The Mondrian Detection Algorithm for Sonar Imagery

David P. Williams

NATO Science and Technology Organization

Centre for Maritime Research and Experimentation (CMRE)

Viale San Bartolomeo 400, 19126 La Spezia (SP), Italy

Phone: +39 0187-527-439, Fax: +39 0187-527-330

E-mail: david.williams@cmre.nato.int

Abstract— **A new algorithm called the Mondrian detector has been developed for object detection in high-frequency synthetic aperture sonar (SAS) imagery. If a second (low) frequency-band image is available, the algorithm can seamlessly exploit the additional information via an auxiliary prescreener test. This flexible single-band and multi-band functionality fills an important capability gap. The algorithm's overall prescreener component limits the number of potential alarms. The main module of the method then searches for areas that pass a subset of pixel-intensity tests. A new set of reliable classification features has also been developed in the process. The overall framework has been kept uncomplicated intentionally in order to facilitate performance estimation, to avoid requiring dedicated training data, and to permit delayed real-time detection at sea on an autonomous underwater vehicle (AUV). The promise of the new algorithm is demonstrated on six substantial data sets of real SAS imagery collected at various geographical sites that collectively exhibit a wide range of diverse seafloor characteristics. The results show that – as with Mondrian's art – simplicity can be powerful.**

Index Terms— **Object detection, mine countermeasures (MCM), synthetic aperture sonar (SAS), multi-band data.**

I. Introduction

A. Motivation

The detection of underwater objects in side-looking sonar imagery is a common goal across diverse applications, from habitat mapping [1] and archaeology [2] to pipeline monitoring [3] and unexploded ordnance (UXO) remediation [4]. In this work, we address the topic of object detection in the context of underwater mine countermeasures (MCM), which is concerned with locating dangerous man-made targets on (or buried in) the seabed. However, our interest is in developing a viable general-purpose detection algorithm that is flexible enough to be used widely in various domains, not just for a single application. Thus, the driving motivation for – and contribution of – this work is the development of a new algorithm that can be adapted for detecting in sonar imagery a wide range of "interesting" objects of different sizes, whether they be mines and UXO, coral outcroppings and pipelines, or shipwrecks and other smaller man-made junk.

Present-day sonar-based underwater data collection usually employs autonomous underwater vehicles (AUVs) [5], and increasingly these platforms are imbued with the ability to adapt their survey routes based on the conditions encountered

in-mission [6]. This autonomy enables the collection of more informative data while simultaneously reducing mission times. But in order to achieve this more efficient use of resources, the onboard data processing algorithms whose results inform the decision-making for adaption must be fast. For example, adapting a survey to re-inspect a suspicious object detected *in situ* [7] is contingent on there being a detection algorithm able to execute quickly onboard with limited computational power in the first place.

At the same time, significant effort in synthetic aperture sonar (SAS) research has been devoted to developing dual-band systems capable of exploiting multiple frequency bands simultaneously [8]. These newer systems enable the creation of co-registered images comprising the usual high-frequency (HF) band sonar data, but also additional low-frequency (LF) band data that can provide sub-sediment information. Nevertheless, many single-band sonar systems remain in use, so there is a desire to develop a flexible method that can function with only a single HF sonar image, but when available, leverage data from an additional band in an auxiliary manner.

One reason a new detection algorithm ought to be compatible with both types of systems is that collecting new underwater data is expensive. Additionally, as with many remote sensing modalities [9], sonar data exhibits a strong environmental dependence, predominantly on seafloor conditions [10]. Taken together, these two factors mean that there is still a relative paucity of sonar data that sufficiently covers the variability in the universe of environmental conditions that can be encountered.

As a result, oftentimes the training data available do not necessarily match the test data on which automatic target recognition (ATR) algorithms are to be subsequently applied. That is, a data mismatch exists in which the underlying statistics of the training and test sets differ. This situation can lead to overfitting, and in general, the algorithms being brittle with unreliable performance [11]. To instead ensure that the detection algorithm is robust, we purposely want to refrain from relying on a set of dedicated training data.

This particular issue is crucial for our application and hence warrants a more concrete explanation. We are so keen to avoid needing training data because in real MCM operational settings data mismatch arises from several sources. The type or variety of clutter objects at a site can be different from what

has been observed in training exercises in "friendly" waters. Furthermore, it is also common to encounter novel, previously unseen classes of targets for which training data examples were not possessed. But another issue rarely acknowledged in the literature is that obtaining a pristine, canonical image of a target in an operational setting is rare. Sonar processing defects resulting from imperfectly compensated vehicle motion can cause distorted, poor-quality imagery [12]. Seafloor features, such as sand ripples or proximate rocks, often blend with a target's response. Moreover, the frequent presence of deployment accouterments – such as trolleys, shackles, or chains – still attached to the target can also significantly alter its signature. In summary, it is unrealistic to expect true "test" data to resemble that which is obtained under controlled conditions.

This perspective dovetails with the topic of performance estimation. Before automated detection and classification algorithms can be exploited for dangerous military MCM operations, there is a stringent need to know the expected performance of the algorithms [13]. That is, one must be able to accurately predict the performance of those algorithms in order to properly assess the risk to the fleet. Thus, any new detection algorithm intended for this application should be designed to facilitate performance estimation, for example as a function of image quality or environmental complexity.

To conclude, there is a desire for a detection algorithm that can be used for generic objects of interest spanning different sizes, avoids requiring explicit training data, is computationally fast, functions with both single-band and multi-band imagery, and enables performance estimation. However, none of the existing techniques sufficiently satisfy all of these requirements.

B. Existing Approaches

In the context of underwater MCM, several different target detection approaches have been developed over the years. The conventional stream of techniques, from seminal work like [14] up to even the recent research of [15], merge the detection and classification stages into a unified system that follows a traditional machine-learning path. The details and algorithms may vary, but the approaches are characterized by a common general framework in which feature extraction is followed by a supervised classifier. Unfortunately, perhaps due to the fact that only limited sets of sonar data were investigated, there is also a shared failure to address the vitally relevant operational issue about the potential effects of concept drift arising from training data mismatch [16]. This same drawback plagues other recent approaches relying on wavelet-based multi-scale matched subspace detectors [17] and on graph-based local neighborhoods in the training data [18], for which strong generalization is unproven.

Other research based on Markov random fields [19] avoided the need for training data in the detection stage by leveraging domain knowledge about the size and shape of targets. However, this approach was overly restrictive in assuming that target highlights would always be accompanied by a relatively well-defined (acoustic) shadow region. The method breaks down when a strong shadow is not present, which is often the

case with real sonar data, especially in complex environments or other challenging conditions.

More recent work [20] instead adopted a brain-computer interface approach incorporating human operators in conjunction with support vector machines. This too relies on training data, but also eliminates the possibility of autonomous operations with its prominent human-in-the-loop component. In addition, this and all of the other aforementioned approaches were developed for only single-band sonar imagery, with no straightforward way to extend to the multi-band case.

In contrast, another class of recently developed approaches [21], [22] have focused on exploiting multi-channel (*i.e.,* dual-band) SAS imagery via the concept of coherence. The obvious drawback to these methods is that they cannot be used with single-band imagery, which limits their applicability. Moreover, the algorithms are also very computationally demanding.

In our previous work, a fast algorithm that exploited integral images was developed for real-time detection of mines in sonar imagery [23]. This algorithm employed a cascaded architecture consisting of three stages in which subsequent stages operated on increasingly smaller regions of the imagery. The approach located shadows in the image, detected and eliminated those shadows consistent with sand ripples, and then performed highlight tests on the remaining regions under consideration. Although the algorithm proved successful for the intended purpose [7], it was not compatible with performance estimation or dual-band imagery.

To address these shortcomings, in this work a new detection algorithm largely inspired by [23] has been developed. We call the new method the "Mondrian detector" because the driving design aesthetic of the approach shares characteristics with works by the Dutch painter Piet Mondrian [24], namely the fact that an image is divided into various blocks of different sizes (*cf.* Mondrian's famous "Composition II in Red, Blue, and Yellow"). Another key commonality between the algorithm and Mondrian's art is conceptual simplicity. Given that this year marks the centennial of the art movement known as De Stijl, which Mondrian helped found, this homage seems apropos.

C. Organization of Paper

The remainder of this paper is organized as follows. Sec. II describes the newly developed detection algorithm in detail. Experimental results showing the algorithm's detection performance on several substantial sets of sonar data collected at sea are presented in Sec. III. Concluding remarks are made in Sec. IV.

