

Bayesian Target Detection Algorithms for Solid Subpixel Targets in Hyperspectral Images

Stefania Matteoli^{ID}, *Senior Member, IEEE*, and James Theiler^{ID}, *Member, IEEE*

Abstract—We investigate the use of Bayesian methods for hyperspectral subpixel target detection, where the uncertainty associated with the target fill factor is “probabilized” by a suitable prior. Specifically, we present a general framework for Bayesian target detection by employing different models for the background distribution, comparing different choices for the Bayesian prior, and investigating different numerical schemes for evaluating the Bayesian integral. The Bayesian methods are furthermore compared to their generalized likelihood ratio test (GLRT)-based counterparts. Experiments performed over real hyperspectral imagery, with both real and implanted subpixel targets, show that incorporating prior knowledge by means of nonuniform priors emphasizing smaller target fill factors outperforms usage of the “noninformative” uniform prior and enhances Bayes performance beyond the GLRT, a result observed for both parametric and nonparametric background models. We find that even “rough” priors can successfully leverage the context-based information by emphasizing target sizes that are of most interest. We further observe that the Gauss–Legendre numerical integration scheme provides efficient integral approximation while maintaining the desirable admissibility property of Bayesian methods.

Index Terms—Bayesian statistics, generalized likelihood ratio test (GLRT), hyperspectral, kernel density estimation, likelihood ratio, target detection.

I. INTRODUCTION

SEARCHING a hyperspectral image for a target based on its spectral signature is a data processing task with many applications, ranging from forestry and mineral prospecting to pollution monitoring and public safety. With the richness of information content offered by hundreds of spectral bands at every image pixel, hyperspectral imaging has shown great potential in material discrimination. In many applications, the targets to be searched for are *subpixel*, which is to say that they are smaller than the pixel size on the ground. Detecting subpixel targets means handling a spectral signal including not only the target component but also the signal component pertinent to the *background*, which is composed of the nontarget materials in the given pixel [1].

Manuscript received 27 February 2023; revised 1 June 2023; accepted 29 June 2023. Date of publication 4 July 2023; date of current version 18 July 2023. The work of James Theiler was supported by the Laboratory Directed Research and Development (LDRD) Program at the Los Alamos National Laboratory, Los Alamos, NM, USA. (*Corresponding author: James Theiler.*)

Stefania Matteoli is with the Institute of Electronics, Information Engineering and Telecommunications (IEIIT), Consiglio Nazionale delle Ricerche (CNR), 56122 Pisa, Italy (e-mail: stefania.matteoli@cnr.it).

James Theiler is with the Space Remote Sensing and Data Science Group, Los Alamos National Laboratory, Los Alamos, NM 87545 USA (e-mail: jt@lanl.gov).

Digital Object Identifier 10.1109/TGRS.2023.3292067

Among the various approaches proposed to deal with subpixel targets, this article focuses on decision theory-based statistical algorithms for which the likelihood ratio test (LRT) [2] between the probability density functions (pdfs) of the pixel conditioned to the two competing (target present and target absent) hypotheses is applied at each image pixel. By modeling the spectral variability of a pixel containing a subpixel target with the replacement target model (RTM) [1], [3], the target-present pdf can be expressed as a function of the background pdf and the target signature and the resulting LRT becomes a composite hypothesis testing problem depending on the target *fill factor* (or abundance), which is the RTM key parameter. The most widely employed approach to composite hypothesis testing is to replace the unknown target fill factor with its maximum likelihood (ML) estimate, thus obtaining the generalized LRT (GLRT) [2]. Other approaches exist such as penalized LRT and clairvoyant fusion [4], [5], [6], which result in a “weighted GLRT” benefiting from a certain degree of design flexibility. Unlike these GLRT-based detectors, the Bayesian detectors integrate the likelihood over the range of fill-factor values by weighting it with a suitable prior pdf for the fill factor itself. An important advantage of Bayesian detectors is that they are guaranteed to be *admissible* [7]. This is an attractive benefit: if a detector is admissible, then no other detector is *uniformly* more powerful, that is to say, for a given false alarm rate (FAR), no other detector has a higher detection rate (DR) for *all* target fill factors. While this does not mean an admissible detector is unambiguously optimal, it does mean that no other detector is unambiguously better. Thus, when searching for a detector that is well-suited to a particular operational scenario, it is desirable to restrict attention to the class of admissible detectors.

