FUSING SONAR IMAGES FOR MINE DETECTION AND CLASSIFICATION

Gerald J. Dobeck

NSWC Coastal Systems Station, Dahlgren Division, Panama City, FL 32407-7001

ABSTRACT

An image fusion method is developed for sonar systems that collect multiple images in a single pass. This research builds on our past work with automated detection and classification of sea mines in side-looking sonar imagery.

The new method has the following processing steps. Each image is processed separately by our automated detection and classification algorithm[2]. Next detections in the images are collated into two general categories: (1) detections appearing in only one image and (2) detections appearing in multiple images. A fuzzy-logic procedure is used to fuse the detections that are common to multiple images. The final "fused" classifications will fall under two general descriptions: (1) classifications where sufficient mine-like evidence was contained in a single image and (2) classifications where sufficient mine-like evidence had to be accumulated over multiple images. This fusion process dramatically reduces false alarms.

Keywords: sea mines, naval mines, automatic detection, automatic classification, sensor fusion, data fusion, side-looking sonar, side-scan sonar, synthetic-aperture sonar, sonar imagery

1. INTRODUCTION

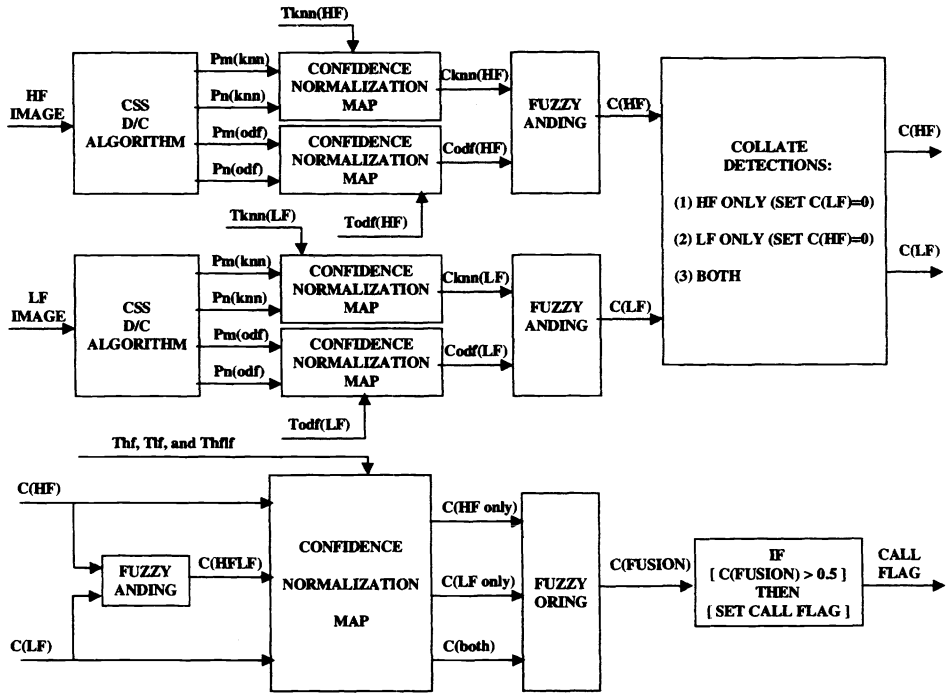
The Navy is moving more and more towards unmanned minehunting sonar systems because they are portable, can operate in hard-to-reach shallow water regions, and keep people out of harms way. They often carry high-resolution side-looking sonars that are effective at detecting and classifying bottom sea mines in the complex littoral environment. However, this environment gives rise to many false alarms caused by natural, biologic, and man-made clutter resulting in a slow down of mine countermeasure operations. There has been much research focused on the development of automated detection and classification methods for these systems that reduce false alarms while still maintaining a high probability of mine detection and classification[1-21].

Current side-looking sonar systems collect a single image during one pass. In our current research, the benefits of image fusion on false alarm reduction for sonar systems, which collect multiple images in one pass, are being investigated. This effort builds on our past research on automated detection and classification of sea mines in side-looking sonar imagery as reported at AeroSense'95[1] and AeroSense'97[2].

The focus of this paper is the problem of fusing images from a high-frequency synthetic-aperture sonar (HFSAS) and a low-frequency synthetic-aperture sonar (LFSAS).

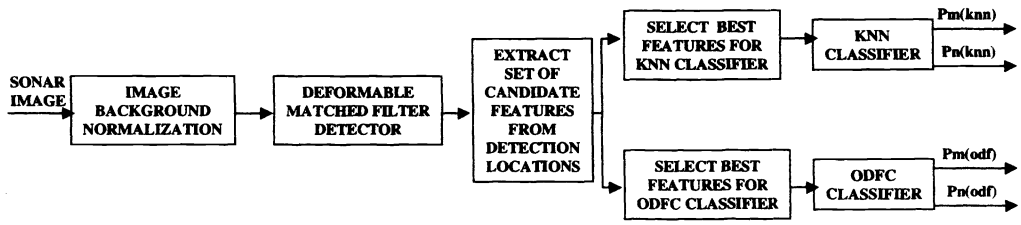
Our new method has the following processing steps. The HF and LF images are processed separately by our automated detection and classification algorithm[2]. Next detections in the images are collated into three categories: (1) detections appearing in the HF image, (2) detections appearing in the LF image, and (3) detections appearing in both images. Then a fuzzy-logic procedure is used to fuse the detections that are common to both images. The final "fused" classifications will fall under two general descriptions: (1) classifications where sufficient mine-like evidence was contained in a single image and (2) classifications where sufficient mine-like evidence had to be accumulated over multiple images. This fusion process dramatically reduces false alarms.