II. MONDRIAN DETECTOR

A. Overview

The Mondrian detector is designed for HF SAS imagery, either single-band or multi-band, where the latter case implies that an LF image is also available. To ensure that the algorithm can function in either case, the LF image is used only in an auxiliary manner, as an additional prescreener that flags suspicious regions in the scene. The main prescreener, which exploits the HF image, is based on lacunarity. The core module
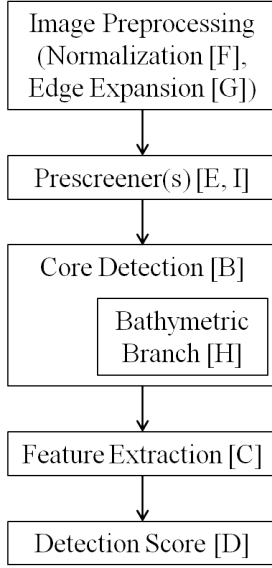
Fig. 1. The components of the Mondrian detection algorithm; the subsection within Sec. II in which each is described is also noted.

TABLE I
DETECTOR BLOCK ARRANGEMENT DETAILS

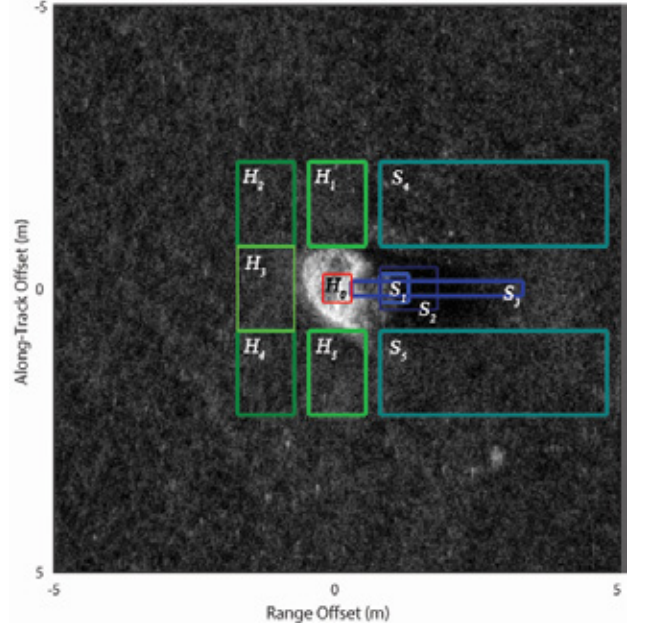| Block | Block Size (m) | | Relative Center (m) | |
Name	Along-Track	Range	Along-Track	Range
H_0	γ_1	γ_2	0	0
H_1	$3\gamma_1$	$2\gamma_2$	$-(2\gamma_1 + \gamma_3)$	0
H_2	$3\gamma_1$	$2\gamma_2$	$-(2\gamma_1 + \gamma_3)$	$-(1.5\gamma_2 + \gamma_4)$
H_3	$3\gamma_1$	$2\gamma_2$	0	$-(1.5\gamma_2 + \gamma_4)$
H_4	$3\gamma_1$	$2\gamma_2$	$2\gamma_1 + \gamma_3$	$-(1.5\gamma_2 + \gamma_4)$
H_5	$3\gamma_1$	$2\gamma_2$	$2\gamma_1 + \gamma_3$	0
S_1	γ_1	γ_2	0	$\gamma_2 + \gamma_4$
S_2	$1.5\gamma_1$	$2\gamma_2$	0	$1.5\gamma_2 + \gamma_4$
S_3	$0.5\gamma_1$	x	0	$0.5\gamma_2 + 0.5x$
S_4	$3\gamma_1$	$8\gamma_2$	$-(2\gamma_1 + \gamma_3)$	$4.5\gamma_2 + \gamma_4$
S_5	$3\gamma_1$	$8\gamma_2$	$2\gamma_1 + \gamma_3$	$4.5\gamma_2 + \gamma_4$

Fig. 2. The arrangement of blocks used in the Mondrian detector (overlaid on a target-like object from a SAS image).

of the detector consists of a set of pixel-intensity tests, six related to highlights and three related to shadows. These tests are based on the average pixel intensity in different blocks, or regions, that share a special predefined spatial arrangement in the image. Essentially the algorithm searches for areas characterized by relatively strong highlight and nearby shadow regions, with more emphasis placed on the former. The image pixel regions (or "blobs") that pass the required proportion of tests are converted into discrete alarms. A small set of features is then extracted for each alarm, from which a final detection score is computed. The design of the detector is intentionally kept simple so that analytical performance prediction would be feasible if basic scattering models – such as of the seafloor and targets – were available. A basic block diagram showing the components of the Mondrian detection algorithm appears in Fig. 1.

B. Core Algorithm

The core of the Mondrian detection algorithm exploits integral images [25] to enable the rapid calculation of pixel sums in rectangular areas of the sonar image. (It should be noted that the pixel values must be non-negative, and sufficient numerical precision must be maintained.) Specifically, eleven different blocks that share a specially defined relative arrangement are constructed. The sizes and locations of these blocks are summarized in Table I, where $\gamma = \{\gamma_1, \ldots, \gamma_5\}$ are fixed parameters that dictate the size of objects to be detected.

The algorithm is designed to detect objects that have an approximate length in $[\gamma_1, \gamma_1 + 2\gamma_3]$ in the along-track direction and a width in $[\gamma_2, \gamma_2 + 2\gamma_4]$ in the range direction. The role of $\{\gamma_3, \gamma_4\}$ is to effectively create a guard region that imbues the algorithm with the flexibility to detect objects over a range of sizes. For the objects of interest here, the settings used are $\gamma_i = 0.5$ for $i \in \{1, \ldots, 4\}$. In contrast, a pipeline detection application, for example, would likely set $\gamma_1 > \gamma_2$ and $\gamma_3 > \gamma_4$.

The block S_3 is range-dependent to reflect the geometry of the sensor-object relationship; in Table I,

$$x = \gamma_5 r/(a - \gamma_5), \qquad (1)$$

where γ_5 is the assumed nominal height of targets of interest, a is the altitude of the sonar sensor, and r is the range to the potential object.

The arrangement of blocks from Table I is also shown visually in Fig. 2.

The average pixel intensity in each of those blocks is computed when the origin of the entire system of blocks is at each position in the image. As a result, the average pixel intensity contained in a given block centered at each position in the image can be displayed as a map. Let $B(W)$ be the mean value in block W centered at some specified position. The detection algorithm then considers nine different tests based on the pixel intensity in various blocks. The nine tests $\Theta = \{\Theta_1, \ldots, \Theta_9\}$ are summarized in Table II, where ν is the (normalized) image median, and $\tau = \{\tau_1, \tau_2, \tau_3\}$ are fixed

TABLE II

DETECTOR TESTS

Test Name	Test Expression
Θ_1	$B(H_0) > (1 + \tau_1)B(H_1)$
Θ_2	$B(H_0) > (1 + \tau_1)B(H_2)$
Θ_3	$B(H_0) > (1 + \tau_1)B(H_3)$
Θ_4	$B(H_0) > (1 + \tau_1)B(H_4)$
Θ_5	$B(H_0) > (1 + \tau_1)B(H_5)$
Θ_6	$B(H_0) > \tau_2 + \nu$
Θ_7	$B(S_1) < (1 - \tau_1)B(S_4)$ $\wedge \quad B(S_1) < (1 - \tau_1)B(S_5)$
Θ_8	$B(S_3) < (1 - \tau_1)B(S_4)$ $\wedge \quad B(S_3) < (1 - \tau_1)B(S_5)$
Θ_9	$B(S_2) < \tau_3 \nu$

model parameters (*i.e.,* thresholds).

The highlight tests essentially compare the pixel intensity of block $B(H_0)$ to local regions nearby (tests Θ_1-Θ_5) and to the overall image (test Θ_6). The shadow tests compare the pixel intensity in areas expected to be in the shadow region cast by an object (centered at H_0) to local background regions nearby (tests Θ_7 and Θ_8) and to the overall image (test Θ_9).

A binary map indicating the pixel locations that passed the required number of tests, n_s, is thus constructed. The resulting map of success "blobs" is subsequently converted to discrete alarms, with the centroid of a given blob taken to be the alarm location. The size of the blob is retained as a feature. A $5(\gamma_1 + \gamma_3) \times 5(\gamma_2 + \gamma_4)$ mugshot is then extracted for each alarm, from which a set of four additional features is computed. (So here, the mugshot is 5 m × 5 m.)

C. Feature Extraction

The set of features extracted for each alarm was developed with the idea that changes in image quality and environmental complexity would affect the features in specific predictable ways, thereby permitting performance estimation. The complete list of features is given in Table III.