Although the potential of the Bayesian approach for detecting targets in hyperspectral images was acknowledged several years ago [8], [9], Bayesian detectors are not widely employed in the hyperspectral detection literature. One reason may be the difficulty in finding closed-form solutions (mostly due to the integration they require); another may be that the flexibility to choose the prior turns out to be a double-edged sword—it is rarely obvious which prior should be associated with a given problem. One issue is the very interpretation of the prior. The traditional interpretation is that it represents the distribution of fill factors that one expects to see, a priori, in the imagery to be searched. A slightly different view is provided in [7], where the prior is seen as a means to express the importance that is attached to the various values of the fill factor. Despite this

difficulty, we consider the flexibility of the Bayesian detector to be more of a blessing than a curse; and combining that with the built-in admissibility property, we believe the Bayesian detection of subpixel targets merits further investigation.

In recent preliminary studies [10], [11], [12], we began exploring the potential of Bayesian detection in hyperspectral imagery: a basic implementation of a numerical solution to Bayesian detection with uniform prior was introduced in [10]; by employing a specific prior with fixed parameters along with certain simplifying assumptions, an approximate closed-form Bayesian detector was derived in [11]; and the utility of a single delta-function prior was investigated in [12]. This article follows the thread outlined by those studies by presenting a more general and structured framework for Bayesian target detection: 1) by enabling the use of arbitrary background models (as demonstrated by experiments using nontrivial parametric and nonparametric background models); 2) by enabling the flexibility to use arbitrary prior pdfs (as demonstrated by the use of a uniform, a nonuniform, and even a strictly “improper” prior); 3) by investigating and comparing the effectiveness of two different numerical integration schemes; 4) by presenting experimental results over both synthetic and real scenarios featuring rare targets; and 5) by comparing detection performance of Bayesian detectors with their more conventional GLRT-based counterparts. We find that Bayesian detectors represent a valid alternative to the more widely employed GLRT, offering not only admissibility but also flexibility with respect to the users’ knowledge and interest about the scenario, and we observe that this flexibility can result in improved performance.

The rest of this article is organized as follows. Section II explains the Bayesian approach to solid subpixel target detection, with examples of two Bayesian target detectors. Section III describes the experiments and discusses their results. Finally, Section IV summarizes the main outcomes and contributions of this work.

II. BAYESIAN TARGET DETECTORS FOR SUBPIXEL TARGETS

A. Detection of Solid Subpixel Targets

The LRT between the target-present (H_1) and the target-absent (H_0 , i.e., just background) pdfs is expressed as follows:

$$\Lambda_{\text{LRT}}(\mathbf{x}, \mathbf{t}) = \frac{f_{\mathbf{x}|H_1}(\mathbf{x}, \mathbf{t})}{f_{\mathbf{x}|H_0}(\mathbf{x})} \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^d$ is the generic pixel (with d being the spectral dimension) and $\mathbf{t} \in \mathbb{R}^d$ is the target spectral signature— \mathbf{t} and \mathbf{x} are assumed to be radiometrically compatible. Both conditional pdfs are unknown and need to be estimated from the data. Because targets are *rare* in the scene, the number of available realizations of $\mathbf{x}|H_1$ generally does not suffice to the $f_{\mathbf{x}|H_1}(\mathbf{x}, \mathbf{t})$ estimation. Thus, the target-present pdf is generally expressed as a function of the background pdf and the target spectral signature \mathbf{t} by means of target spectral variability models, which specify how the observed pixel depends on the background pixel in combination with the target signature and its abundance [3], [13]. One widely

employed spectral variability model that is physically consistent with the phenomenology of subpixel targets is the replacement target model (RTM) [1], [3], which fractionates the pixel in two components, the pixel fraction where the target obscures the background and the residual background-only fraction, as follows:

$$\begin{cases} H_1, & \mathbf{x} = \alpha\mathbf{t} + (1 - \alpha)\mathbf{z} \\ H_0, & \mathbf{x} = \mathbf{z} \end{cases} \quad (2)$$

where $\mathbf{z} \in \mathbb{R}^d$ is the generic background spectrum and $\alpha \in (0, 1]$ denotes target *fill factor* and quantifies the fraction of the pixel occupied by the target itself.