Part of the SPIE Conference on Detection and Remediation Technologies for Mines and Minelike Targets IV • Orlando, Florida • April 1999
SPIE Vol. 3710 • 0277-786X/99/$10.00



Pm designates degree of mine-likeness
Pn designates degree of nonmine-likeness
C designates mine-like confidence

Figure 1. Image Fusion



Pm designates degree of mine-likeness
Pn designates degree of nonmine-likeness

Figure 2. CSS's Detection and Classification Algorithm

2. DETECTION AND CLASSIFICATION

The strategy employed for detection (D), classification (C), and image fusion is diagramed in Figure 1. As shown in the figure each image is processed separately by the D/C algorithm. In the detection stage, the image is screened for mine-like objects. The main purpose of the detection stage is to screen the entire image quickly and reduce it to a few target-size regions that will be examined in much more detail by the

classification stage. In the classification stage, the degree of mine-likeness, Pm, and nonmine-likeness, Pn, are computed based on features extracted from each of the detection regions. CSS's D/C algorithm is illustrated in Figure 2, and more details can be found in Reference 2. Next, the list of detections found in both images, containing Pn, Pm, and the detection locations, is input to the fusion stage. In this stage detections are collated into the three categories described above (HF only, LF only, and both).

2.1 Modified Image Normalizer

Since our last publication, the detection stage has undergone two major modifications to improve robustness. These will be described next. The first modification was to the image normalizer[3]. The purpose of the normalizer is to reduce the variability of the local mean throughout the image. The improved normalizer produces a more consistent background so that highlight and shadow pixels are more discernable. This is a crucial step in preparing the image for the matched filter. The normalizer uses a forwards and backwards filter to estimate the local background level. This estimate is divided into the image, producing a nominal background level, B_{Level}, of about unity. The normalization process preserves sharp edges on the target highlight and shadow. Figure 3 shows an example of the normalization. Note that the signature of mine-like object near the center of image has not been smeared by the normalization process, and the highly reflective sand has been normalized.

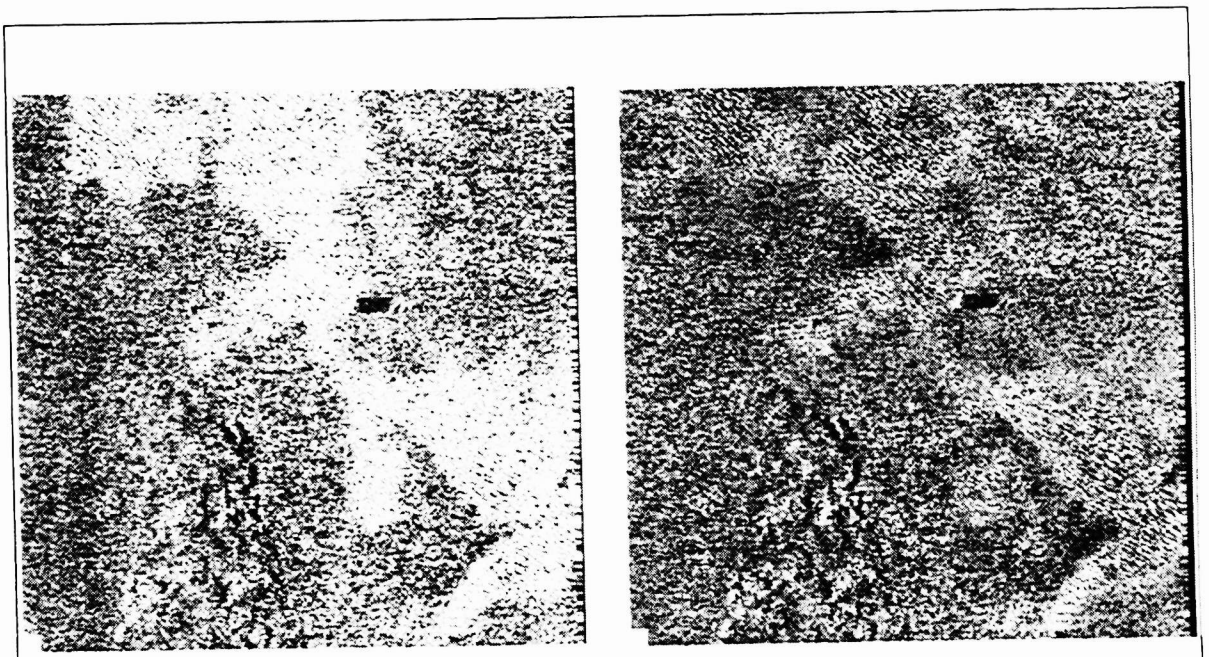


Figure 3. Image Normalization

2.2 Detection Strategy

The workhorse of our detection method is the nonlinear matched filter. Prior to matched-filtering, the normalized image is often decimated (after pixel averaging) to decrease the computations of the matched-filter. The matched filter is scanned over the entire normalized image (and possibly decimated). After the matched filter image is computed, it is normalized. This is done by first dividing the matched-filter image into range regions (in our applications, three regions are used: short, medium, and long range). The mean and variance of the pixels in each range region are computed. These are used to normalize the matched-filter image to zero mean and unit variance. Finally, the normalized pixel values are limited to be greater than or equal to zero.

To determine where the detected objects are in the image, a target-size window is scanned over the normalized matched-filter image. Within a window, pixels are counted that are greater than a detection-amplitude threshold. If this count exceeds a detection-count threshold, then a detection is declared. All closely located detections are clustered into a single detected object. A typical window size is about 1.5-meters by 1.5-meters.