TABLE III

FEATURES

Feature	Description
f_1	Blob Area
f_2	Highlight Concentration
f_3	Highlight Area
f_4	Shadow Concentration
f_5	Shadow Area

The blob-area feature is a convenient summary measurement that expresses the degree (and physical extent) to which the detection tests were satisfied. The highlight concentration feature estimates the fraction of highlight pixels that are in the expected object highlight area (compared to the full mugshot), while the highlight area feature estimates the size of the object highlight. The shadow concentration and shadow size features measure analogous quantities for shadows instead of highlights. These features are computed quickly, without the need of employing a computationally intensive segmentation algorithm, as follows.

For a detected object, let $\mathbf{F}^M$ denote a Gaussian-smoothed version of its $5(\gamma_1 + \gamma_3) \times 5(\gamma_2 + \gamma_4)$ mugshot extracted from the (normalized) SAS image scene, $\mathbf{F}$. Let $\mathbf{F}^H$ and $\mathbf{F}^S$ denote sub-images extracted from $\mathbf{F}^M$ that cover what we refer to as an expected highlight region and an expected shadow region, respectively. The region $\mathbf{F}^H$ is centered about the alarm location, whereas that same position corresponds to the center of the left-most edge of the region $\mathbf{F}^S$. It should be noted that the sizes and locations of these regions are fixed, identical for all mugshots. Specifically, $\mathbf{F}^H$ is $(\gamma_1 + 2\gamma_3) \times (\gamma_2 + 2\gamma_4)$ and $\mathbf{F}^S$ is $(\gamma_1 + 2\gamma_3) \times 2(\gamma_2 + 2\gamma_4)$, meaning the region sizes are driven by the sizes of objects of interest.

Let $|H|$, $|S|$, and $|M|$ denote the numbers of pixels in the expected highlight region, the expected shadow region, and the entire mugshot, respectively. Let δ_x and δ_y denote the pixel dimensions in the image, in m.

The highlight area and shadow area are calculated as

$$h_a(k) = \delta_x \delta_y \sum_{i=1}^{|H|} \mathbb{I}(\mathbf{F}_i^H \geq \phi_h(k)) \tag{2}$$

$$s_a(k) = \delta_x \delta_y \sum_{i=1}^{|S|} \mathbb{I}(\mathbf{F}_i^S \leq \phi_s(k)), \tag{3}$$

respectively, where $\phi_h(k)$ and $\phi_s(k)$ are thresholds, $\mathbb{I}$ is an indicator function that is unity when its argument is true (and zero otherwise), and $\mathbf{F}_i$ is the value of the ith pixel in the image specified by the superscript.

Sonar imagery is significantly affected by the characteristics of the seafloor in the scene, so appropriate thresholds for determining highlights and shadows cannot be established *a priori*. Therefore, a set of three thresholds that are a function of the scene image median, $\nu(\mathbf{F})$, is used. This approach of using multiple thresholds makes the feature calculations more robust, as does incorporating local statistics (*i.e.,* the median) from the scene image.

The set of highlight thresholds is defined to be $\phi_h(k) = \phi_1(k) + \nu(\mathbf{F})$, where $\phi_1 = \{8, 10, 12\}$, indexed by k. Analogously, the set of shadow thresholds is defined to be $\phi_s(k) = \phi_2(k)\nu(\mathbf{F})$, where $\phi_2 = \{0.001, 0.01, 0.1\}$.

The final highlight area feature, f_3, is then defined as the mean of the three highlight areas, $h_a(k)$. The final shadow area feature, f_5, is defined as the mean of the three shadow areas, $s_a(k)$.

The highlight concentration and shadow concentration are calculated similarly as

$$h_c(k) = \frac{\sum_{i=1}^{|H|} \mathbb{I}(\mathbf{F}_i^H \geq \phi_h(k))}{\sum_{j=1}^{|M|} \mathbb{I}(\mathbf{F}_j^M \geq \phi_h(k))} \tag{4}$$

$$s_c(k) = \frac{\sum_{i=1}^{|S|} \mathbb{I}(\mathbf{F}_i^S \leq \phi_s(k))}{\sum_{j=1}^{|M|} \mathbb{I}(\mathbf{F}_j^M \leq \phi_s(k))}, \tag{5}$$

respectively.

The final highlight concentration feature, f_2, is defined as the mean of the three highlight concentrations, $h_c(k)$. The final shadow concentration feature, f_4, is defined as the mean of the three shadow concentrations, $s_c(k)$.

D. Detection Score

For an alarm represented by the feature vector $\boldsymbol{f} = [f_1, \ldots, f_5]^\mathsf{T}$ extracted from the HF image, the final

detection score, t, is computed as

$$t = \xi_1 f_1 + \xi_2 h + \xi_3 s, \qquad (6)$$

where

$$h = (f_2 f_3)^{1/2} \qquad (7)$$
$$s = (f_4 f_5)^{1/2} \qquad (8)$$

and $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_3\}$ are parameters. The detection score is a useful way to rank the alarms and, when feasible, inform an AUV re-inspection strategy.

It can be seen that the geometric means of (7) and (8) are readily understandable as concise summary measures of an alarm's highlight region and shadow region, respectively. Thus, the parameters $\boldsymbol{\xi}$ can be set based on the perceived relative importance of highlight and shadow information.

E. Prescreener

The main (*i.e.,* HF-based) prescreener Λ_1 exploits the concept of lacunarity [26], a measure of pixel-intensity variation. If the lacunarity $\ell(i, j)$ of a block centered at position (i, j) in the HF image is above a threshold, α_1, the pixel passes the prescreener. The block is chosen to be half of the maximum size of objects of interest, $0.5(\gamma_1 + 2\gamma_3) \times 0.5(\gamma_2 + 2\gamma_4)$, so 0.75 m $\times 0.75$ m here.

When an LF image is available, an additional prescreener Λ_2 is employed, as will be explained in Sec. II-I. Only alarms that pass at least one of the prescreener tests $\boldsymbol{\Lambda} = \{\Lambda_1, \Lambda_2\}$ can subsequently be declared valid detections.

F. Normalization

The normalization procedure employed performs median normalization in each direction, applies a common nonlinear transformation, subtracts the mean, and truncates the pixel values to the range $[0, 40]$.

More specifically, for an image $\mathbf{X}$, let $\nu(\mathbf{X}_{i,:})$ denote the median computed from all pixels of the image at the ith range position; let $\nu(\mathbf{X}_{:,j})$ denote the median computed from all pixels of the image at the jth along-track position; and let $\nu(\mathbf{X})$ and $\mu(\mathbf{X})$ denote the median and mean of the entire image, respectively.

Let $A(i, j)$ be the magnitude of the raw sonar image at position (i, j), and let $\mathbf{A}$ denote the entire image. The normalization applied to the pixel at (i, j) is then

$$B(i,j) = A(i,j)/\nu(\mathbf{A}_{i,:}) \qquad (9)$$
$$C(i,j) = B(i,j)/\nu(\mathbf{B}_{:,j}) \qquad (10)$$
$$D(i,j) = 20\log_{10} C(i,j) \qquad (11)$$
$$E(i,j) = D(i,j) - \mu(\mathbf{D}) \qquad (12)$$
$$F(i,j) = \min\{\max\{E(i,j), 0\}, 40\}. \qquad (13)$$

The image median used in the Mondrian detector (*e.g.,* in Table II) is $\nu = \nu(\mathbf{F})$, and $\mathbf{F}$ is the normalized image used for detection and feature extraction. (The exception is that the lacunarity prescreener uses the unnormalized image $\mathbf{A}$ instead of $\mathbf{F}$.)

It should be noted that additional subtleties are included in the normalization implementation to avoid using "bad" data, namely regions at long range characterized by scalloping due to platform motion and regions at short range that correspond to the water column.

G. Handling Image Edges

The original sonar image is expanded around its perimeter to ensure that the H_0 block can be formed everywhere in the original image, and in turn, that targets can be detected anywhere in the original image. The image expansion is performed by simply reflecting the image about the edge in question, by copying the relevant pixel values.

Because the image expansion is performed with only H_0 in mind, it will still not be possible to form some other blocks in certain locations of the image. When this is the case, the tests in Table II that would be affected are elided, reducing the effective number of test successes possible at particular locations. For instance, at locations along the top edge of the image, block means $B(H_1)$, $B(H_2)$, and $B(S_4)$ would not be computable, so tests $\{\Theta_1, \Theta_2, \Theta_7, \Theta_8\}$ would not be performed for those pixels. An alarm would still be generated in those regions if the fraction of tests passed (to tests performed) was above the required threshold.

