Multiframe GLRT-Based Adaptive Detection of Multipixel Targets on a Sea Surface

Marco Rodriguez-Blanco, *Member, IEEE*, and Victor Golikov, *Member, IEEE*

***Abstract*—Ship-based automatic video/infrared detection of small floating objects on a sea surface remains a challenge. We developed the GLRT-based adaptive multiframe detector for multipixel targets embedded in the channel Gaussian noise plus the background Gaussian clutter with unknown covariance matrix using a sequence of images as input data. We used video spatial-temporal patches called bricks to characterize both the target appearance and parameters estimates of the background clutter. The proposed detector is based on the return model from a sea surface in the presence of a floating object. The proposed algorithm combines the multipixel adaptive subspace detector (ASD) and adaptive multipixel background-plus-noise power change detector in a unique scheme. Experiments on simulated data and real video demonstrate the ability of the proposed detector and show that this detector considerably outperforms the known ASD, especially in a real-life situation, when the size, shape, and position of the object are unknown. We used four different types of real targets in the experiment, and the proposed detector shows better performance than the ASD, the modified mean subtraction filter and focused correlation detector. We evaluated the performance degradation in the presence of the mismatches between the actual and designed one-lag correlation coefficient of background.**

***Index Terms*—Multiframe adaptive subspace detection, hypothesis-dependent background power, sea surface.**

I. Introduction

THIS work studies integrated detection of an unknown (unknown target size, shape, and position) low contrast target in the presence of Gaussian channel noise plus background clutter using a sequence of images as input data. The images may be collected by electromagnetic sensors, such as optical sensors or infrared devices. If one or more targets are present during a scan, the corresponding image contains the returns from the targets plus the returns from the background clutter. Otherwise, if no target is present, the sensor return consists exclusively of background clutter.

The most important threats in sea-based video/infrared search are small low contrast targets. If the target is at a far distance, then its projected image is very small and low-contrast. Moreover, such images generally have a very low signal-to-background ratio (SBR), where the pixel noise contains the sea background clutter and channel noise and, in practice, these targets have unknown size, shape, and position. Conventional small target detection methods, such as mean subtraction filter [1], max-mean/max-median filter [2], least mean square filter [3], several spatial and temporal-based methods [4], and morphology-based methods [5], are used to reduce only the background clutters. Although those methods can somewhat reduce the background clutter, they do not consider the problem of selecting filters that are suitable for the targets. Application of such methods to problems of the target detection shows limited detection performance enhancement since these methods use fixed filters. Recently, Kim and Lee [6] have improved the mean subtraction filter by inserting a target enhancement and noise reduction filter called modified mean subtraction filter (M-MSF). In [7], Kadyrov *et al*. develop the focused correlation (FC) method. This method is based on the observation that two images of the background (for example, sea surface) do not correlate, whereas the two images of the same floating object do correlate.

Another kind of the target detection algorithms is based on a statistical hypothesis test that is used in the case of a heavy dynamic background environment [8]. The well-known matched and matched subspace filters [8] are such algorithms. In [9], Pohlig introduces the detection algorithm using a three-dimensional (3-D) statistical generalized likelihood ratio test (GLRT) based on batch processing: all available frames are stacked in a data volume, and then the GLRT decides on the presence or absence of a target anywhere in that volume. In practice, the assumption of the *a priori* knowledge of the background covariance matrixes is often not valid. In the literature, two different cases have been studied depending on whether the background clutter is supposed to be known or not. Several papers have addressed adaptive detecting schemes, such as the detection algorithm [10], adaptive GLRT-based algorithm [11], and adaptive subspace detector (ASD) [12]. Key assumption in [10] is that the background covariance matrix is the same under the two statistical hypotheses of the presence (H_1) or absence (H_0) of the target. In [12], Kraut *et al*. assume that the secondary data (free of target components) can be different from the primary data. It was shown [12] that ASD is a GLRT-based detector with a scaling parameter of the primary background data. Usually, the detector estimates the local mean of background and then subtracts this value from a signal of each pixel. In this case, if local mean of a background equals to the mean of a signal from object, then the object has a zero global contrast. The performances of these algorithms depend on signal-to-background-plus-noise ratio, and, therefore, these algorithms exhibit one drawback: noncontrasting targets are not detected.

In optical detection, the received background power depends on the target presence. This phenomenon is used in the recently proposed detectors [13], [14] that use different pixel background variances under two statistical hypotheses. The modified ASD

Manuscript received March 6, 2016; revised May 24, 2016 and June 14, 2016; accepted June 16, 2016. Date of publication July 10, 2016; date of current version November 30, 2016. *(Corresponding author: Victor Golikov)*

The authors are with the Engineering Faculty, Autonomous Carmen University, Ciudad del Carmen 24115, Mexico (e-mail: mrodriguez@pampano.unacar.mx; vgolikov@pampano.unacar.mx).

Digital Object Identifier 10.1109/JSTARS.2016.2582383

1939-1404 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

of the subpixel targets in a Gaussian environment was designed in [13]. It was assumed that the background covariance matrix could be estimated by using secondary dataset and, hence, the background power under the null hypothesis is *a priori* knowledge (estimated); but under the alternative hypothesis, there cannot be *a priori* knowledge of the background power. The authors of [13] considered the detection problem mathematically, as the one-dimensional adaptive subpixel detection problem, by analyzing only one pixel-vector obtained from a number of images. In [14], the detection problem is the 3-D nonadaptive multipixel problem that analyzes a number of vector-pixels received from consecutive images. In [14], *a priori* knowledge of the background covariance matrix is assumed, and the secondary dataset is not considered. In the multipixel detection, the total observed area is divided into spatial-temporal patches, and the detector analyzes each patch to detect the target presence in the patch. Each pixel of the observed patch under H_1 contains the target signal plus channel noise or background clutter plus channel noise.