2.3 Deformable Matched Filter

The second modification was to the matched filter. The need for this change arose as more experience was gained with a wider-variety of mine types and mine orientations. The original matched filter approach called for three fixed matched filter templates: one for short range, one for medium range, and one for long range. This fixed template structure did not permit the flexibility needed to handle a greater variety of highlight, shadow, and deadzone lengths. Here "length" refers to down-range length, and "deadzone" refers to the region between the end of the highlight and the start of the shadow. Extending this approach by increasing the number of templates, as a function of range and mine type, did not seem an efficient approach. It would be computationally intensive, and more templates would be needed as new mine types were added. What was desired was a single generic template whose principal purpose was screening; i.e., a single generic filter that could quickly scan the entire image, eliminate most of it, and identify regions of interest.

A deformable matched filter was constructed to accomplish this. For the HFSAS, the deformable matched filter uses 3 shifting templates: (1) pre-highlight, (2) highlight, and (3) shadow. Because shadow is not a reliable phenomenon in LFSAS (due to diffraction and multi-path effects), pre-target, highlight, and post-target templates were used for the LFSAS. The post-target template is handled like the pre-target template. In the remainder of this section, the HFSAS deformable matched filter is described; the LFSAS filter is similarly defined.

The 3 match-filter templates are one-pixel wide in the cross-range direction. As discussed in our earlier publications, this generates a robustness to mine orientation. The match filter is operated only range-wise on the image.

The computation of the deformable matched-filter output, corresponding to a given pixel location (referred to as the reference location), is done in three steps. First, the pre-highlight template is computed over a pre-specified region relative to the reference location. The maximum value over the region is saved. For the second and third steps, this procedure is repeated for the highlight and shadow template. Finally, the value of the matched filter at the reference location is determined from the 3 maximum values. In general, the pre-specified regions for the pre-target, highlight, and shadow templates are different, but they can overlap.

The specific equations for the deformable matched filter are:

$$
\begin{aligned}
S_{ij} &= \max_{k\in[M_S,N_S]} \sum_{t=m_S}^{n_S} \max\left[C_S\left(\min\{I(i+k+t,j),\ I_{max}\}-B_{Level}\right),\ 0\right] \\
H_{ij} &= \max_{k\in[M_h,N_h]} \sum_{t=m_h}^{n_h} \max\left[C_h\left(\min\{I(i+k+t,j),\ I_{max}\}-B_{Level}\right),\ 0\right] \qquad (1) \\
P_{ij} &= \max_{k\in[M_p,N_p]} \sum_{t=m_p}^{n_p} \max\left[C_p\left(\min\{I(i+k+t,j),\ I_{max}\}-B_{Level}\right),\ 0\right]
\end{aligned}
$$

$$
I_{MF}(i,j) = \text{matched filter image} = \max[\ S_{ij} + H_{ij} - P_{ij},\ 0\ ] \qquad (2)
$$

where

$I(i,j)$ = normalized image; i = range index; j = cross-range index

S_{ij} = shadow contribution to matched-filter output

H_{ij} = highlight contribution to matched filter output

P_{ij} = pre-target contribution to matched filter output

$C_{p/h/s}$ = matched-filter coefficients for pre-target/highlight/shadow templates

$M_{p/h/s}$ & $N_{p/h/s}$ = shift limits (in range direction) of pre-target/highlight/shadow templates

$n_{p/h/s} - m_{p/h/s} + 1$ = length of pre-target/highlight/shadow templates

B_{Level} = normalized image's average background level

I_{max} = saturation limit

$C_S = 1 / [(T_S - B_{Level})(n_S - m_S + 1)]$

$C_h = 1 / [(T_h - B_{Level})(n_h - m_h + 1)]$

$C_p = 1 / [(T_p - B_{Level})(n_p - m_p + 1)]$

$T_{p/h/s}$ = reference pre-target/highlight/shadow pixel levels

2.4 Discussion of Deformable Matched Filter

The coefficients, $C_{p/h/s}$ are determined such that the contribution of the pre-target, highlight, and shadow templates (S_{ij}, H_{ij}, and P_{ij}) are equal to unity when evaluated at their respective reference levels, $T_{p/h/s}$. The thresholds $T_{p/h/s}$ are selected based on a mine's average pre-target/highlight/shadow characteristics. Pixels, which are within the highlight template and less than the background level, are considered "don't care" pixels. That is, we do not want them to reduce the impact of bright highlight pixels that are also within the template. This is accomplished via the **max** function inside the summation of equation (1). Similar reasoning applies to the "don't care" pixels within the shadow and pre-target templates. The usual linear matched filter does not permit "don't care" pixels. The saturation limit, I_{max}, prevents high pixels from dominating the contribution.

The deformable matched filter finds the "darkest " shadow contribution, "brightest" highlight contribution, and most anomalous pre-target contribution over their respective regions (defined by $M_{p/h/s}$ & $N_{p/h/s}$). Properly selected values for $T_{p/h/s}$ fall in the following ranges: $T_p > B_{Level}$, $T_h > B_{Level}$, and $T_s < B_{Level}$. Note from equation (2) that the shadow and highlight contributions increase the matched-filter output while the pre-target contribution decreases the matched-filter output. For our application, the pre-target/highlight/shadow template lengths ($n_{p/h/s} - m_{p/h/s} + 1$) and shifts ($M_{p/h/s}$ & $N_{p/h/s}$) have been chosen so that the matched-filter will produce a high output for a wide-variety of mine types and orientations and a low value for most background textures. For example, consider the following specifications. Lengths are given in terms of number of pixels, where the pixel size (range direction) is typical for a high-resolution side-looking sonar.