By the change of variable $\mathbf{x} = \psi(\mathbf{z})$, we obtain $f_{\mathbf{x}|H_1}(\mathbf{x}, \mathbf{t}) = f_{\mathbf{z}}[\psi^{-1}(\mathbf{x})]|d\psi/d\mathbf{x}|^{-1}$ and $f_{\mathbf{x}|H_0}(\mathbf{x}) = f_{\mathbf{z}}(\mathbf{x})$, and thus, (1) turns into

$$\Lambda_{\text{LRT}}^{\text{RTM}}(\mathbf{x}, \mathbf{t}; \alpha) = \frac{(1 - \alpha)^{-d} f_{\mathbf{z}}\left(\frac{\mathbf{x} - \alpha\mathbf{t}}{1 - \alpha}\right)}{f_{\mathbf{z}}(\mathbf{x})}. \quad (3)$$

If the fill factor α were known, the detector in (3) (known as *clairvoyant*) would provide optimal detection [2]. In real scenarios, α is not known and we need to solve the composite hypothesis testing problem and produce a detector that does not depend on α .

B. Bayesian Approach to Subpixel Target Detection

A Bayesian LRT detector can be derived from the composite hypothesis testing in (3) by posing a prior pdf $f_{\alpha}(\alpha)$ for α and marginalizing α out as follows:

$$\Lambda_{\text{Bayes}}^{\text{RTM}}(\mathbf{x}, \mathbf{t}) = \frac{\int_{\alpha=0}^1 (1 - \alpha)^{-d} f_{\mathbf{z}}\left(\frac{\mathbf{x} - \alpha\mathbf{t}}{1 - \alpha}\right) f_{\alpha}(\alpha) d\alpha}{f_{\mathbf{z}}(\mathbf{x})}. \quad (4)$$

In order to fully specify this detector, not only the prior $f_{\alpha}(\alpha)$ needs to be specified but also a model for the background pdf $f_{\mathbf{z}}(\cdot)$. Here, we specifically investigate two Bayesian detectors by leveraging two very well-known background models, namely, a long-tailed (parametric) distribution and a data-driven kernel-based (nonparametric) model.

1) *Parametric Bayesian Target Detector*: Invoking parametric models means assuming a specific parametric form for the background pdf and replacing the parameters with their estimates made from the available data. Long-tailed distributions have been shown to be particularly suitable for modeling hyperspectral backgrounds [14] and have been successfully employed within the GLRT approach to detect rare weak targets [15], [16]. For the multivariate t distribution with mean vector $\boldsymbol{\mu}$, covariance matrix \mathbf{C} , and number of degrees of freedom ν , the background pdf is given as follows:

$$f_{\mathbf{z}}^{(P)}(\mathbf{z}; \boldsymbol{\mu}, \mathbf{C}, \nu) = c|\mathbf{C}|^{-\frac{1}{2}} \left[1 + \frac{(\mathbf{z} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{z} - \boldsymbol{\mu})}{\nu - 2} \right]^{-\frac{d+\nu}{2}} \quad (5)$$

where c is a constant (depending on ν and d).

In practice, the parameters ($\boldsymbol{\mu}$, \mathbf{C} , and ν) are not known and need to be estimated from the available data, i.e., from a sufficiently large set $\{\mathbf{z}_n\}_{n=1}^N$ of target-free secondary data,

assumed to be identically distributed as \mathbf{z} . Due to the *target rarity assumption*, all image pixels are generally employed for this purpose.