In contrast to these articles [13], [14], in our paper, we consider an adaptive 3-D multipixel detection problem in the presence of normal background with unknown covariance matrix. The basic premises of used linear mixture modeling are that within a given sequence of pixels: 1) the background surface consists of a number of the oscillating elements (multipoint oscillating surface); 2) the target has a solid surface and, hence, oscillates as a one element; and 2) the background spectrum consists of more components than the target spectrum. We offer a different treatment of the GLRT, where the secondary data is combined with primary, but the secondary data is used only to estimate the background-plus-channel-noise variance and covariance under the null hypothesis. We use the primary data to estimate the background-plus-channel-noise variance under H_1 for each pixel within the observed patch. We assume that the width of the background Fourier spectrum is more than the target spectrum width and, hence, it is possible to estimate the background-plus-channel-noise variance for each pixel of the patch under H_1. The derived adaptive detector has to combine the adaptive energy detector (AED) structure and the background power change detector in a unique scheme. Both phenomena contribute to detection quality and, hence, the proposed detector considerably outperforms the well-known and recently proposed detectors that are based on the background suppressions and the assessment of the SBR. In this work, we compare the performances of the proposed detector with recently proposed M-MSF, FC detector [6], [7] and well-known ASD in the presence of both, different size of the secondary dataset and the mismatch between the designed and actual correlation coefficients of the background.

II. Problem Formulation and Detector Designs

The goal is to design and assess the adaptive scheme to detect unknown, relatively small low contrast targets in a given sequence of images with a random homogeneous background. We use video spatial-temporal patches, called bricks, to determine whether the observed brick contains a target in the presence of the random background. We assume that some or all pixel areas in the analyzed brick can be covered by the target under hypothesis H_1. These pixels contain the target signal plus channel noise, and the remaining pixels contain the background clutter plus channel noise. The variance of each pixel of the brick under H_1 is equal to the background variance plus channel noise variance or only channel noise variance. We focus attention on a situation when the presence of the target in the pixel changes the pixel noise covariance matrix. In this case, the pixel noise is equal to the channel Gaussian white noise. It is assumed that the optical background clutter is the predominant noise factor and is modeled as a correlated Gaussian noise. For simplicity, all the image data under H_0 is assumed to be a zero-mean Gaussian random vector. Specifically, the local mean of each pixel has been subtracted from the original pixel value. The whole dataset is divided into primary and secondary datasets. The primary dataset (all pixels in the observed brick) consists of a real pixel-vectors $\boldsymbol{x}_{\mathrm{j}}$, $N \times 1$, $j = 1, \ldots, L$. The secondary dataset consists of G bricks (*GL* pixel-vectors or *GLN* pixels), does not contain any targets, and exhibits the same structure of the background covariance structure. Under H_0, i.e., the background-plus-channel-noise hypothesis, we have $\boldsymbol{x}_j = \sigma_{0,j}\boldsymbol{n}_j$, where the noise pixel-vector $\boldsymbol{n}_j \sim N[\boldsymbol{0}, \boldsymbol{R}]$, $\boldsymbol{R}$ is the normalized covariance matrix of the background clutter plus channel noise for the homogeneous environment, and $\sigma_{0,j}$ is the variance of the background plus channel noise. Under H_1, i.e., the signal-plus-background-plus-channel-noise hypothesis, we have $\boldsymbol{x}_j = \boldsymbol{s}_j + \sigma_{1,j}\boldsymbol{n}_j$, where $\boldsymbol{s}_j = \boldsymbol{H}\theta_j$ is the deterministic signal of interest that belongs to a known subspace $\langle \boldsymbol{H} \rangle$ of size p [6], $\boldsymbol{H}$ is an orthogonal $N \times p$ target mode matrix, p is the rank of a projection matrix [6] $\boldsymbol{P_s} = \boldsymbol{H}\left(\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H, \boldsymbol{\theta}_j$ is the unknown amplitude vector, and $j = 1, \ldots, U$. The noise pixel-vector of the primary data under H_1, $\sigma_{1,j}\boldsymbol{n}_j$ is scaled by $\sigma_{1,j}$, where we have $\sigma_{1,j} = \sigma_{0,j}$ for the pixel-vectors without target $(\boldsymbol{n}_j \sim N[\boldsymbol{0}, \boldsymbol{R}])$; and $\sigma_{1,j} = \sigma_{ch,j}$ for the pixel-vectors with target $(\boldsymbol{n}_j \sim N[\boldsymbol{0}, \boldsymbol{I}])$, where $\sigma_{ch,j}$ is the standard deviation of the channel noise and $\boldsymbol{I}$ is the identity matrix.

We consider the problem of detecting a multipixel optical target in the sequence of N digital images with a random homogeneous Gaussian background and channel noise. We assume that the multipixel target may be present completely or partially anywhere in the brick of size L of the pixel-vectors. We assume that the area of the observed brick is partially covered by the target, and the presence of the target changes the background-plus-noise power in the pixels covered by the target. Let $j = 1, \ldots, U$ be the subset of integers indexing the pixel-vectors (U is unknown), which may contain an unknown object under the H_1 hypothesis and $j = U{+}1, \ldots, L$ be the subset of the pixel-vectors, which do not contain the object under the H_1. The secondary dataset consists of *GL* pixel-vectors $\boldsymbol{x}_j$, $N \times 1$, $j = L+1, 2, \ldots, (G+1)L$, which are assumed to have the background clutter and white channel noise components only. In this case, we assume that $\boldsymbol{x}_j \sim N[\boldsymbol{0}, \sigma^2\boldsymbol{R}]$ are independent and identically distributed Gaussian vectors and that $\boldsymbol{x}_{\mathrm{j}}$ are independent of the primary data $\boldsymbol{x}_j$. Note that we

assume the homogeneous Gaussian environment and assume that the secondary dataset is available to estimate both the normalized covariance matrix $\boldsymbol{R}$ and $\sigma^2 = \sigma^2_{0,j}$.