Typical length of a bright highlight segment: 7 pixels
Maximum range over which to search for the highlight segment: 25 pixels
Typical length of a dark shadow segment: 25 pixels
Maximum length over which to search for the shadow segment (includes deadzone): 60 pixels
Typical length of an anomalous pre-target segment: 25 pixels
Maximum length over which to search for the anomalous pre-target segment: 48 pixels

For such a class of mine-like targets, a reasonable selection of matched-filter parameters are shown in Table 1.

Typical Parameters of the Deformable Matched Filter			
	p	h	s
$n_{p/h/s}$	12	3	12
$m_{p/h/s}$	-12	-3	-12
$[M, N]_{p/h/s}$	[-60, -13]	[-12, 12]	[13, 72]
$T_{p/h/s}$	3 B_{Level}	2 B_{Level}	0
I_{max}	4 B_{Level}	4 B_{Level}	4 B_{Level}

Table 1

2.5 Classification Strategy

Our classification strategy is described in our 1997 SPIE paper[2]. A brief outline follows. After the detected objects of interest have been found in the image, these objects are fed to the classification stage to determine the confidence that they are mine-like. The principal purpose of the classification stage is to eliminate the false detections while preserving the detected mines. Our classification stage uses two classifiers: (1) K-Nearest Neighbor Attractor-based Neural Network (KNN), and the Optimal Discrimination Filter Classifier (ODFC). The KNN is a probabilistic-type neural net that uses hyper-spheres to define class regions in feature space. The ODFC is a classifier based on linear discrimination theory that uses hyper-planes to define class regions in feature space. At each detection location, a large set of features is extracted from the neighborhood surrounding the detection. Features are extracted from both the normalized image and matched-filter image. A "best" subset is selected for the KNN, and another "best" subset is selected for ODFC. When both classifiers declare the detected object as mine-like, the object is classified as mine-like.

3. IMAGE FUSION STRATEGY

In this section the image fusion process, depicted in Figure 1, is developed. The HF and LF images are processed separately by the D/C algorithm. The D/C algorithm produces a list of detections for each image. In order to proceed, our fusion strategy requires that each detection be assigned a mine-like confidence that can be properly compared and combined within the fusion process. Specifically, the mine-like confidence of a HF detection, referred to as C(HF), and the mine-like confidence of a LF detection, referred to as C(LF), need to be determined. This is accomplished using the outputs of the KNN and ODFC classifiers. C(HF) and C(LF) are determined in two steps. First the outputs of KNN and ODFC are normalized, and second they are combined using a fuzzy ANDING.

3.1 Determination of C(HF) and C(LF)

The KNN and ODFC classifiers each have two outputs: degree of mine-likeness, Pm, and the degree of nonmine-likeness, Pn, which are assumed to be nonnegative. Refer to these as Pm(knn), Pn(knn), Pm(odf), and Pn(odf). In general, one cannot directly compare the outputs of KNN and ODFC; e.g., the fact that Pm(knn) is greater than Pm(odf) does not have significant meaning since the classifiers are quite different. To produce a legitimate comparison, the outputs of the KNN and ODFC classifiers must be normalized. This is done by selecting two thresholds, Tknn and Todf such that the performance of each classifier operating alone is similar in terms of probability of classification (Pc) and probability of false alarm (Pfa). Specifically,

If Pm(knn) > Tknn Pn(knn), then "mine-like" is declared.

If Pm(odf) > Todf Pn(odf), then "mine-like" is declared.

Tknn and Todf are selected so both these classification rules perform similarly in terms of Pc and Pfa. These rules are equivalent to,

$$\frac{Pm(knn)}{Pm(knn) + Tknn\ Pn(knn)} > 0.5 \quad (\text{The relation is the same for ofd.})$$

This suggests the following mapping from Pm(knn) and Pn(knn) to a confidence, Cknn, that will produce the desired normalizing effect. Namely,

$$Cknn = \frac{Pm(knn)}{Pm(knn) + Tknn\ Pn(knn)} \quad (\text{Codf is defined similarly.}) \tag{3}$$

This mapping has the following desirable properties when $Pm \geq 0$, $Pn \geq 0$, and $T \geq 0$:

1. $0 \leq C \leq 1$.
2. For any classifier, $C = 0.5$, if $Pm = T\ Pn$.
3. $C = 0$, if $Pm = 0$, and $Pn > 0$.
4. $C = 1$, if $Pn = 0$, and $Pm > 0$.
5. C increases monotonically with Pm.
6. C decreases monotonically with Pn.

This implies the confidence of all classifiers are equivalent (normalized) (1) at their respective thresholds, (2) at Pm = 0, and (3) at Pn = 0. Furthermore, (5) and (6) imply an appropriate monotonic extrapolation as Pm and Pn vary.

The above procedure is applied to the HF detections to generate Cknn(HF) and Codf(HF), and separately to the LF detections to generate Cknn(LF) and Codf(LF).

