "Automated detection/classification of sea mines in sonar imagery," *Proc. SPIE*, vol. 3079, pp. 90–110, April 1997.

[2] G. J. Dobeck, "Fusing sonar images for mine detection and classification," *Proc. SPIE*, vol. 3710, pp. 602–614, April 1999.

[3] T. Aridgides, P. Libera, M. Fernandez, and G. J. Dobeck, "Adaptive filter/feature orthogonalization processing string for optimal LLRT mine classification in side-scan sonar imagery," *Proc. SPIE*, vol. 2765, pp. 110–121, April 1996.

[4] T. Aridgides and M. Fernandez, "Enhanced atr algorithm for high resolution multi-band sonar imagery," *Proc. SPIE*, vol. 6953, pp. 0H1–0H10, March 2008.

[5] V. Chandran, S. Elgar, and A. Nguyen, "Detection of mines in acoustic images using higher order spectral features," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 3, pp. 610–618, July 2002.

[6] M. Azimi-Sadjadi and J. D. Tucker, "Target detection from dual disparate sonar platforms using canonical correlations," *Proc. SPIE*, vol. 6953, pp. 0J1–0J10, March 2008.

[7] L. Scharf and C. Mullis, "Canonical coordinates and the geometry of inference, rate, and capacity," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 824–891, March 2000.

[8] A. Pezeshki, M. Azimi-Sadjadi, and L. Scharf, "Undersea target classification using canonical correlation analysis," *Oceanic Engineering, IEEE Journal of*, vol. 32, no. 4, pp. 948–955, Oct. 2007.

[9] J. D. Tucker, M. R. Azimi-Sadjadi, and G. J. Dobeck, "Canonical coordinates for detection and classification of underwater objects from sonar imagery," *OCEANS 2007 - Europe*, pp. 1–6, 18–21, June 2007.

[10] L. L. Scharf and B. D. Van Veen, "Low rank detectors for Gaussian random vectors," *IEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1579–1582, Nov 1987.

[11] A. Pezeshki, L. Scharf, J. K. Thomas, and B. D. Van Veen, "Canonical coordiantes are the right coordinates for low-rand Gauss-Gauss detection and estimation," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4817–4820, Dec 2006.

[12] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. John Hopkins University Press, 1996.

[13] W. Key, "Side scan sonar technology," *Proc. Oceans'00*, vol. 2, pp. 1029–1033, Sept 2000.

[14] S. Butler, "Triply resonant broadband transducers," *Oceans '02 MTS/IEEE*, vol. 4, pp. 2334–2341 vol.4, 29-31 Oct. 2002.