We develop a hypothesis test that distinguishes the signal-plus-noise hypothesis (H_1) from the background-plus-noise hypothesis (H_0). Then, the hypotheses are expressed as

$$\begin{cases} H_0 : \boldsymbol{x}_j = \boldsymbol{c}_j + \boldsymbol{n}_j, j = 1, 2, \ldots, (G+1)L \\ H_1 : \begin{cases} \boldsymbol{x}_j = \boldsymbol{s}_j + \boldsymbol{n}_j, j = 1, \ldots, U \\ \boldsymbol{x}_j = \boldsymbol{c}_j + \boldsymbol{n}_j, j = U+1, \ldots, (G+1)L. \end{cases} \end{cases} \tag{1}$$

Our approach may be evaluated against: 1) the well-known statistic GLRT of Kelly [11] that assumes the same noise scaling in both primary and secondary data ($\sigma = \sigma_0 = \sigma_1 = 1$), 2) the ASD ($\sigma = \sigma_0 = \sigma_1 \neq 1$) [12], and 3) the constant false alarm ASD [12] that assumes $\sigma \neq \sigma_0 = \sigma_1$. The statistic test ACE assumes that the variance σ_0^2 is estimated using the primary data. In contrast to these papers, we assume that $\sigma_{0,j}$ may be estimated using the secondary dataset, but $\sigma_{1,j}s$ may be estimated by the primary data and may be different with respect to $\sigma_{0,j}$. It is well known that GLRT is obtained by inserting maximum-likelihood (ML) estimates for unknown parameters into the likelihood ratio. Following the generally accepted approach, we derive the GLRT by considering the joint probability density function (pdf) of the measurement and the secondary dataset. This paragraph is devoted to the derivation of the two-step detector. In particular, we first derive the GLRT based on primary data, assuming that the variance ($\sigma^2_{0,j} = \sigma^2$) and normalized covariance matrix ($\boldsymbol{R}$) of the background plus channel noise are known. Fully adaptive detector is obtained by substituting the unknown matrix by the sample covariance matrix based on secondary data only.

Step1): Under the hypotheses H_1 and H_0, pdf of the $(G+1)L$ vectors may be written as

$$\begin{aligned} &p_{1,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}|\boldsymbol{R},\sigma^2_{1,j},\boldsymbol{\theta}_j,U,H_1\right) \\ &= \frac{(2\pi)^{-\frac{N(G+1)L}{2}}}{\|\boldsymbol{I}\|^{\frac{U}{2}}\|\boldsymbol{R}\|^{\frac{[(G+1)L-U]}{2}}\left[\prod_{j=1}^{U}\sigma^2_{ch,j}\prod_{j=1+U}^{N(G+1)L}\sigma^2_{1,j}\right]^{\frac{N}{2}}} \\ &\quad\times\exp\left[-\frac{1}{2}tr(\boldsymbol{T}_1+\boldsymbol{R}^{-1}\boldsymbol{T}_2)\right] \end{aligned} \tag{2}$$

$$\begin{aligned} &p_{0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}|\boldsymbol{R},\sigma^2_{0,j},H_0\right) \\ &= \frac{(2\pi)^{-N(G+1)L/2}}{\|\boldsymbol{R}\|^{(G+1)L/2}\left[\prod_{j=1}^{N(G+1)L}\sigma^2_{0,j}\right]^{N/2}}\exp\left[-\frac{1}{2}tr(\boldsymbol{R}^{-1}\boldsymbol{T}_0)\right] \end{aligned} \tag{3}$$

where $\|\|$ and *tr*() denote the determinant and the trace of a square matrix, respectively

$$\boldsymbol{T}_1 = \sum_{j=1}^{U}\sigma^{-2}_{ch,j}(\boldsymbol{x}_j - \boldsymbol{H}\theta_j)(\boldsymbol{x}_j - \boldsymbol{H}\theta_j)^T \tag{4}$$

$$\boldsymbol{T}_2 = \sum_{j=U+1}^{L}\sigma^{-2}_{1,j}\boldsymbol{x}_j\boldsymbol{x}_j^T + \sum\nolimits_{j=L+1}^{(G+1)L}\sigma^{-2}\boldsymbol{x}_j\boldsymbol{x}_j \tag{5}$$

$$\boldsymbol{T}_0 = \sum_{j=1}^{L}\sigma^{-2}_{0,j}\boldsymbol{x}_j\boldsymbol{x}_j^T + \sum_{j=L+1}^{(G+1)L}\sigma^{-2}\boldsymbol{x}_j\boldsymbol{x}_j^T. \tag{6}$$

The GLR statistic for the problem at hand is

$$\mathrm{L} = \frac{\max\limits_{\sigma^2_{1,j},\boldsymbol{\theta}_j,U} p_{1,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}|\boldsymbol{R},\sigma^2_{1,j},\boldsymbol{\theta}_j,U,H_1\right)}{p_{0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{(G+1)L}|\boldsymbol{R},\sigma^2_{0,j},H_0\right)} \tag{7}$$

where the numerator is maximized by independent varying $\sigma^2_{1,j}$, $\boldsymbol{\theta}_j$, and U. Using the primary dataset, we obtain two MLE of the $\sigma^2_{1,j}$: the first estimate for the subset $j = 1, \ldots, U$, where the possible object is present (in his case the pixel noise is equal to channel noise), and the second estimate for the subset $j = U+1, \ldots, L$, where the possible object is not present (in this case the pixel noise is equal to the background-plus-channel noise). Then, the MLE of the $\sigma^2_{1,j}$ has two solutions [8]:

$$\hat{\sigma}^2_{1,j} = \hat{\sigma}^2_{ch,j} = \frac{\boldsymbol{x}_j^T\boldsymbol{P}_s^{\perp}\boldsymbol{x}_j}{N-p} \text{ for } j = 1, \ldots, U, \tag{8}$$

and

$$\hat{\sigma}^2_{1,j} = \frac{\boldsymbol{x}_j^T\boldsymbol{x_j}}{N} \text{ for } j = U+1, \ldots, L \tag{9}$$

where p is the rank of the target subspace H, $\boldsymbol{P}_s^{\perp} = \boldsymbol{I} - \boldsymbol{P}_s$ is $N \times N$ orthogonal projection matrix onto the subspace orthogonal to the signal subspace $\langle H\rangle$, and $\boldsymbol{P}_s$ is $N \times N$ orthogonal projection matrix onto the signal subspace H. The MLE of the vector $\boldsymbol{\theta}_j$ has an explicit solution [8], when the target is present ($j = 1, .., U$):

$$\hat{\boldsymbol{\theta}}_j = \left(\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{x}_j. \tag{10}$$