The parametric Bayesian detector based on the multivariate t distribution can thus be written as

$$\Lambda_{\text{Bayes}}^{\text{RTM}(P)}(\mathbf{x}, \{\mathbf{z}_n\}_{n=1}^N, \mathbf{t}) = \frac{\int_{\alpha=0}^1 (1-\alpha)^{-d} h\left(\frac{x-\alpha t}{1-\alpha}; \hat{\boldsymbol{\mu}}, \hat{\mathbf{C}}, \hat{\nu}, d\right) f_{\alpha}(\alpha) d\alpha}{h(\mathbf{x}; \hat{\boldsymbol{\mu}}, \hat{\mathbf{C}}, \hat{\nu}, d)} \quad (6)$$

with

$$h(\mathbf{z}; \hat{\boldsymbol{\mu}}, \hat{\mathbf{C}}, \hat{\nu}, d) = \left[1 + \frac{(\mathbf{z} - \hat{\boldsymbol{\mu}})^T \hat{\mathbf{C}}^{-1} (\mathbf{z} - \hat{\boldsymbol{\mu}})}{\hat{\nu} - 2} \right]^{-\frac{d+\hat{\nu}}{2}}.$$

2) *Nonparametric Bayesian Target Detector*: The nonparametric approach does not rely upon a specific distributional form for the background pdf but directly estimates it in a data-driven fashion. Here, we consider the variable-bandwidth kernel density estimator (VKDE) [17], [18], which has been successfully employed within the LRT for detecting rare targets and spectral anomalies in hyperspectral images [19], [20]. More specifically, the background pdf is obtained as follows:

$$f_{\mathbf{z}}^{(\text{NP})}(\mathbf{z}, \{\mathbf{z}_n\}_{n=1}^N) = \frac{1}{N} \sum_{n=1}^N \frac{1}{r_k^d(\mathbf{z}_n)} \kappa \left[\frac{\mathbf{z} - \mathbf{z}_n}{r_k(\mathbf{z}_n)} \right] \quad (7)$$

where $\kappa[\cdot]$ is a kernel function whose bandwidth $r_k(\mathbf{z}_n)$ depends on the local data density in the feature space and is specifically evaluated at each \mathbf{z}_n as its Euclidean distance to its k -nearest neighbor (k -NN) in $\{\mathbf{z}_i\}_{i=1, i \neq n}^N$. The k -NN-based bandwidth selection allows for an automated tuning of the kernel smoothing with resulting performance weakly dependent on the free parameter k [17], [18].

The nonparametric Bayesian detector based on the VKDE can thus be written as

$$\Lambda_{\text{Bayes}}^{\text{RTM}(\text{NP})}(\mathbf{x}, \{\mathbf{z}_n\}_{n=1}^N, \mathbf{t}) = \frac{\int_{\alpha=0}^1 (1-\alpha)^{-d} \sum_{n=1}^N \frac{1}{r_k^d(\mathbf{z}_n)} \kappa \left[\frac{x-\alpha t - \mathbf{z}_n}{r_k(\mathbf{z}_n)} \right] f_{\alpha}(\alpha) d\alpha}{\sum_{n=1}^N \frac{1}{r_k^d(\mathbf{z}_n)} \kappa \left[\frac{x - \mathbf{z}_n}{r_k(\mathbf{z}_n)} \right]}. \quad (8)$$

3) *Prior PDFs*: Choosing the prior is a classic (and sometimes even controversial) issue in Bayesian statistics. Ideally, prior choice should be driven by a priori knowledge and, traditionally, it is set based on the best guess for what the unknown parameter distribution is likely to be. For instance, if targets are expected to be much smaller than a pixel, one should design a prior that puts more weight on small fill factors.

In principle, if no prior knowledge is available, a *non-informative* prior should be used, this is a prior as “flat” as possible [2]. Because of the bounded domain of the fill factor α , the most natural choice for a noninformative prior is the uniform pdf.

In [11], the prior was chosen to simplify the mathematical derivation of an approximated closed-form solution, thus trading analytical tractability for the user’s freedom to set the prior based on the actual available knowledge. Here, by contrast,

we do not impose any constraint on the prior, giving up on the derivation of (even approximated) closed-form solutions in favor of providing the user with the maximum flexibility in designing the prior. In fact, we believe that even when the available knowledge may be vague, scarce, or somehow not readily expressible as a pdf, a properly set “rough” prior would provide benefits relative to usage of a noninformative prior. In addition, for detection problems, an alternative interpretation is possible [7], where prior choice should not be limited to specify how likely some given fill-factor values are expected to be, but should also specify how important those values of fill factor are to the user. In this case, for example, putting more weight on small fill factors aims at enhanced DRs at those small fill factors, possibly at the expense of reduced DRs at larger fill factors.