In the interior of an image, all blocks can be formed and hence all $n_t = 9$ tests can be performed. Near image edges, since some blocks cannot be formed, the minimum number of tests that must be passed is made to be a function of the number of tests that cannot be conducted, n_x. The general expression that determines the minimum number of tests that must be passed, n_s, to generate an alarm is

$$n_s = \lfloor (2/3)(n_t - n_x) \rfloor, \qquad (14)$$

where $\lfloor \cdot \rfloor$ indicates the mathematical floor. As a result, (14) dictates that between $1/2$ and $2/3$ of the viable tests must be passed for a pixel to generate an alarm, depending on its location relative to image edges.

H. Bathymetric Branch

The design of the core Mondrian detection algorithm is premised on the assumption that targets will reliably exhibit strong highlights. Shadows implicitly play a lesser role in the approach because fewer shadow tests are involved than highlight tests. (In fact, this is why tests Θ_7 and Θ_8 were intentionally defined with conjunctions.) For example, it is possible to generate an alarm while passing only highlight tests, but the same is not true if only shadow tests are passed. Unfortunately, targets sometimes produce weak highlights.

In order to enable the detection of targets with weak highlights, an additional route based on bathymetric anomalies is included in the overall detection scheme. The guiding principle of this alternative component is that an object proud of the seafloor can generate a strong shadow despite having a weak highlight. Therefore, a search of the sonar image for shadows that, based on geometry, suggest the presence of a proud object of a relevant size (*i.e.,* height and length) is undertaken in what we call the bathymetric branch of the detector.
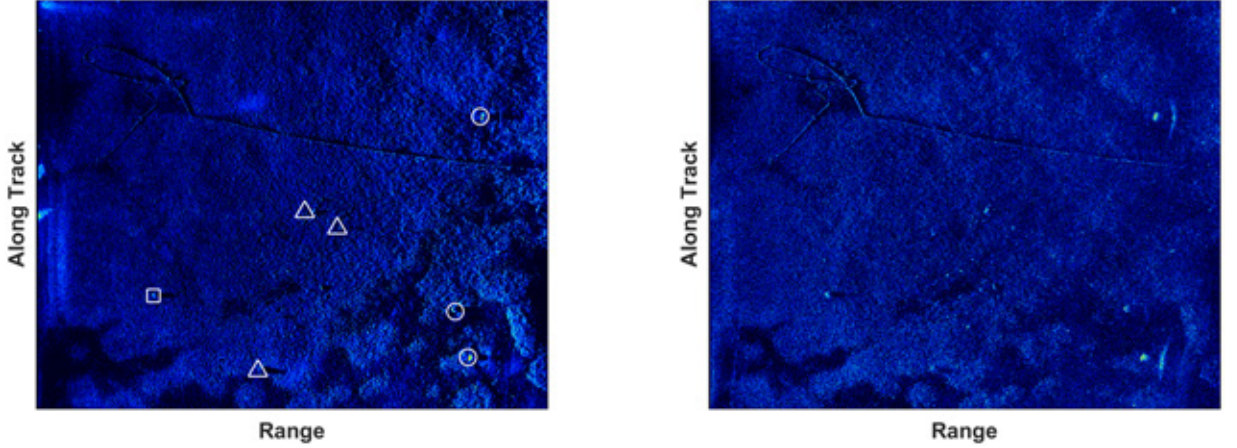
Fig. 3. For an image from a dual-frequency sonar system, detections generated by the Mondrian detector as a result of the LF prescreener (marked by circles), as a result of the HF lacunarity prescreener (marked by squares), and as a result of the (HF) bathymetric branch (marked by triangles). The left and right images are the HF and LF images, respectively, with range increasing to the right. The alarms correspond to three true targets and four false alarms.

More specifically, a map indicating the pixels that passed shadow-test Θ_8 is formed during the core detection module. Each shadow "blob" in this map is then subjected to a set of bathymetric tests $\Psi = \{\Psi_1, \ldots, \Psi_6\}$ that assess characteristics of the shadow region. The tests are shown in Table IV, where z. denotes properties of the shadow region being evaluated (obvious from the table), and $\boldsymbol{\zeta} = \{\zeta_1, \ldots, \zeta_9\}$ are fixed parameters (thresholds).

TABLE IV

BATHYMETRIC BRANCH TESTS

Test Name	Property	Test Expression
Ψ_1	Area	$\zeta_1 \leq z_a \leq \zeta_2$
Ψ_2	Orientation	$|z_o| \leq \zeta_3$
Ψ_3	Along-Track Length	$\zeta_4 \leq z_\ell \leq \zeta_5$
Ψ_4	Solidity	$z_s \geq \zeta_6$
Ψ_5	Extent	$z_e \geq \zeta_7$
Ψ_6	Estimated Object Height	$\zeta_8 \leq z_h \leq \zeta_9$

The tests reject those (shadow) blobs that are inconsistent – in terms of shape or size – with a target. (Solidity measures the proportion of pixels in the convex hull that are also in the shadow region. Extent measures the proportion of pixels in the shadow region to pixels in the corresponding bounding box.)

All of the shadow properties are straightforward to measure, except for z_h, the estimated height of the object that would have caused the shadow, which is computed as

$$z_h = a(r_f - r_i)/r_f, \tag{15}$$

where a is the altitude of the sensor, and r_i and r_f are the ranges at the start and end of the shadow, respectively.

The blobs that pass every test are subsequently converted to discrete alarms and added to those generated by the "main" branch of the detector.

I. Dual-Band Extension

When an LF image is available with the HF image, an additional prescreener Λ_2 is employed. This prescreener uses the core module of the Mondrian detector on the LF image, but

given the physics involved, considers only the six highlight-based tests. All six tests must be satisfied for a pixel to pass this prescreener. In addition, the resulting blob of success pixels must also be of an appropriate size.

This prescreener test is shown in Table V, where $\{\alpha_2, \ldots, \alpha_5\}$ are fixed model parameters, the prime notation denotes that the LF-band image is used, and f_1' is understood to be the area of the blob formed by the pixels that satisfied the earlier operands of the conjunction.

TABLE V

LF PRESCREENER TEST

Test Name	Test Expression
Λ_2	$B(H_0') > (1 + \alpha_2)B(H_k') \quad \forall k \in \{1, \ldots, 5\}$ $\wedge \quad B(H_0') > \alpha_3 + \nu'$ $\wedge \quad \alpha_4 \leq f_1' \leq \alpha_5$

As a demonstration of the Mondrian detector's full functionality, example SAS imagery from a dual-band sonar system is shown in Fig. 3. The alarms generated by each component of the Mondrian detector for this imagery, including the LF-based prescreener, are also shown. In this particular imagery, all three alarm-sources of the algorithm – lacunarity prescreener, LF-based prescreener, and bathymetric branch – generated alarms.

J. Setting Parameters

The Mondrian detection algorithm admittedly contains numerous fixed parameters: prescreener parameters $\boldsymbol{\alpha}$, detector model parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\tau}$, feature-extraction parameters $\boldsymbol{\phi}$, detection score parameters $\boldsymbol{\xi}$, and bathymetric branch parameters $\boldsymbol{\zeta}$. As noted throughout the preceding subsections, appropriate values can be determined for many of the quantities based on the geometry of objects of interest or common domain knowledge. But for completeness, here we provide insight regarding how to set reasonable values for all of the parameters.

The purpose of the lacunarity prescreener is to flag regions in the image with pixel-intensity variation exceeding that

which can be explained by the seafloor alone. The pixel intensity of the seafloor in SAS imagery has been shown to be well-modeled by a k-distribution (with some shape parameter β) [27], for which the lacunarity is given by $2+2/\beta$. As $\beta \to \infty$, a Rayleigh distribution is obtained; smaller values of β produce heavier-tailed distributions, and hence higher lacunarity. (Lacunarity can be shown to be related to scintillation index, ψ, by $\ell = \psi + 1$.) Because the appropriate value of β for a site's seafloor is usually not known *a priori*, we assume a lower bound in terms of seafloor statistics when setting the prescreener lacunarity threshold $\alpha_1 = 2$. If the seafloor statistics (*i.e.,* β) of a site are known – for example, from historical data – this information can be used to set α_1 more appropriately (namely, higher).

As noted previously, the detector model parameters $\boldsymbol{\gamma}$ dictate the size of the objects to be detected. More specifically, the algorithm is designed to detect objects that have an approximate length in $[\gamma_1, \gamma_1 + 2\gamma_3]$ in the along-track direction, a width in $[\gamma_2, \gamma_2 + 2\gamma_4]$ in the range direction, and (of lesser importance) an approximate height of γ_5.

The detector model parameters $\boldsymbol{\tau}$ are related to the expected target-to-background ratio and (target) shadow-to-background ratio involving objects to be detected and the seafloor. As a guide, the parameters should satisfy $\tau_1 \in (0,1)$, $\tau_2 > 0$, and $\tau_3 \in (0,1)$, but the particular values would be dictated by the aforementioned ratios expected for the targets of interest. When available, sonar target-strength models can inform the setting of the highlight-related parameters. Knowledge of environmental conditions that cause multipath interference, and hence reduce shadow purity, can aid in setting the shadow-related parameters appropriately. For instance, if the water depth is shallow and the sea state is calm – favorable conditions for multipath – the parameters can be set to expect a worse (*i.e.,* higher) shadow-to-background ratio.