Substitutions of $\boldsymbol{\theta}_j$ (10) and $\hat{\sigma}^2_{1,j}$ (8) into (4) yield

$$tr(\boldsymbol{T}_1) = \mathrm{tr}\left(\sum_{j=1}^{U}\frac{\boldsymbol{x}_j\boldsymbol{P}_s^{\perp}\boldsymbol{x}_j^T}{\hat{\sigma}^2_{ch,j}(N-p)}\right) = NU \tag{11}$$

and $\hat{\sigma}^2_{1,j}$ (9) into (5) yield

$$\begin{aligned} &\mathrm{tr}\left(\boldsymbol{R}^{-1}\boldsymbol{T}_2\right) = \\ &\mathrm{tr}\left[\boldsymbol{R}^{-1}\left(N\sum_{j=U+1}^{L}\frac{\boldsymbol{x}_j\boldsymbol{x}_j^T}{\boldsymbol{x}_j^T\boldsymbol{x}_j} + \sum_{j=L+1}^{(G+1)L}\sigma^{-2}\boldsymbol{x}_j\boldsymbol{x}_j^T\right)\right] \end{aligned} \tag{12}$$

.

Maximizing the numerator over $\boldsymbol{\theta}_j$ and $\sigma^2_{1,j}$, the GLR statistic can be recast

We use a straightforward maximization in (13) shown at the bottom of the next page. Since $L(G+1) > pN$, it follows immediately that the maximum of L with respect to U is obtained by replacing the unknown parameter U by known L.

Then, the GLR statistic can be rewritten as (14) shown at the bottom of the next page.

Step 2): In the case of the homogeneous environment, we can make the detector fully adaptive by plugging the ML estimate of $\boldsymbol{R}$ and σ^2 based on the secondary data $\boldsymbol{x}_{\mathrm{j}}$

$$j = L+1, \ldots, (G+1)L, \text{i.e.,}$$
$$\hat{\sigma}^2_{0,j} = \hat{\sigma}^2 = \frac{1}{NGL} \sum_{j=L+1}^{(G+1)L} \boldsymbol{x}_j^T \boldsymbol{x}_j,\ \hat{\boldsymbol{R}} = \frac{\frac{1}{NGL}\sum_{j=L+1}^{(G+1)L} \boldsymbol{x}_j \boldsymbol{x}_j^T}{\hat{\sigma}^2}. \tag{15}$$

Following this guidance, we receive

$$L = \frac{\|\hat{\boldsymbol{R}}\|^{L/2} \exp\left[\frac{N}{2\hat{\sigma}^2}\left(\sum_{j=1}^{L} \boldsymbol{x}_j^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{x}_j - L\right)\right]}{\left(\prod_{j=1}^{L} \frac{\boldsymbol{x}_j^T \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j}{\hat{\sigma}^2 (N-p)}\right)^{N/2}}. \tag{16}$$

Taking the logarithm of the *N/2*th root of *L*, the proposed GLRT-based detector, named multipixel ASD (MASD), is given by the following decision rule:

$$\sum_{j=1}^{L} \left(\frac{1}{\hat{\sigma}^2} \boldsymbol{x}_j^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{x}_j - ln \frac{\boldsymbol{x}_j^T \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j}{(N-p)\,\hat{\sigma}^2} \right) \underset{<_{H_0}}{>^{H_1}} t \tag{17}$$

where *t* is the threshold. One can see from (17) that the detection quality depends on the relation between the object energy contribution (the first term) and the background power change contribution (the second term). Obviously, the first term contribution may be insufficient to detect the low-contrast object. Note that the second term does not depend on the object contrast and its return energy. The contribution of the second term depends on relation between the channel noise variance estimate $\hat{\sigma}^2_{ch,j} = \frac{\boldsymbol{x}_j \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j^T}{(N-p)}$ in the each pixel-vector and background-plus-channel-noise variance $\hat{\sigma}^2$. The second term contribution achieves the high level when $\hat{\sigma}^2_{ch,j} \ll \hat{\sigma}^2$, number of pixels occupied by object and number of frames within the brick are large. We can adjust the background power change sensitivity of detector with respect to the target presence sensitivity using a factor *m*. Varying the factor *m,* we can adapt the background power change to the detection performance. We can introduce the factor of signal detection sensitivity *m* and rewrite the decision rule in the following form:

$$T_{\mathrm{MASD}} = \sum_{j=1}^{L} \left(\frac{m}{\hat{\sigma}^2} \boldsymbol{x}_j^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{x}_j - ln \frac{\boldsymbol{x}_j^T \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j}{(N-p)\,\hat{\sigma}^2} \right) \underset{<_{H_0}}{>^{H_1}} t. \tag{18}$$