In this work, we will leverage the context-based available information while at the same time directing the detection toward smaller target sizes. In doing so, besides proper priors, we also allow for “improper” priors (not finite or with infinite integral), as long as they can produce posterior pdfs with bounded integrals.

4) *Integration Schemes*: Two numerical integration schemes are considered in this article, the midpoint (MP) rule and Gaussian–Legendre (GL) quadrature [21], to numerically approximate the integral at the numerator of the Bayesian detectors. For simplicity, let us write $q(\alpha)$ as the integrand function, and observe that both rules can be expressed in the form

$$\int_0^1 q(\alpha) d\alpha \approx \sum_{i=1}^{n_{\text{pts}}} w_i q(\alpha_i) \quad (9)$$

but with different choices for w_i and α_i . Since the computation in (9) is dominated by the n_{pts} evaluations of the integrand, the cost depends on n_{pts} , but not on whether the MP or GL rule is employed.

According to the MP rule, the integration interval is divided into equally wide subintervals identified by their n_{pts} uniformly spaced MPs $\{\alpha_i^{(\text{MP})} = i/n_{\text{pts}} - 1/(2n_{\text{pts}})\}_{i=1}^{n_{\text{pts}}}$ and the integral is numerically approximated by the sum of the areas of n_{pts} rectangles of width $w_i^{(\text{MP})} = 1/n_{\text{pts}}$ and height $q(\alpha_i^{(\text{MP})})$.

Approximation of an integral with the GL rule means, instead, adopting a nonuniform spacing of the n_{pts} points where the function is evaluated. Specifically, the positioning of the sampling points within the integration domain is given by $\{\alpha_i^{(\text{GL})} = (\xi_i + 1)/2\}_{i=1}^{n_{\text{pts}}}$, where $\{\xi_i\}_{i=1}^{n_{\text{pts}}}$ are the roots of the n_{pts} -degree Legendre polynomial $L_{n_{\text{pts}}}(\xi)$, and the weights are given by

$$w_i^{(\text{GL})} = \frac{1}{(1 - \xi_i)^2 \left[L'_{n_{\text{pts}}}(\xi_i) \right]^2}. \quad (10)$$

The roots $\{\xi_i\}_{i=1}^{n_{\text{pts}}}$ can be found with recursive methods (e.g., the Newton–Raphson method), but both roots and weights are only calculated once (and have been extensively tabulated; e.g., [21, Table 25.2]), whereas the integrand evaluations are performed for every pixel.

It should be noted that approximating the integral at the numerator of a Bayesian detector with either the MP or the

GL rule does not impair the admissibility of the detector itself. In fact, approximating the integral with a sum over discrete terms is equivalent to a Bayesian detector with a different prior (namely, a weighted delta-comb function) and, thus, the admissibility property is maintained.

III. EXPERIMENTAL RESULTS

This section describes the experimental design and discusses the experimental results obtained.

A. Experimental Design

1) *Algorithmic Comparison*: The Bayesian detectors derived in this article from the RTM are compared in this section to their RTM-based GLRT counterparts. For the parametric background model in (6), the GLRT counterpart is a closed-form expression, given by the elliptically contoured finite target matched filter (EC-FTMF) [16], [22]. For the nonparametric background in (8), the RTM-based NP-GLRT is not solvable in the closed form [19], but a numerical solution was implemented here, using the same sampling points $\{\alpha_i^{(*)}\}_{i=1}^{n_{\text{pts}}}$ (with $*$ representing either the MP or GL rule) that were used for the numerical integration in the Bayesian algorithms. In the experiments with controlled fill factors (see Sections III-A4.a and III-B1), the clairvoyant versions of both parametric and nonparametric algorithms were introduced in the comparison as references. While a large variety of detectors based on different approaches have been proposed in the literature (e.g., [23], [24], [25], and [26]), in order to keep focus, we have circumscribed our attention to statistical LRT RTM-based detectors. As mentioned, the goal of this article is to investigate Bayesian methods as a robust, flexible, and admissible alternative to GLRT. The comparison of Bayes methods to other different detectors will be the subject of future work.