The feature-extraction parameters $\boldsymbol{\phi}$ are used to characterize an alarm's highlight and shadow. The use of multiple thresholds for ϕ_1 and ϕ_2 is intended to reduce the risk of setting the parameters to inappropriate values. (The highlight and shadow thresholds, ϕ_h and ϕ_s, that ϕ_1 and ϕ_2 help define also incorporate the image median.) As a rule of thumb, these parameters should satisfy $\phi_1 > \tau_2$ and $\phi_2 < \tau_3$. The expected target-to-background ratio of objects of interest should inform the setting of ϕ_1. Selecting logarithmic thresholds for ϕ_2 allows different degrees of shadow purity to be assessed. If no good intuition is possessed about the values to select, one can even integrate over additional (*i.e.,* more than three) threshold values to further guard against improper settings.

The detection score parameters $\boldsymbol{\xi}$ can be set based on the perceived importance of highlight and shadow information provided by the SAS imagery. Let d_{hs} be the ratio of the contribution of the highlight information to the contribution of the shadow information, for an alarm's detection score. That is, from (6),

$$d_{hs} = \frac{\xi_2 h}{\xi_3 s}. \tag{16}$$

When the maximum values of h and s in (7) and (8), respectively, are achieved, $s_{\max} = \sqrt{2}h_{\max}$, because the region $\mathbf{F}^S$ from which the shadow area is calculated is twice as large

TABLE VI
MONDRIAN DETECTOR PARAMETER SETTINGS

Parameter	Value	Units
α_1	2	–
γ_1	0.5	m
γ_2	0.5	m
γ_3	0.5	m
γ_4	0.5	m
γ_5	0.25	m
τ_1	0.4	–
τ_2	4	–
τ_3	0.5	–
ϕ_1	{8, 10, 12}	–
ϕ_2	{0.001, 0.01, 0.1}	–
ξ_1	1	–
ξ_2	3	–
ξ_3	1	–
ζ_1	0.2	m^2
ζ_2	4	m^2
ζ_3	40	degrees
ζ_4	0.2	m
ζ_5	2	m
ζ_6	0.75	–
ζ_7	0.5	–
ζ_8	0.1	m
ζ_9	0.8	m

as the region $\mathbf{F}^H$ from which the highlight area is calculated. Thus, when $h = h_{\max}$ and $s = s_{\max}$,

$$\frac{\xi_2}{\xi_3} = \frac{d_{hs} s_{\max}}{h_{\max}} = \sqrt{2} d_{hs}. \tag{17}$$

If it is believed that the highlight information is twice as valuable as the shadow information, or $d_{hs} = 2$, then $\xi_2/\xi_3 \approx 2.83$. We have rounded this value to one significant digit, and hence set $\xi_2 = 3$ and $\xi_3 = 1$ (which implies the highlight information is actually 2.12 times more valuable than the shadow information).

A similar argument can be used to set ξ_1. Let d_{bs} be the ratio of the contribution of the blob information to the contribution of the shadow information, for an alarm's detection score. That is, from (6),

$$d_{bs} = \frac{\xi_1 f_1}{\xi_3 s}. \tag{18}$$

Simple algebra will reveal that the maximum value of s will be $s_{\max} = \sqrt{2b_{\max}}$, where

$$b_{\max} = (\gamma_1 + 2\gamma_3)(\gamma_2 + 2\gamma_4), \tag{19}$$

the size of the largest object for which detection is desired. Suppose an alarm's blob feature achieves this maximum area, so $f_1 = b_{\max}$. When the maximum value of s in (8) is also achieved, it holds that

$$\frac{\xi_1}{\xi_3} = \frac{d_{bs} s_{\max}}{b_{\max}} = \sqrt{\frac{2}{b_{\max}}} d_{bs}. \tag{20}$$

For the values of $\boldsymbol{\gamma}$ used here, this means

$$\frac{\xi_1}{\xi_3} = \frac{2\sqrt{2}}{3} d_{bs} \approx 0.94 d_{bs}. \tag{21}$$

If it is believed that the blob information – which is a summary of the extent to which an alarm passed the detection tests – has

TABLE VII

DATA SET CHARACTERISTICS

Data Set Code	Name of Sea Experiment	Survey Dates (months / year)	Survey Location	Number of Images	Dominant Seafloor Characteristic	Average Image Complexity
CAT1	Catharsis 1	3 / 2009	Palmaria, Italy	40	clay	0.105 ± 0.015
AMI1	AMiCa	5-6 / 2010	Tellaro, Italy	404	flat sand	0.074 ± 0.021
ARI1	ARISE 1	5 / 2011	Bonassola, Italy	321	sand ripples	0.126 ± 0.083
ARI2	ARISE 2	10-11 / 2012	Elba, Italy	327	posidonia	0.142 ± 0.095
ONM1	ONMEX	9 / 2016	Hyères, France	94	shells	0.055 ± 0.008
GAM1	GAMEX	3-4 / 2017	Patras, Greece	139	mud	0.049 ± 0.008

the same worth as the shadow information, or $d_{bs} = 1$, then $\xi_1/\xi_3 \approx 0.94$. We have rounded this value to one significant digit, and hence set $\xi_1 = 1$ since $\xi_3 = 1$ already (which implies the blob information is actually 1.06 times more valuable than the shadow information).

Finally, the bathymetric branch parameters ζ can be set based on the expected size of the targets of interest in conjunction with the inviolable geometry created by the side-looking sensor and seafloor, from which shadows will manifest further in range, necessarily, from a proud object.

All of the parameter settings used in the following experiments are given in Table VI. No attempt to optimize the parameters was made, and no training data from the MUSCLE sensor were used.

III. EXPERIMENTAL RESULTS

A. Data Sets

To assess the performance of the Mondrian detector for target detection in sonar imagery, we utilize a substantial amount of data collected by CMRE's MUSCLE AUV. This experimental, state-of-the-art AUV is a 21-inch diameter vehicle from Bluefin that is equipped with an HF SAS system developed by Thales. The center frequency of the (single-band) SAS is 300 kHz, and the bandwidth is 60 kHz. The system enables the formation of high-resolution sonar imagery with a theoretical along-track resolution of 2.5 cm, and a theoretical across-track resolution of 1.25 cm, usually out to a range of 150 m. (Example imagery will be shown later.)

In the experiments, measured SAS data collected at six different geographical sites in the Mediterranean Sea between 2009 and 2017 is used. At each site, man-made targets of various realistic mine shapes – cylinders, truncated cones, wedges – had been deployed before MUSCLE AUV data-collection surveys were executed. Each data set is characterized by a different dominant seafloor characteristic, in terms of either composition (*e.g.,* clay or sand) or cover (*e.g.,* posidonia or shells). These six data sets span a large proportion of the most common seabed conditions, and are particularly notable and challenging because many target opportunities are marked by poor image quality.

For an objective sense of the relative difficulty of performing target detection at each site, a measure of the average image complexity (mean $\pm$ standard deviation) across each data set's images was also computed; an explanation of this metric is given in the Appendix. A summary of the data sets is shown in Table VII, where each image corresponds to a wide-area scene spanning 5500 m^2. The number of target detection opportunities is comparable to the number of images, making the scope of this study significant.

B. Performance Comparison

To establish the relative merit of the proposed approach fairly, we compare the performance of the Mondrian detector to that of the integral-image detector from [23], since neither method (unlike others noted previously) requires training data. The experimental results on the six MUSCLE data sets is shown in Fig. 4, where the curves for the Mondrian detector are generated by varying the threshold on the detection score in (6). To demonstrate the utility of the bathymetric branch, the performance of the Mondrian detector when the bathymetric branch is deactivated is also shown. It can be seen that the Mondrian detector consistently performs better than, or comparably to, the integral-image detector. It can also be seen that the performance of the Mondrian detector roughly tracks the image complexity of the data sets (*e.g.,* higher false alarm rates for data sets ARI1 and ARI2, where the average image complexity is relatively high).