Now the overall mine-like confidence, C(HF), for HF detections is determined by the fuzzy ANDING of Cknn(HF) and Codf(HF) (C(LF) is defined similarly),

$$C(HF) = \textbf{min}\ (\ Cknn(HF), Codf(HF)\ ) \tag{4}$$

3.2 HF/LF Fusion

The mapping in equation (3) and the fuzzy ANDING in equation (4) produce the mine-like confidences, C(HF) and C(LF), that are used in the subsequent image fusion process. First, C(HF) is determined for all HF detections, and C(LF) is determined for all LF detections. Second, the HF and LF detections are collated into 3 categories: (1) detections appearing only in the HF image, (2) detection appearing only in the LF image, and (3) detections appearing in both images. For detections from category (1), C(LF) is set to zero; for detections from category (2), C(HF) is set to zero. The confidence, C(HFLF), that both the HF and LF jointly show mine-likeness is given by the fuzzy ANDING of C(HF) and C(LF),

$$C(HFLF) = \textbf{min}\ (\ C(HF), C(LF)\ )$$

The image fusion rule can now be stated,

IF { (C(HF) > Thf) **or** (C(LF) > Tlf) **or** (C(HFLF) > Thflf) },
THEN { The detection is declared as mine-like. }, (5)

where Thf, Tlf, and Thflf are thresholds with values between 0 and 1.

With properly selected thresholds (generally, Thf > Thflf and Tlf > Thflf), this decision rule provides the flexibility to make a call (1) when there is sufficient evidence in either image alone or (2) when evidence needs to be accumulated over both images.

3.3 Determining the Fusion ROC Curve and Selecting Thresholds Thf, Tlf, and Thflf

The thresholds are determined as follows. First, a receiver operating characteristics (ROC) curve (Pc versus number of false alarms, Nfa) is determined from training data. The training data is a detection list containing C(HF), C(LF), and class label (mine, nonmine). It should be somewhat large. The fusion ROC curve, (Pc, Nfa), is defined as follows. Using equation (5) the value for Nfa associated with a specific Pc is the minimum Nfa over the space of all three thresholds that gives the same Pc. For each (Pc, Nfa) pair, the minimizing thresholds are saved. Since there are only 3 thresholds, a practical way to generate the ROC curve is to search over a discretized threshold space. To do this each threshold interval [0,1] is divided into N equally spaced points. Then the training detection list is evaluated for each of the NxNxN threshold combinations. For the example described in section 4.1, N was 31; the computation of the ROC curve took only a few seconds on our computer workstation. Finally, the thresholds are chosen corresponding to a point on the ROC curve that produces good Pc and Nfa performance.

3.4 Computing the Mine-Like Confidence of the Fusion Process

The fusion confidence, C(fusion), is computed as follows. Using the same mapping formulation that was used above, define

$$C(\text{HF only}) = \frac{C(\text{HF})}{C(\text{HF}) + Ahf\,(\,1 - C(\text{HF})\,)}$$

$$C(\text{LF only}) = \frac{C(\text{LF})}{C(\text{LF}) + Alf\,(\,1 - C(\text{LF})\,)} \qquad (6)$$

$$C(\text{both}) = \frac{C(\text{HFLF})}{C(\text{HFLF}) + Ahflf\,(\,1 - C(\text{HFLF})\,)}$$

where Ahf = Thf / (1 - Thf). Alf and Ahflf are defined similarly.

Then C(fusion) is given by the fuzzy ORING of C(HF only), C(LF only), and C(both),

$$C(\text{fusion}) = \textbf{max}\,[\,C(\text{HF only}),\, C(\text{LF only}),\, C(\text{both})\,] \qquad (7)$$

Note that the image fusion rule, stated by equation (5), is equivalent to

IF { C(fusion) > 0.5 }, **THEN** { The detection is declared mine-like. }, (8)

4. RESULTS

4.1 SAS Database Evaluation

The HF/LF SAS databases, Sonar 6 and Sonar 7, from the Naval Sensor Data database at the NSWC Coastal Systems Station, Panama City, FL, were used to evaluate the fusion method. The SAS data was collected at sites in St. Andrews Bay and the Gulf of Mexico near Panama City, FL. The Bay site had a sand/mud bottom type, and the Gulf site had a dominantly sand bottom. All mine-like objects were on the sea floor.

The Sonar 6 and 7 contains 78 pairs of HF/LF SAS images. Each image is about 661 range pixels by about 1200 cross-range pixels. Both the HF and LF images are coregistered with same pixel size.

There are 236 mine-like bottom objects in the databases. These objects are of 4 types (of different geometrical shape). Two are large cylindrical-like objects and two are small. The range length of these object's highlights varied from about 8 to 24 pixels; the cross-range length varied from about 10 to 48 pixels. The range length of these object's shadows varied from about 24 to 72 pixels; the cross-range length varied like the highlight's cross-range extent. The range length of the deadzone between the end of the object's highlight and the start of its shadow varied from about 0 to 24 pixels. The HF images of the objects possessed shadows, while the LF images of the object's did not (due to diffraction and multi-path effects). A pair of HF and LF image segments are shown in Figure 4; the images are normalized. The 4 objects near the bottom in the diamond pattern are the 4 mine-like types.