2) *Methodology Settings*: For the parametric Bayesian detector, ML estimates $(\hat{\boldsymbol{\mu}}, \hat{\mathbf{C}})$ were used for $\boldsymbol{\mu}$ and \mathbf{C} , whereas a simple strategy based on the method of moments was employed for ν . Specifically, as done in [27], ν was estimated by the following ratio between moments of the EC- t distribution:

$$\hat{\nu} = 2 + \frac{p\tilde{\kappa}_p}{\tilde{\kappa}_p - (L + p)} \quad (11)$$

where $\tilde{\kappa}_p = (\hat{m}_{p+2}/\hat{m}_p)$ and \hat{m}_p is the sample estimate made on $\{\mathbf{z}_n\}_{n=1}^N$ of $E\{r^p\}$, with $r = [(\mathbf{z} - \hat{\boldsymbol{\mu}})^T \hat{\mathbf{C}}^{-1} (\mathbf{z} - \hat{\boldsymbol{\mu}})]^{0.5}$. We chose p to be equal to one.

For the nonparametric detectors, because the VKDE in (8) employs a spherically symmetric kernel, the data were linearly transformed to have unit variances in all spectral directions before detector application, as suggested in [28] and in the KDE literature [17]. The kernel function was taken as the Epanechnikov kernel function, which exhibits several desirable properties [17]. The number k of NNs in VKDE was chosen to satisfy $N^{1/3} < k \ll N^{1/2}$, as recommended in [19].

With regard to the integration to be computed in both Bayesian detectors, we employed the GL quadrature rule with a number of points $n_{\text{pts}} = 6$. The resulting points $\{\alpha_i^{(\text{GL})}\}_{i=1}^6$

were also employed to sample the likelihood function in the NP-GLRT detector. Investigations into the integration strategies are addressed in detail in Sections III-A4.c and III-B3.

For the prior pdf functions for the Bayesian detectors, the uniform prior was employed in all experiments, allowing the comparison to be carried out at the “integral” versus “peak” level (see Sections III-A4.a and III-B1) and at the integration scheme level (see Sections III-A4.c and III-B3). In another experiment (see Sections III-A4.b and III-B2), Bayes flexibility with respect to context-based a priori information was explored by testing two further priors, both suitable to model a subpixel target detection scenario.

The first nonuniform prior function tested was a beta pdf [21]

$$f_{\alpha}(\alpha) = f_{\alpha}^{\text{beta}}(\alpha; a, b) = \alpha^{a-1}(1-\alpha)^{b-1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \quad (12)$$

where a and b are the nonnegative shape parameters and $\Gamma(\cdot)$ is the gamma function. The beta pdf naturally models subpixel target detection scenarios because its support is $[0, 1]$. By adjusting the shape parameters, a variety of configurations may be obtained, acting not only on the mean value $E\{\alpha\} = a/(a+b)$ but also on the entire pdf shape.

We also tested the following prior, which follows a power law and puts more weight on small fill factors:

$$f_{\alpha}(\alpha) = 1/\alpha^m \quad (13)$$

where $m > 0$ tunes the subpixel weighting. Note that even though this would be an improper prior over a continuous interval that included $\alpha = 0$, our finite-sample approach is not strictly improper since we use a finite sum of delta functions whose amplitudes scale like $1/\alpha^m$. Because the finite sum avoids $\alpha = 0$, the actual prior has a bounded integral.

3) *Performance Measures*: Target detection performance is typically measured by making reference to two main basic quantities that leverage ground-truth data of target positions and are evaluated by thresholding the detection test statistic, namely, the DR and the FAR. DR is the ratio between the number of target pixels exceeding the threshold in the detection statistic (correct detections) and the total number of target pixels in the scene—the higher the better. FAR is the ratio between the number of nontarget (background) pixels exceeding the threshold (false detections) and the total number of background pixels—the lower the better. In our experiments, we built the *empirical* receiver operating characteristic (ROC) curves by plotting the DR versus the FAR and evaluated over the detection statistic while varying the detection threshold. ROC curves embody the fundamental tradeoff in target detection performance, i.e., DR and FAR cannot both be optimized, and provide an overall perspective on detection performance.