In general, the Mondrian detection algorithm is more robust than the integral-image detector across different environmental conditions. The Mondrian detector can succeed even when target shadow quality is poor or target highlights are weak (relative to the background). The Mondrian detector also performs better near image edges, where only *partial* target signatures may be present. The low detection rates for the integral-image detector in data sets ARI2 and ONM1 are due to low target-to-background ratios, caused by seafloors covered with posidonia and shells, respectively. The lower false alarm rate for the Mondrian detector (compared to the integral-image detector) in data sets CAT1 and AMI1, which are relatively benign environments, can partially be attributed to more discerning detection scores that incorporate both intensity and *size* of alarm highlight and shadow regions. To illustrate these insights more concretely, the alarms generated by the Mondrian and integral-image detectors on three markedly different MUSCLE images are shown in Figs. 5-7.
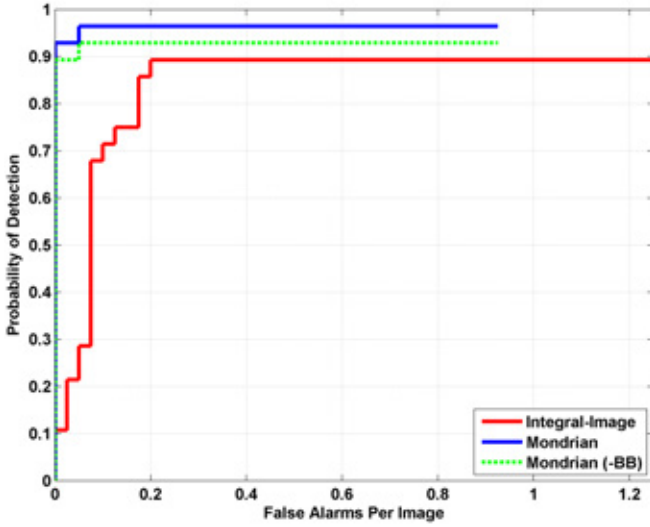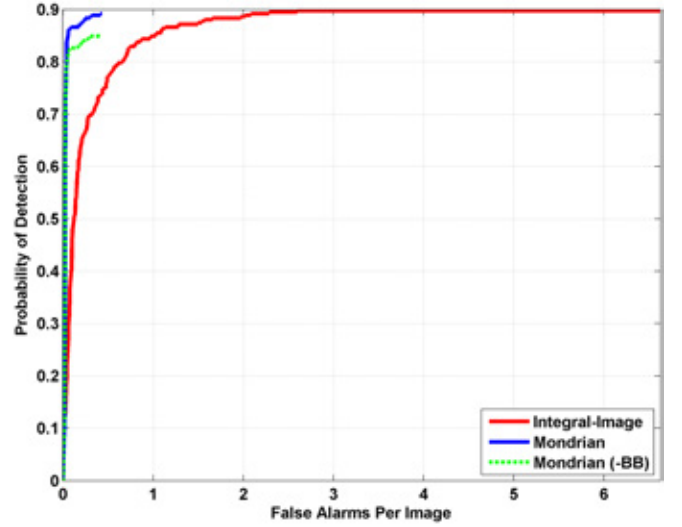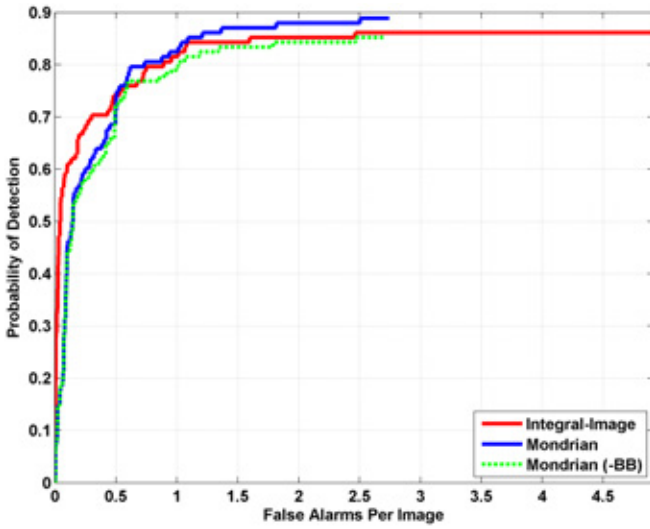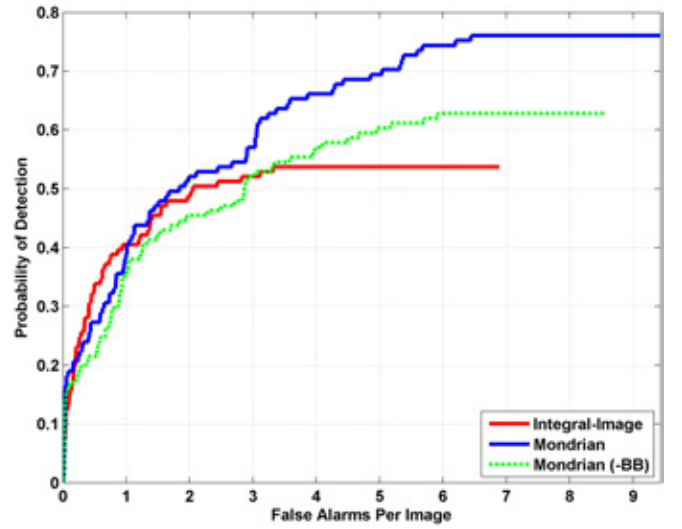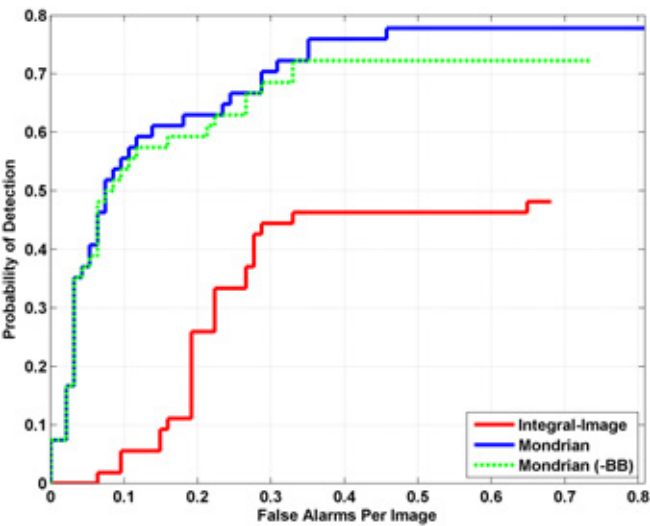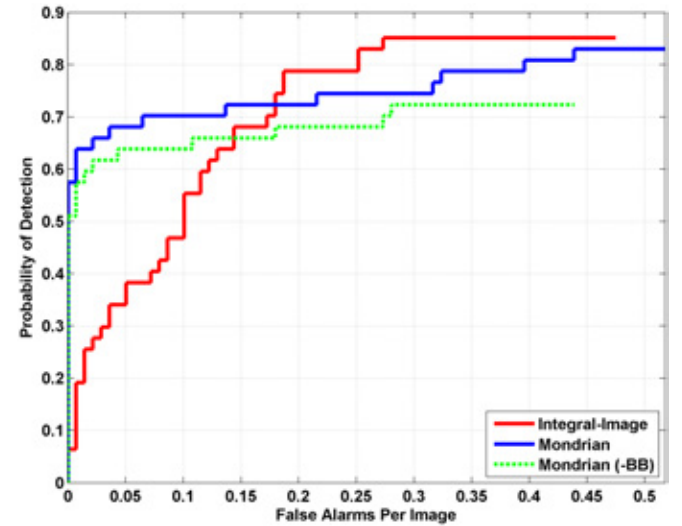
Fig. 4. Detection performance of the proposed Mondrian detector and the integral-image detector from [23] on six data sets, each with different dominant seafloor characteristics. Also shown is the performance of the Mondrian detector when the bathymetric branch is deactivated (denoted "Mondrian (-BB)").
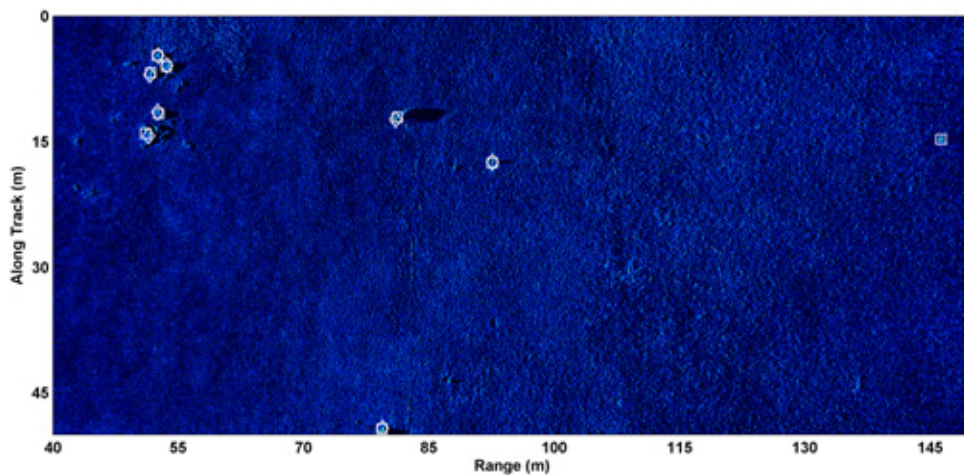
Fig. 5. Detections generated by the Mondrian detector (marked by squares) and by the integral-image detector (marked by diamonds) for a MUSCLE SAS image with sandy seafloor. There are two true targets, around 15 m along-track at ranges 80 m and 145 m, the latter of which is missed by the integral-image detector. (The image complexity value was 0.139.)
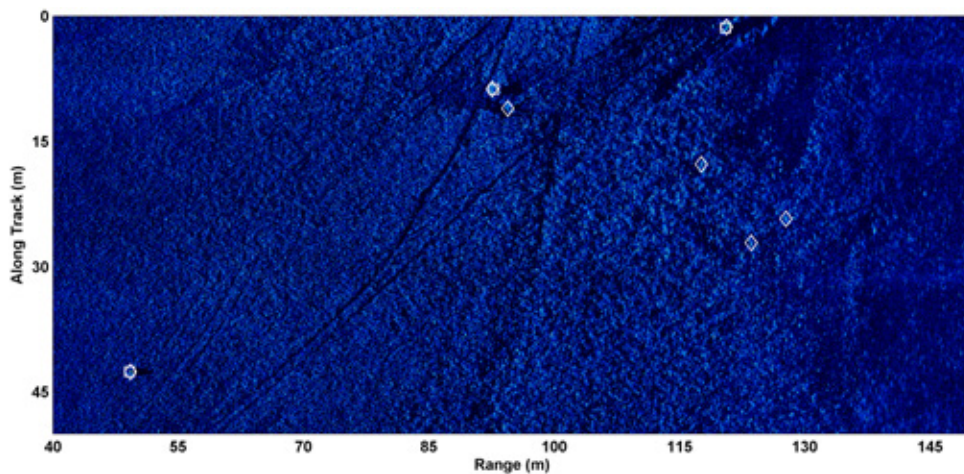