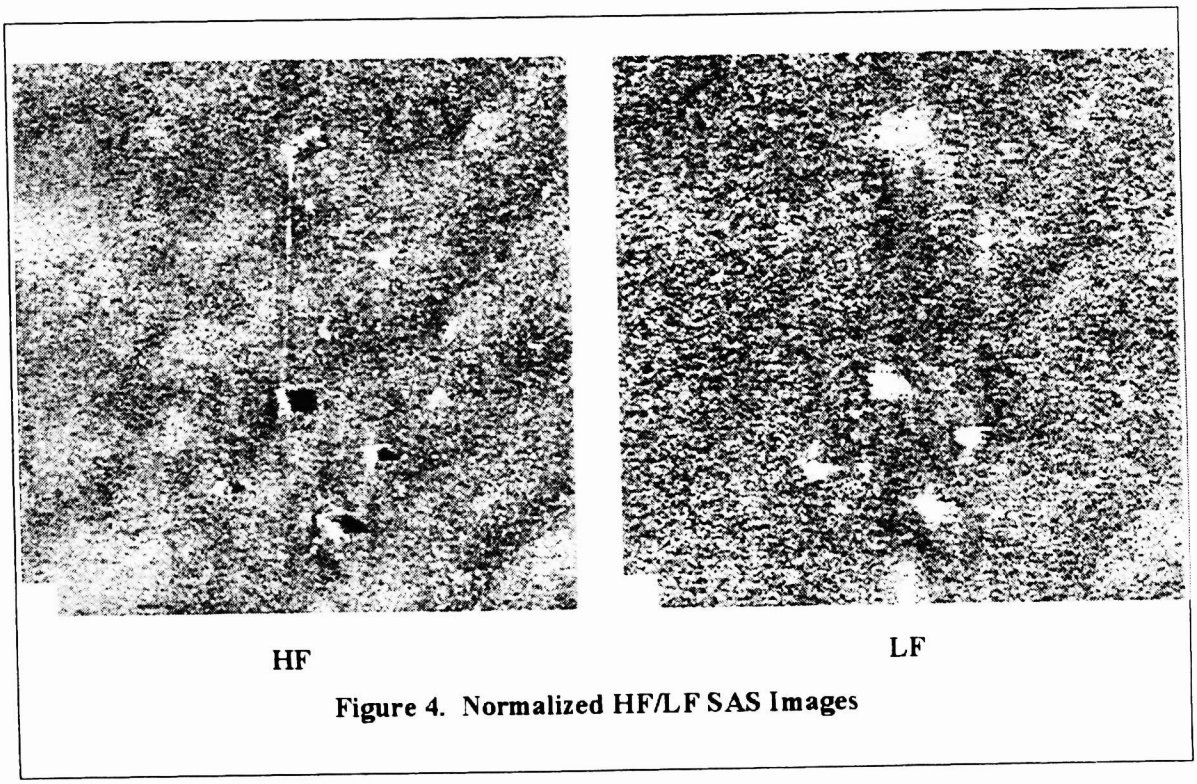


Figure 4. Normalized HF/LF SAS Images

Half the image pairs were used for training and half for testing. After the D/C algorithm was trained with the HF training images, all the HF images were processed by the algorithm. Similarly for the LF images. For the HFSAS database, the deformable matched filter parameters were selected that gave a probability of detection, Pd, of 99% at an average number of false alarms per image (FAI) of about 5.0. For the LFSAS database, parameters were selected that gave a Pd of 100% at an FAI of about 3.0. The fusion process was conducted in the following steps in accordance with Figure 1.

1. Thresholds Tknn(HF), Todf(HF), Tknn(LF), and Todf(LF) are determined as described in the beginning of section 3.1.
2. Detections between the HF and LF images are collated into 3 categories: detections only in HF, detection only in LF, and detections in both HF and LF.
3. For each detection the mine-like confidences, C(HF), C(LF), and C(HFLF), are determined as described in section 3.2.
4. The image fusion ROC curve is determined as described in section 3.3, and an operating point was selected.
5. For each detection the mine-like fusion confidence, C(fusion), is determined as described in section 3.4 using the thresholds Thf, Tlf, and Thflf assosiated with the selected operating point.

6. All detections, where C(fusion) > 0.5, are declared mine-like.

The fusion results are shown in Table 2 for two operating points on the fusion ROC curve. Rather than Pc versus Nfa, the table is given in terms of probability on mine detection and classification (PdPc) versus the average number of false alarms per image (FAI). Also shown in the table, for comparison purposes, are results of using HF alone and the results for using LF alone.

Average False Alarms per Image			
PdPc	**HF alone**	**LF alone**	**HF/LF Fusion**
95%	**0.41**	**0.49**	**0.13**
100%	**1.00**	**0.72**	**0.56**

Table 2. Image Fusion Results

4.2 Algorithm Fusion and MUDSS Evaluation

Under ONR research support, CSS has been advocating algorithm fusion as a means to dramatically reduce false alarm rate and produce robust mine detection and classification in a variety of environments. Under a 6.2 MCM research program, the Office of Naval Research (ONR 321TS) has funded the development of three D/C algorithms: (1) CSS algorithm, described in this paper, (2) the Lockheed Martin (LM) algorithm[5,16], and the Draper Lab algorithm[7,14]. The fusion of these algorithms is currently under investigation.

Substantial benefit will arise from the fusion of D/C algorithms, when they are significantly different. By "significantly different" is meant that the algorithms have been developed based on different mathematical, geometrical, and statistical theories. This results in feature spaces and class boundaries that are constructed in very different ways. Thus, each algorithm is keying on substantially different mine-like characteristics. When a single algorithm's parameters are set for a high PdPc, then it usually follows that the false alarm rate is also high. This is often seen when the D/C system operates in new environment that differs from the ones used in training and development. However, our initial research suggests that, if the algorithms are "significantly different" and set with high PdPc parameter settings, then their and-ing will be very beneficial. They will have few false alarms in common while calling most mines in unison.

This can be explained as follows. Each algorithm looks for its own special set of mine-like characteristics that it can find reliably in noise. Since a mine possesses the mine-like characteristics of all the algorithms, it follows that the algorithms are very likely to see the same mines when their parameters are set for high PdPc. However, few nonmines will possess the mine-like characteristics of all the algorithms.