Note that the first term $T_{\mathrm{AED}} = \sum_{j=1}^{L} \boldsymbol{x}_j^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{x}_j$ is the known classical AED, and the second term $T_{\mathrm{ACHD}} = \sum_{j=1}^{L} ln \frac{\boldsymbol{x}_j^T \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j}{(N-p)\hat{\sigma}^2}$ is the adaptive background power change detector (ACHD) that depends on the data projection on the orthogonal subspace. In the following section, we evaluate performance of the MASD and assess it against the two recently proposed M-MSF and FC detectors and classical ASD [12]:

$$T_{\mathrm{ASD}} = \sum_{j=1}^{L} \boldsymbol{x}_j^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{H} \left(\boldsymbol{H}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{H} \right)^{-1} \boldsymbol{H}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{x}_j \underset{<_{H_0}}{>^{H_1}} t_1. \tag{19}$$

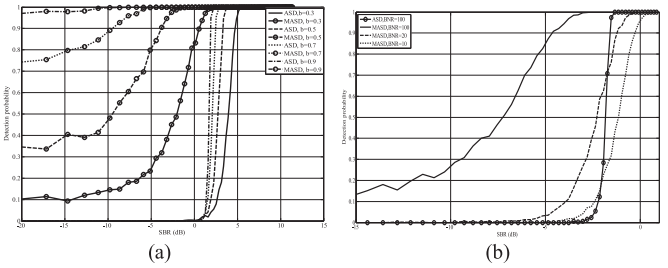
III. Detection Performance and Numerical Illustrations

In the following, we assess the performance of the proposed detector and compare it to the ASD by utilizing the standard Monte Carlo counting techniques. Specifically, in order to evaluate the threshold needed to ensure a preassigned value of the false alarm probability (P_F) and the probability of detection (P_D), we resort to $100/P_{\mathrm{F}}$ and 10^3 independent trials, respectively. We model the background and channel noise as Gaussian process. We assume that the background normalized covariance matrix is exponentially shaped, i.e., $\boldsymbol{R} = [r_{i,k}] = \left[\rho^{|i-k|}\right]$, where ρ is the one-lag correlation coefficient. We assume that the optical observed brick is partially covered by the target, and the presence of the target in any pixel reduces the total noise power to the channel noise power. We introduce the target fill factor $b = U/L$ and the mismatched factor $(MF = \rho_{\mathrm{ac}} - \rho_d)$ of the actual ρ_{ac}; design ρ_d one-lag background correlation coefficient; and then analyze, as the detection performance depends on *b* and *MF*. The primary data is modeled as the deterministic target signals with the uniform amplitude vector $\boldsymbol{\theta} = [1, 1, \ldots, 1]^T$ and with *p* low-frequency independent basis modes of the matrix $\boldsymbol{H}$. This matrix is a Vandermonde matrix with discrete complex exponential elements. The secondary data model is a sequence of the *G* subimages (frames), where each frame is represented as the *L*, independent, identically distributed correlated Gaussian vectors. We investigate the detectability of the detector at different parameters: the target fill factor $b = U/L$, the background-to-noise ratio (BNR), the SBR, brick size *NL*, secondary data size *GL*, the mismatched factor *MF*, and the sensitivity factor *m*. The MASD performance

$$\mathrm{L} = \max_{U} \frac{\|\boldsymbol{R}\|^{U/2} \left(\prod_{j=1}^{U} \sigma^2_{0,j}\right)^{N/2} \exp\left\{tr\left[\frac{\boldsymbol{R}^{-1}}{2}\left(\sum_{j=1}^{L} \sigma^{-2}_{0,j} \boldsymbol{x}_j \boldsymbol{x}_j^T + \sum_{j=L+1}^{(G+1)L} \sigma^{-2} \boldsymbol{x}_j \boldsymbol{x}_j^T\right)\right]\right\}}{\left(\prod_{j=1}^{U} \frac{\boldsymbol{x}_j^T \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j}{N-p}\right)^{N/2} \exp\left[tr\left(\boldsymbol{R}^{-1} \frac{N}{2} \sum_{j=U+1}^{L} \frac{\boldsymbol{x}_j \boldsymbol{x}_j^T}{\boldsymbol{x}_j^T \boldsymbol{x}_j} + \frac{\boldsymbol{R}^{-1}}{2} \sum_{j=L+1}^{(G+1)L} \sigma^{-2} \boldsymbol{x}_j \boldsymbol{x}_j^T\right) + \frac{NU}{2}\right]}. \tag{13}$$

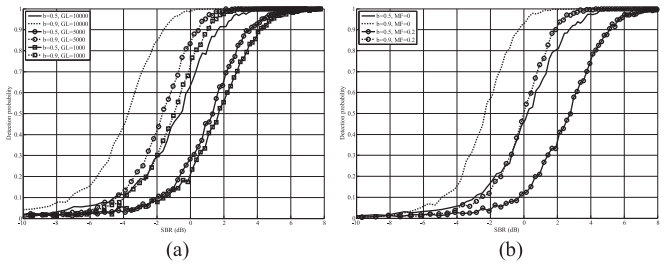
$$L = \frac{\|\boldsymbol{R}\|^{L/2} \left(\prod_{j=1}^{L} \sigma^2_{0,j}\right)^{N/2} \exp\left\{tr\left[\frac{\boldsymbol{R}^{-1}}{2}\left(\sum_{j=1}^{L} \sigma^{-2}_{0,j} \boldsymbol{x}_j \boldsymbol{x}_j^T + \sum_{j=L+1}^{(G+1)L} \sigma^{-2}_{0,j} \boldsymbol{x}_j \boldsymbol{x}_j^T\right)\right]\right\}}{\left(\prod_{j=1}^{L} \frac{\boldsymbol{x}_j^T \boldsymbol{P}_s^{\perp} \boldsymbol{x}_j}{N-p}\right)^{N/2} \exp\left[tr\left(\frac{\boldsymbol{R}^{-1}}{2} \sum_{j=L+1}^{(G+1)L} \sigma^{-2} \boldsymbol{x}_j \boldsymbol{x}_j^T\right) + \frac{NL}{2}\right]} \tag{14}$$

.