Fig. 6. Detections generated by the Mondrian detector (marked by squares) and by the integral-image detector (marked by diamonds) for a MUSCLE SAS image with posidonia-covered seafloor. Only the nearest-range alarm is a true target. The integral-image detector generates four additional false alarms compared to the Mondrian detector. (The image complexity value was 0.266.)
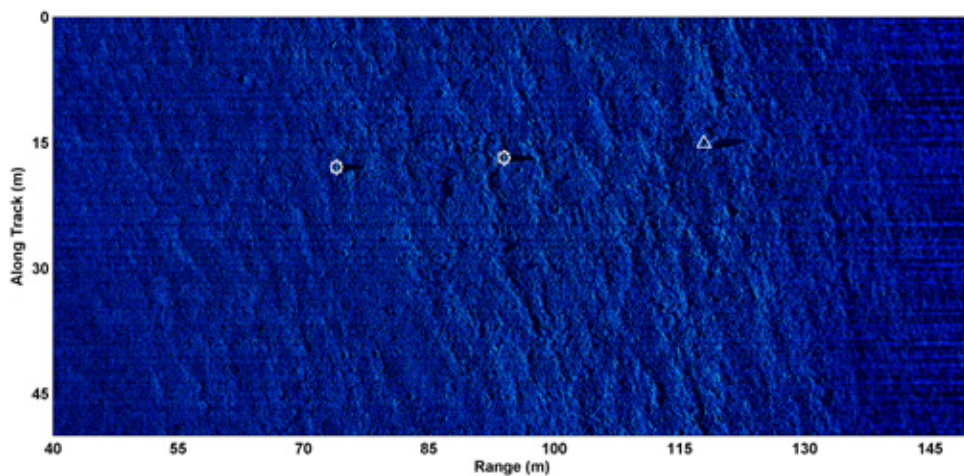

Fig. 7. Detections generated by the Mondrian detector (marked by squares and triangles) and by the integral-image detector (marked by diamonds) for a MUSCLE SAS image with shell-covered seafloor. All three alarms are true targets. The integral-image detector misses one of them. (The image complexity value was 0.073.)
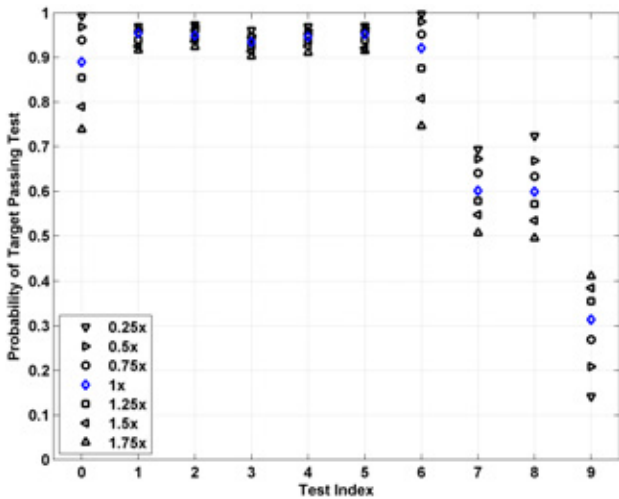
Fig. 8. Sensitivity of the prescreener test Λ_1 (index 0) and the detection tests $\{\Theta_1, \ldots, \Theta_9\}$ (indices 1-9) to the parameter values used, in terms of the fraction of targets (across all six data sets) that would pass a given test. The legend indicates the value of the parameter used, where x corresponds to the value listed in Table VI.

C. Parameter Sensitivity

To assess the sensitivity of the Mondrian algorithm parameters that most affect detection capability, we determine the fraction of target opportunities (across all six data sets) that would pass the prescreener test Λ_1 or the detection tests $\{\Theta_1, \ldots, \Theta_9\}$ for various threshold parameter settings. Specifically, we consider the cases in which the relevant parameters are set to values $c \in \{0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75\}$ times the value actually used in the experiments. The parameter in question for test Λ_1 is α_1, while for tests $\{\Theta_1, \ldots, \Theta_9\}$ they are, in order, $\{\tau_1, \tau_1, \tau_1, \tau_1, \tau_1, \tau_2, \tau_1, \tau_1, \tau_3\}$. The results of this study are shown in Fig. 8.

The variance on the prescreener test Λ_1 can be attributed to the fact that the sites have different seafloor statistics, for which the implied k-distribution model is satisfied to different degrees. For example, using a threshold of $1.75x$, or $\alpha_1 = 3.5$, assumes that the seafloors are well-modeled by a (quite heavy-tailed) k-distribution with shape parameter $\beta = 4/3$. The limited variation observed in tests $\{\Theta_1, \ldots, \Theta_5\}$, which involve highlights, indicate low sensitivity to the parameter τ_1 (over the range of values considered). Instead, there is greater sensitivity to τ_1 in tests $\{\Theta_7, \Theta_8\}$, which involve shadows, with this due to the fact that the shadow quality (*i.e.,* purity) is not consistently strong across the data sets. The variation in test Θ_6 (parameter τ_2) effectively reveals the true target-to-background ratio of the targets in the data sets, while the variation in test Θ_9 (parameter τ_3) hints at the true purity of the shadows cast by the targets in the data sets.

D. Miscellany

To demonstrate the algorithm's applicability to other side-looking sonars, a typical result showing the alarms generated by the Mondrian detector on an image collected by a REMUS AUV equipped with a Marine Sonic 900 kHz sidescan sonar is shown in Fig. 9. The same parameter settings used for
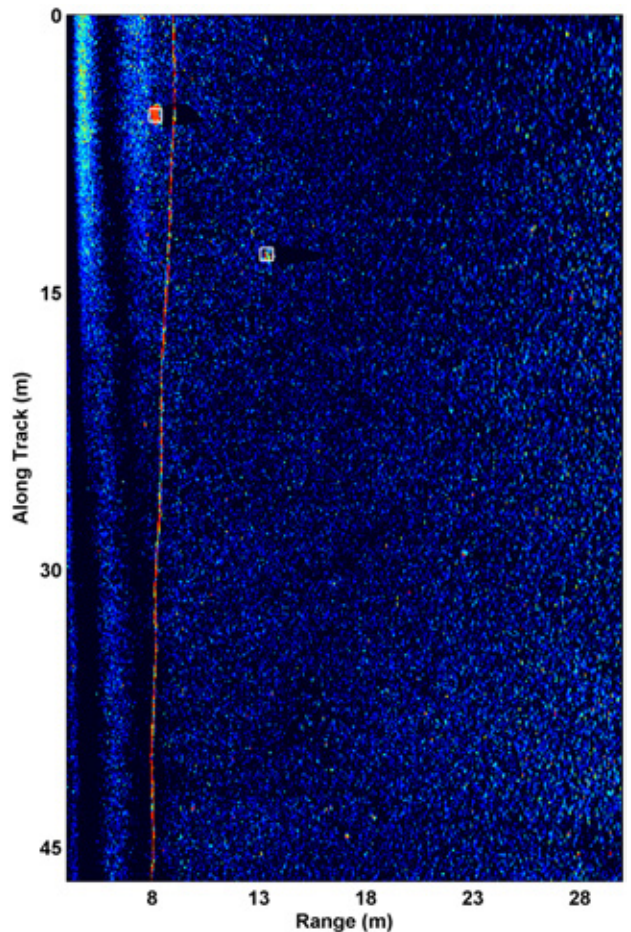

Fig. 9. Detections generated by the Mondrian detector (marked by squares) for a REMUS sidescan sonar image. Both alarms are true targets.

the MUSCLE experiments were used for the REMUS image. Despite the lower resolution of this simple sidescan sonar (as opposed to SAS), and the drastically shorter range covered by the sonar swath, the detector still functions as desired and successfully detects the targets present.