It is extremely difficult to develop a single D/C algorithm from limited databases that will be robust in a wide-variety of environments. By "robust" is meant that algorithms, developed using data collected in the in one area of the world (e.g., Gulf of Mexico), will perform similarly in the another (e.g., Mediterranean). Furthermore, limited databases are the rule because of cost and other practical concerns. Algorithm fusion appears to be a viable solution to this problem.

This concept was recently tested under Mobile Underwater Debris Survey System (MUDSS) program[4]. The MUDSS system was equipped with the same HF/LF SAS's used to collect database described above (Sonar 6 and Sonar 7), as well as electro-optic and magnetic sensors. It is designed to find unexploded ordinance in underwater defense sites used since World War I. The site for the MUDSS test was a shallow-water bay site near Eglin Air Force Base, in Northwest Florida. The CSS, LM, and Draper algorithms were implemented in real-time on 2 Dec Alpha 21164-EV56 processors that were part of the MUDSS real-time processing system. Each processor was rated at about 960 MFLOPS; one processor executed the 3 algorithms for the port-side HF/LF SAS's and the second processor was used for the starboard-side.

Among the objects planted at the test site to evaluate the MUDSS system were 500, 1000, and 2000 pound bombs. These were very mine-like and offered an opportunity to evaluate D/C algorithms in a new environment. CSS and LM developed and trained their algorithms with the Sonar 6 and Sonar 7 databases (78 image pairs). Draper developed and trained their algorithm with databases Sonar 1 and Sonar 2 from CSS that were collected by different sonars (215 pairs of images). The very shallow water environment at the MUDSS test site was acoustically very difficult, and the bay was very cluttered. The environment was much worse than the environments where Sonar 1, Sonar 2, Sonar 6, and Sonar 7 data were collected. Based on the databases, the D/C algorithms' parameters were tuned to high PdPc levels (about 95% for CSS and about 98 % for LM and Draper).

Preliminary analysis for a 12-kilometer run from the MUDSS test has been done. At the high PdPc parameter levels, results from this difficult environment show that 2781 objects were called by at least one algorithm (CSS called 338, LM called 927, and Draper called 2015). The comparison of the number of calls made by each algorithm make sense because CSS set their parameters for a lower PdPc level than LM and Draper, and Draper's parameters were tuned for different sonars. However, only 79 calls were made by all 3 algorithms, and these 79 calls contained all the mine-like bombs. Thus, algorithm fusion (simple and-ing) resulted in a 97% reduction in false alarm calls.

The benefit of algorithm fusion was further demonstrated by the following. The use of CSS's algorithm alone resulted in 338 calls. The and-ing of CSS and LM's algorithms resulted in 157 calls. The and-ing of CSS, LM and Draper's algorithms resulted in 79 calls. In each stage, the addition of a new algorithm to the fusion process substantially reduced the number of false calls.

In this test, algorithm fusion demonstrated remarkable robustness with a 97% false alarm reduction (while still calling all mine-size bombs), despite the fact that (1) the new environment was very difficult and different than that used for training, (2) the training databases were small, (3) the algorithms were tuned to have High PdPc for the benign training data, and (4) Draper's algorithm was tuned to different sonars.

It is believed that, if the thresholds in the algorithms were adjusted for a 90% PdPc (which is more appropriate for a difficult environment), then the fused results would be significantly less that 79, and all bombs would still have been called by all 3 algorithms. This is currently under study.

5. SUMMARY AND CONCLUSIONS

In summary, this paper presented the following.

1. A new sonar image normalizer was described that equalizes the background in a sonar image, thus making highlight and shadow pixels more discernable.
2. A deformable matched filter was described that adapts to a wide-variety of mine types and mine orientations.
3. A method was described on how to normalize the outputs of different classifiers so that they can be compared and combined in the fusion process.
4. A method for fusing multiple sonar images was described.
5. The benefits of the fusion multiple detection and classification algorithms were discussed.
6. Image fusion results for high-frequency and low-frequency synthetic-aperture sonars (HF/LF SAS's) demonstrated the benefits of the new normalizer, the deformable matched filter, and image fusion.
7. A second set of results for HF/LF SAS's demonstrated the benefits of the fusion of the 3 different detection and classification algorithms developed by CSS (described in this paper), Lockheed Martin (LM), and Draper Lab. The 3 algorithms were implemented for real-time execution on two Dec Alpha 21164-EV56 processors (each rated at about 960-MFLOPS).

In conclusion, it was shown that image fusion could significantly reduced false alarm rate. Using the CSS HF/LF SAS database, the fusion of HF and LF imagery reduced false alarm rate by more than a factor of 3 over using a single sonar alone. In addition, fusion of multiple D/C algorithms was shown to be a very robust means to reduce false alarms while still preserving a high probability of mine detection and

classification. This was demonstrated in a recent MUDSS test using HF/LF SAS's. The test site was a difficult, very shallow water environment. For the test, the CSS, LM, and Draper algorithms were successfully implemented for real-time execution. Test results showed that the fusion (simple and-ing) of the 3 algorithms reduced false alarms by 97%.

ACKNOWLEDGEMENTS

This work was supported by the Office of Naval Research (Dr. Randy Jacobson, ONR 321TS) as part of the 6.2 mine countermeasures program element.