(a) (b)

Fig. 1. Probability of detection versus SBR for MASD and ASD: (a) at different *b,* BNR = 100 and *GL* = 5000; (b) at different BNR, $b = 0.9$, $GL = 1000$. Simulation results for $N = 10,\ L = 10,\ m = 1,\ p = 3,\ \rho_{\rm ac} = \rho_d = 0.8$, $MF = 0$, $P_F = 10^{-3}$.



(a) (b)

Fig. 2. Probability of detection versus SBR for MASD: (a) at different *GL* and *b*, $\rho_{\rm ac} = \rho_d = 0.75$, $MF = 0$; (b) at different *MF* and *b*, *GL* = 10 000. Simulation results for $N = 10$, $L = 10$, $m = 1$, $p = 3$, $P_F = 10^{-3}$, BNR = 10, $P_F = 10^{-3}$.

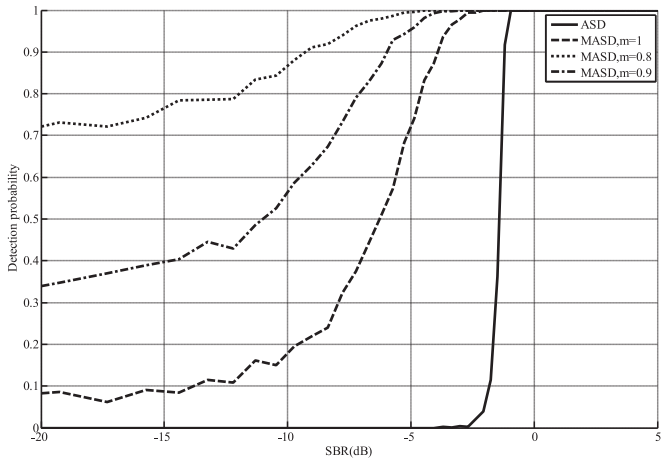
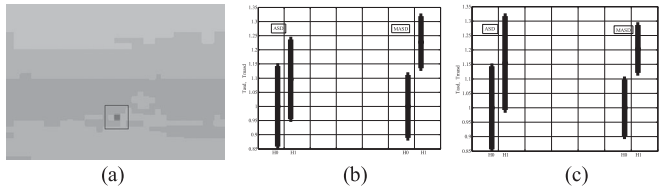


Fig. 3. Probability of detection versus SBR for MASD and ASD at different sensitive factors *m*, $MF = 0, BNR = 100, L = 10, b = 0.9, N = 10, p = 3$, $GL = 10000$, and $P_F = 10^{-3}$. Simulation results.



(a) (b) (c)

Fig. 4. First image with the ball (a) into the observed bricks (the square): of size 4 × 4 × 10 (b) and 4 × 4 × 25 (c). The gray ball separability analysis for MASD and ASD (*N* = 10 (b) and *N* = 25 (c)). Real experiment results.

is compared to the ASD performance. In Fig. 1(a), P_D is shown as a function of SBR for several *b* values, given the secondary data size *GL* = 5000. These curves show that, for a fixed SBR, increasing of the target fill factor enhances the performance of both detectors. One can see that the MASD performance is improved more than ASD. The analysis of the (18) shows [see Fig. 1(a)] that the second term in (18), which is based on the ratio between background powers in the null and alternative hypotheses, is the cause of improving the detection quality. An important parameter of any detector is the gain factor (*GF*) defined as the horizontal displacement at $P_D = 0.9$ between different curves in the Figs. 1–3. Fig. 1(a) shows that the *GF* of the MASD, with respect to ASD, depends on target fill factor *b*. One can see that $GF = 5$ dB at the $b = 0.3$, and for larger $b = 0.7$ the *GF* is increased to $GF = 13$ dB. The performance of the MASD depends on BNR too. In Fig. 1(b), one can see that the MASD performance is improved when the BNR is decreased. For example, when the BNR is decreased from BNR = 100 to BNR = 20, the $GF = 3.5$ dB. Fig. 2(a) shows that the MASD performance depends on the secondary dataset size *GL*. When the $GL = 10000$, the gain factor *GF* = 3 dB with respect to *GL* = 1000. Fig. 2(b) shows that the MASD performance depends on the mismatched factor *MF*. The mismatch between actual and designed one-lag background correlation coefficients $\rho_{\rm ac}$ and ρ_d decreases the performance of MASD. In Fig. 4, the gain factor $GF = 2.5$ dB in the case of $MF = 0.2$ for different pixel fill factors. The factor of signal detection sensitivity *m* changes the detectability of the MASD. Fig. 3 shows that changing *m* from 1 to 0.9 improves (*GF* = 2 dB) the performance of the MASD. We compare the MASD and ASD performances for different target fill factors and successive images using the experimental video sequences of different floating objects on a sea surface. The performance evaluation of detection algorithms in practice is challenging due to the limitations imposed by the limited amount of target data. As a result, the establishment of accurate detection probability curves is quite difficult. We consider a signal at output of the video camera composed of *N* = 10 and 25 successive images (frames). Each pixel in the images is represented by 8-bits, that is, each pixel has 256 gray levels. The digital camera (30 frames/s) is located around 8 m above the sea surface to realize the surface sea monitoring which must give a detection distance less than 600 m (to separate the sky from the sea surface). At the first stage, we have experimentally evaluated the target mode subspace size *p* (the number of the considerable Fourier components) for three small floating objects: the gray ball, the yellow plastic container, and the swimmer in a white hat. In all experiments with different objects, the presented altitudes of the sea waves are less than 0.3 m. The maximum frequency of the Fourier spectrum is equal to 2 Hz for all types of the targets and 12 Hz for sea surface. These values are obtained by using the secondary dataset *GL* = 5000. The number of the considerable target spectrum components *p* depends on *N*: $p = 3$ for $N = 10$ and $p = 7$ for $N = 25$. Then we estimate the average and the variance σ^2_{c+n} of the background data from

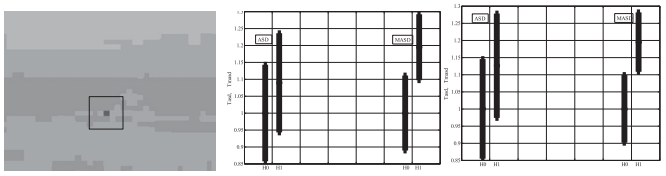


Fig. 5. First image with the ball (a) into the observed bricks (the square): of size $6 \times 6 \times 10$ (b) and $6 \times 6 \times 25$ (for (c)). The gray ball separability analysis for MASD and ASD ($N = 10$ (b) and $N = 25$ (c)). Real experiment results.