One of the advantageous aspects of the integral-image detector was its speed, which enables (delayed) real-time detection in SAS imagery onboard AUVs with limited processing power. The speed of the Mondrian detector is slightly slower than that of the integral-image detector. On the same desktop computer with a 3-GHz Intel Core2 Quad CPU and 8 GB of RAM, the average computation time (mean $\pm$ standard deviation) to execute the Mondrian detector on a MUSCLE image was 46.05 s $\pm\, 8.71$ s, while for the integral-image detector it was 26.01 s $\pm\, 11.04$ s. Nevertheless, given that the Mondrian detection algorithm was implemented in high-level MATLAB and unoptimized, onboard processing should still be feasible.

IV. CONCLUSION

A new algorithm for detecting objects in HF SAS imagery has been developed. The method can also utilize LF-band imagery when it is available. Since the algorithm is, to our knowledge, the first with this flexible single-band and dual-band functionality, it represents our principal contribution and

fills an important capability gap. A few new basic, yet robust, features were also developed in the process. The value of the proposed algorithm was demonstrated on several data sets of real SAS imagery collected at sea in diverse environmental conditions. The results showed that – as with Mondrian's art – simplicity can be powerful.

A concerted effort has been made to keep the approach as conceptually uncomplicated and general as possible in order to facilitate principled performance estimation in the future. For example, given basic models of target and background scattering, it may be possible to derive the probability of detection analytically. Toward this end, future work will pursue establishing the relationship between detection performance and environmental (image) complexity more rigorously. Dual-band sonar data sets will also be sought so that the added benefit of the second band vis-à-vis detection can be assessed experimentally.

V. Acknowledgment

Appendix

The metric used to quantify the image complexity was the mean of the spatial information [28] of the lacunarity [26] of the original sonar image, which we have found to correlate well with subjective human assessments of an image's difficulty. It is computed as follows. Let $\mathbf{H}$ and $\mathbf{V}$ represent the result of filtering an image $\mathbf{A}$ with horizontal and vertical Sobel kernels, respectively. The spatial information at each pixel can then be obtained as $\mathbf{S}(\mathbf{A}) = \sqrt{\mathbf{H}^2 + \mathbf{V}^2}$ [28], where the exponents are applied element-wise. But rather than using the original sonar image $\mathbf{A}$ as the input for calculating the spatial information, we instead use the lacunarity "map," $\mathbf{L}(\mathbf{A})$, computed from image $\mathbf{A}$ over a predefined window whose size – 0.75 m $\times$ 0.75 m here – is related to the objects of interest. The mean of this particular spatial information quantity, averaged over all pixels, is then the metric of interest: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{A})))$. This composite function, which we refer to as the "muesli complexity," much more accurately characterizes image complexity than either lacunarity alone, or spatial information alone.

References

[1] S. Williams, O. Pizarro, M. Jakuba, and N. Barrett, "AUV benthic habitat mapping in south eastern Tasmania," *Field and Service Robotics*, vol. 62, pp. 275–284, 2010.

[2] H. Singh, J. Adams, D. Mindell, and B. Foley, "Imaging underwater for archaeology," *Journal of Field Archaeology*, vol. 27, no. 3, pp. 319–328, 2000.

[3] Y. Petillot, S. Reed, and J. Bell, "Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder," in *Proc. IEEE OCEANS*, vol. 1, 2002, pp. 217–222.

[4] S. Kargl, K. Williams, T. Marston, J. Kennedy, and J. Lopes, "Acoustic response of unexploded ordnance (UXO) and cylindrical targets," in *Proc. IEEE OCEANS*, 2010, pp. 1–5.

[5] R. Marthiniussen, K. Vestgard, R. Klepaker, and N. Storkersen, "HUGIN-AUV concept and operational experiences to date," in *Proc. IEEE OCEANS*, vol. 2, 2004, pp. 846–850.

[6] F. Maurelli, P. Patrón, J. Cartwright, J. Sawas, Y. Petillot, and D. Lane, "Integrated MCM missions using heterogeneous fleets of AUVs," in *Proc. IEEE OCEANS*, 2012, pp. 1–7.

[7] D. Williams, F. Baralli, M. Micheli, and S. Vasoli, "Adaptive underwater sonar surveys in the presence of strong currents," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 2604–2611.

[8] D. Brown, D. Cook, and J. Fernandez, "Results from a small synthetic aperture sonar," in *Proc. IEEE OCEANS*, 2006, pp. 1–6.

[9] C. Ratto, P. Torrione, and L. Collins, "Exploiting ground-penetrating radar phenomenology in a context-dependent framework for landmine detection and discrimination," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 5, pp. 1689–1700, 2011.

[10] D. Williams and E. Fakiris, "Exploiting environmental information for improved underwater target classification in sonar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6284–6297, 2014.

[11] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer Science & Business Media, 2013.

[12] D. Cook and D. Brown, "Analysis of phase error effects on stripmap SAS," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 3, pp. 250–261, 2009.

[13] J. Stack, "Automation for underwater mine recognition: current trends and future strategy," in *SPIE Defense, Security, and Sensing*, vol. 8017. International Society for Optics and Photonics, 2011, pp. 1–21.

[14] G. Dobeck, J. Hyland, and L. Smedley, "Automated detection/classification of seamines in sonar imagery," in *Proc. SPIE International Society of Optics*, vol. 3079, 1997, pp. 90–110.

[15] R. Fandos, A. Zoubir, and K. Siantidis, "Unified design of a feature based ADAC system for mine hunting using synthetic aperture sonar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2413–2426, 2014.

[16] G. Widmer and M. Kubat, "Learning in the presence of concept drift and hidden contexts," *Machine Learning*, vol. 23, pp. 69–101, 1996.

[17] A. Noiboar and I. Cohen, "Anomaly detection based on wavelet domain GARCH random field modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1361–1373, 2007.

[18] G. Mishne, R. Talmon, and I. Cohen, "Graph-based supervised automatic target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2738–2754, 2015.

[19] S. Reed, Y. Petillot, and J. Bell, "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 1, pp. 90–105, 2003.

[20] C. Barngrover, A. Althoff, P. DeGuzman, and R. Kastner, "A brain–computer interface (BCI) for the detection of mine-like objects in sidescan sonar imagery," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 1, pp. 123–138, 2016.

[21] J. Tucker and M. Azimi-Sadjadi, "Coherence-based underwater target detection from multiple disparate sonar platforms," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 1, pp. 37–51, 2011.

[22] M. Azimi-Sadjadi, N. Klausner, and J. Kopacz, "Detection of underwater targets using a subspace-based method with learning," *IEEE Journal of Oceanic Engineering*, 2017, in press; DOI: 10.1109/JOE.2016.2637138.

[23] D. Williams, "Fast target detection in synthetic aperture sonar imagery: A new algorithm and large-scale performance analysis," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 1, pp. 71–92, 2015.

[24] J. Milner, *Mondrian*. Phaidon, 1995.

[25] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[26] D. Williams, "Fast unsupervised seafloor characterization in sonar imagery using lacunarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6022–6034, 2015.

[27] A. Lyons, D. Olson, and R. Hansen, "Quantifying the effect of random seafloor roughness on high-frequency synthetic aperture sonar image statistics," *Acoustic and Environmental Variability, Fluctuations, and Coherence*, vol. 38, pp. 151–158, 2016.

[28] H. Yu and S. Winkler, "Image complexity and spatial information," in *IEEE International Workshop on Quality of Multimedia Experience (QoMEX)*, 2013, pp. 12–17.