REFERENCES

1. J.C.Hyland, G.J.Dobeck, "Sea Mine Detection and Classification Using Side-Looking Sonar," Proceedings of SPIE'95, pp. 442-453, Vol. 2496, Orlando, Florida, 17-21 April 1995.
2. G.J.Dobeck, J.C.Hyland, L.Smedley, "Automated Detection/ Classification of Sea Mines in Sonar Imagery," Proceedings of SPIE'97, Vol. 3079, pp. 90-110, Orlando, Florida, 20-25 April 1997.
3. D.H.Kil, F.B.Shin, and G.J.Dobeck, "Reducing Model- and Data-Mismatch Errors in Image Compression by Combining Region Dependent Transform-Coefficients Encoding with Normalized and Content-Adaptive Coding," to be submitted to IEEE Journal of Oceanic Engineering.
4. J.D.Lathrop, J.F.McCormick, P.J.Bernstein, J.T.Bono, D.J.Overway, G.S.Sammelmann, "Mobile Underwater Debris Survey System (MUDSS) Feasibility Demonstration Report," Proceedings of the UXO FORUM 1996, pp. 427-436, Williamsburg, Virginia, 26-28 March 1996.
5. T. Aridgides, M. Fernandez, G. Dobeck, "Adaptive Clutter Suppression and Fusion Processing String for Sea Mine Detection-Classification in sonar Imagery," Proceedings of SPIE'98, Vol. 3392, pp. 243-254, Orlando, Florida, 13-17 April 1998.
6. M. Fernandez, T. Aridgides, G. Dobeck, "Adaptive Order-Statistics Filters for Sea Mine Classification," Proceedings of SPIE'98, Vol. 3392, pp. 255-263, Orlando, Florida, pp 13-17 April 1998.
7. M. Bello, "Parametric and Non Parametric ROC Confidence Bound Construction in the Context of Acoustic-Magnetic Fusion Systems for Minehunting," Proceedings of SPIE'98, Vol. 3392, pp. 1163-1178, Orlando, Florida, 13-17 April 1998.
8. F. Shin, D. Kil, G. Dobeck, "Compression of High-Frequency Sonar Imagery in Shallow Water using Image-Content Adaptive Transformation and Coding," Proceedings of SPIE'98, Vol. 3392, 264-277, Orlando, Florida, 13-17 April 1998.
9. B. Lau, T. Chao, "Aided Target Recognition Processing of MUDSS Sonar Data," Proceedings of SPIE'98, Vol. 3392, pp. 234-242, Orlando, Florida, 13-17 April 1998.
10. W. Szymczak, W. Guo, J. Rogers, "Mine Detection Using Variational Methods for Image Enhancement and Feature Extraction," Proceedings of SPIE'98, Vol. 3392, pp. 286-296, Orlando, Florida, 13-17 April 1998.
11. W. Guo, W. Szymczak, J. Rogers, "Mutiresolution Neural Networks for Mine Detection in Side Scan Sonar Images," Proceedings of SPIE'98, Vol. 3392, pp. 297-305, Orlando, Florida, 13-17 April 1998.
12. N. Neretti, Q. Huynh, N. Intrator, G. Dobeck, "Image Enhancement and Feature Extraction for Pattern Recognition," Proceedings of SPIE'98, Vol. 3392, pp. pp. 306-315, Orlando, Florida, 13-17 April 1998.
13. T.Aridgides, M.Fernandez, G.Dobeck, "Adaptive 3-Dimensional Range-Crossrange-Frequency Filter Processing String for Sea Mine Classification in Side-Scan Sonar Imagery," Proceedings of SPIE'97, Vol. 3079, pp. 111-122, Orlando, Florida, 20-25 April 1997.
14. M.G.Bello, "Acoustic/Magnetic Fusion System Architecture Variants and Their Classification Performance," Proceedings of SPIE'97, Vol. 3079, pp. 138-159, Orlando, Florida, 20-25 April 1997.
15. F.Shin, D.Kil, "An Integrated Approach to Bandwidth Reduction and Mine Detection in Shallow Water with Reduced-Dimension Image Compression and Automated Target Recognition Algorithms," Proceedings of SPIE'97, Vol. 3079, pp. 203-212, Orlando, Florida, 20-25 April 1997.
16. T.Aridgides, P.Libera, M.Fernandez, G.Dobeck, "Adaptive-Filter/Feature Orthogonalization Processing String For Optimal LLRT Mine Classification in Sonar Imagery," Proceedings of SPIE'96, Vol. 2765, pp. 110-121, Orlando, Florida, 9-12 April 1996.

17. M.G.Bello, "Hierarchical Multilayer Perceptron Network-Based Fusion Algorithms For Detection/Classification of Mines Using Multiple Acoustic Images and Magnetic Data," Proceedings of SPIE'96, Vol. 2765, pp. 84-109, Orlando, Florida, 9-12 April 1996.
18. S.R.Nelson, S.M.Tuovila, "Automated Recognition of Acoustic-Image Clutter," Proceedings of SPIE'96, Vol. 2765, pp. 122-129, Orlando, Florida, 9-12 April 1996.
19. T.Aridgides, D.Antoni, M.Fernandez, G.Dobeck, "Adaptive Filter for Mine Detection and Classification in Side-Scan Sonar Imagery," Proceedings of SPIE'95, Vol. 2496, pp. 475-486, Orlando, Florida, 17-21 April 1995.
20. M.G.Bello, "Markov Random-Field-Based Anomaly Screening Algorithm," Proceedings of SPIE'95, Vol. 2496, pp. 466-474, Orlando, Florida, 17-21 April 1995.
21. S.R.Nelson, S.M.Tuovila, "Fractal-Based Image Processing For Mine Detection," Proceedings of SPIE'95, Vol. 2496, pp. 454-465, Orlando, Florida, 17-21 April 1995.