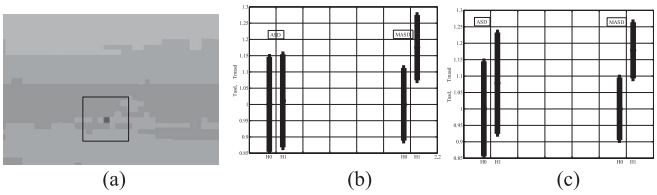


Fig. 6. First image with the ball (a) into the observed bricks (the square): of size $8 \times 8 \times 10$ (b) and $8 \times 8 \times 25$ (c). The gray ball separability analysis for MASD and ASD ($N = 10$ (b) and $N = 25$ (c)). Real experiment results.

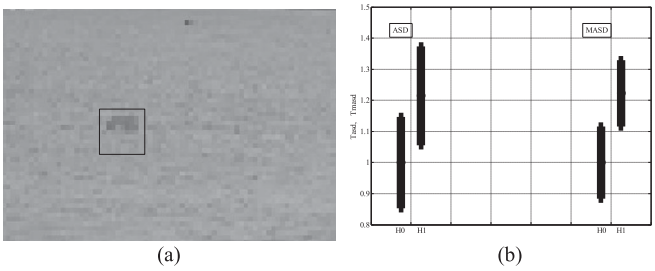


Fig. 7. First image with the plastic container (a) into the observed bricks of size $10 \times 10 \times 100$. The plastic container separability analysis (b) for MASD and ASD.

the observed sea surface by using 5000 successive images. Both detectors, ASD and MASD, use a sequence of optical images, which are first preprocessed by removing the average of the background. The estimates of the parameters p and σ^2_{c+n} are used in ASD and MASD. In order to validate the MASD and ASD, three gray scale image sequences are collected as test samples: the image sequence with floating gray ball, the yellow plastic, and the image sequence with the swimmer. The videos are obtained at a distance of about 200 m from the targets and contain no less than 10 000 images. The set of experiments illustrates the separability between each object type and background for each detector. The figures for each target type are shown in Figs. 6–9. Each figure for gray ball (see Figs. 4–6) contains three subfigures: (a) the first image of the video, (b) the vertical bars for $N = 10$, and (c) the vertical bars for $N = 25$. The bars show the range of detection values under H_0 and H_1 for ASD and MASD. Ideally, these bars should not overlap, indicating that the target is completely separable from background. In the case where overlaps do occur, a number of targets is missed or false alarms occur. One can evaluate as detection performance depends on target fill

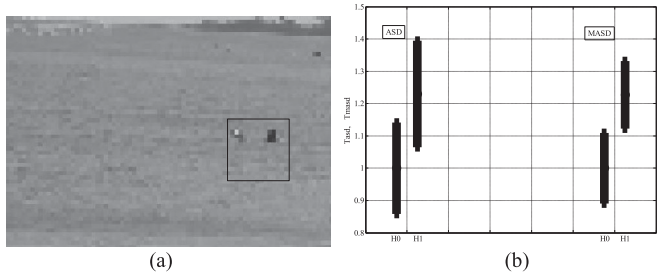


Fig. 8. First image with the swimmer (a) into the observed bricks of size $10 \times 10 \times 100$. The swimmer separability analysis (b) for MASD and ASD.

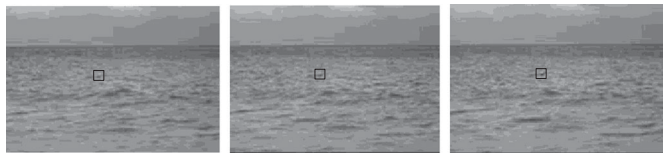


Fig. 9. Images number 77, 80, and 83 of the low-contrast metallic container with the marked observed region.

factors. Figs. 4(a), 5(a), and 6(a) illustrate that the same ball is detected at different brick sizes: $L = 16$ pixels (see Fig. 4), $L = 36$ pixels (see Fig. 5), and $L = 64$ pixels (see Fig. 6). The MASD detected almost without errors at $N = 10$ and $N = 25$ and observed brick size $L = 16$. When the brick size is increased ($L = 36$ and 64), the detection quality is decreased for the same target. The maximum detection performance is achieved when the object occupies all brick pixels L. Minimum size of objects may be known. If we choose the spacial brick size L equal to minimum size of objects, the probability to miss the small object is minimized. Minimum of the number of images in analyzed sequence depends on the spacial size, contrast of the object, dynamic parameters of the sea surface, and determined detection quality. Figs. 4–6 illustrate that the MASD detection performance significantly surpasses the ASD performance. Figs. 7 and 8 illustrate the separability between the plastic container [see Fig. 7(a)], swimmer [see Fig. 8(a)] and background for each detector. Figs. 7(b) and 8(b) show that the MASD detectability significantly outperforms the ASD.

Finally, the detection performances of T_{MASD} and T_{ASD} are compared to two recently proposed M-MSF and FC detectors [6], [7] using real datasets (the floating blue low-contrast metallic container). In Fig. 9, we show the frames number 77, 80, and 83 of the observed brick of size $15 \times 15 \times 100$. It is apparent from Fig. 10 that the reflection intensity from the floating metallic container practically does not change during a shot time-interval between frames. On the contrary, the intensity of the reflections from the sea surface changes quickly. The proposed detector uses this distinction in the best way. For real images, the performances of the four detectors may be shown via the receive operation characteristics curves (see Fig. 11). The results show that T_{MASD} significantly outperforms T_{ASD}, M-MSF, and FC detectors. All the experiment results show that

Fig. 10. Marked observed region of the images number 77, 80, and 83 with the floating metallic container.

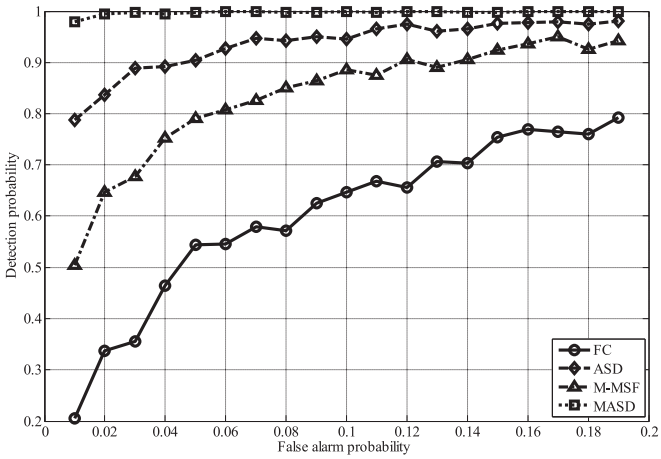


Fig. 11. Receive operation characteristics of the four different detectors for the floating metallic container using the observed brick of size $15 \times 15 \times 100$. Real dataset.

the proposed algorithm T_{MASD} can detect small targets in the presence of intensive background clutter more effectively.

IV. Conclusion

We present an effective and adaptive detector of multipixel objects on the sea surface. The proposed detector offers two principal benefits. The first is that the proposed detector is able to detect no-contrast objects on the sea surface. The second advantage is that the proposed detector is slightly sensitive to the background covariance matrix change and the size of the secondary dataset. We propose the modified 3-D GLR approach based on the analysis of the certain set of the pixel-vectors in the sequence of images to detect the multipixel target with unknown position, size, and shape. The numerical simulations and experimental results show that the proposed adaptive detector considerably outperforms the classical one and the recently proposed M-MSF and FC detectors. The MASD detection efficiency is different under different scenarios and depends on BNR, signal-to-background-plus-noise ratio, target fill factor, and number of images in the analyzed sequence.

Reference

[1] J. Barnett, "Statistical analysis of median subtracting filtering with application to point target detection in infrared backgrounds," *Proc. SPIE*, vol. 1050, pp. 10–15, 1989.
[2] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," *Proc. SPIE*, vol. 3809, pp. 74–83, 1999.
[3] A. Borghgraef, O. Barnich, F. Lapierre, M. V. Droogenbroeck, W. Philips, and M. Acheroy, "An evaluation of pixel-based methods for the detection of floating objects on the sea surface," *EURASIP J. Adv. Signal Process.*, vol. 2010, 2010, Art. no. 978451.
[4] R. Nitzberg, E. H. Takken, D. Friedman, and A. F. Milton, "Spatial filtering techniques for infrared sensors," *Proc. SPIE*, vol. 0178, pp. 40–58, 1979.
[5] V. T. Tom, T. Peli, M. Leung, and J. E. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," *Proc. SPIE*, vol. 1954, pp. 2–11, 1993.
[6] S. Kim and J. Lee, "Small infrared target detection by region-adaptive clutter rejection for sea-based infrared search and track," *Sensors*, vol. 14, no. 7, pp. 13210–13242, 2014.
[7] A. Kadyrov, H. Yu, and H. Liu, "Ship detection and segmentation using image correlation," in *Proc. IEEE Int. Conf. Syst. Man, Cybern.*, Manchester, SMC 2013 U.K., 2013, pp. 3119–3126.
[8] L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Reading, MA, USA: Addison-Wesley, 1991.
[9] S. C. Pohlig, "Spatial-temporal detection of electro-optic moving targets," *IEEE Trans. Aerospace Electron. Syst.*, vol. AES-31, no. 2, pp. 608–616, Jan. 1995.
[10] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace Electron. Syst.*, vol. AES-10, no. 1, pp. 853–863, Nov. 1974.
[11] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerospace Electron. Syst.*, vol. AES-22, no. 1, pp. 115–127, Jan. 1986.
[12] S. Kraut, L. Scharf, and L. McWorther, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.
[13] V.Golikov and O. Lebedeva, "Adaptive detection of subpixel targets with hypothesis dependent background power," *IEEE Signal Process. Lett.*, vol. 20, no. 8, pp. 751–754, Aug. 2013.
[14] V. Golikov, M. Rodriguez-Blanco, and O. Lebedeva, "Robust multipixel matched subspace detection with signal-dependent background power," *J. Appl. Remote Sens.*, vol. 10, no. 1, Jan.–Mar. 2016, Art. no. 015006, doi:10.1117/1.JRS.10.015006.

Marco Rodriguez-Blanco (M'04) was born in Mexico. He received the M.S. and Ph.D degrees in electrical engineering from CENIDET of Cuernavaca, Morelos, Mexico in 2001 and 2009, respectively.

He is currently a Professor at the University Autonomous of Carmen, Ciudad del Carmen, Mexico. His main research interest includes electronic semiconductor devices. He is the author of more than 15 international scientific and conference papers.

Victor Golikov (M'03) received the M.S. degree in radio-physics from the National University of Kharkov, Kharkiv, Ukraine and the Ph.D. degree in radar systems from the National Radio-Electronic University of Kharkiv, Kharkiv, in 1970 and 1976, respectively.

He is a currently Professor at the Autonomous University of Carmen, Ciudad del Carmen, Mexico. He is the author of more than 60 journal papers. His current research interest includes remote optoelectronic systems.