We also extracted scalar-valued summary performance measures from the ROC curves, specifically the FAR evaluated for several given values of DR (hereinafter, FAR@DR) and the area under the ROC curve (AUC), computed in two different ways. The *empirical* ROC curve, with DR and FAR plotted for every possible threshold, is typically a stair-step shape (this is strictly true if all pixels have distinct detection statistic values)

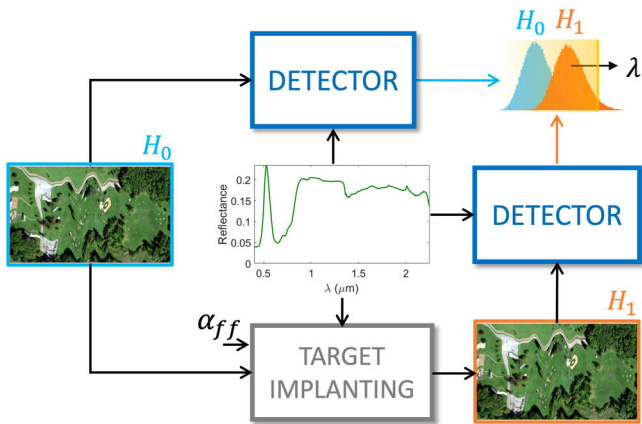


Fig. 1. Scheme of the matched-pair method for target implanting. A matched pair of images is used: the first acts as a background image (embodying the null hypothesis H_0), whereas the second (representing the alternative hypothesis H_1) is a replica of the first image but with targets implanted in each pixel with a given fill factor α_{ff} . The detection algorithm is applied to the pair of images. A ROC curve is obtained by thresholding the histograms of the corresponding detection statistics.

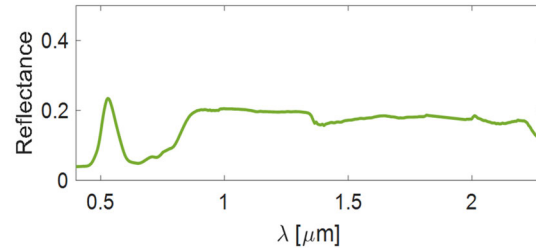
and our standard AUC was evaluated as the area under this stair-step curve. It has been noted in [29], however, that ROC curves should theoretically be “convex-cap,” so we also report convex-AUC, which is the area under the upper convex hull of the empirical ROC curve. While AUC and some of its variants are widely used in the target detection literature, we consider the FAR-based statistics as more informative for scenarios in which the actual targets are relatively rare since it is the low-FAR region of the ROC curve that is more important in that case. Besides ROC curves and the scalar-valued summary performance metrics, we also provide images of detection statistics as well as their histograms, so as to provide effective visual evidence of algorithm behavior.

4) *Subpixel Target Detection Experiments*: Three types of experiments were carried out, which are described in the following subsections. The first set of experiments featured a real hyperspectral image and a subpixel target detection scenario reproduced with controlled fill factors. Then, we further validated the methods on a real hyperspectral image encompassing two ground-truth subpixel target detection scenarios. Finally, the third set of experiments was aimed at providing insights about the Bayes integration scheme. All three sets of experiments were carried out on hyperspectral data fully available to the scientific community.

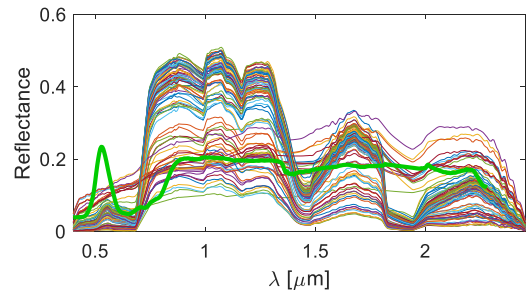
a) *Experiments with controlled fill factor*: In the first experiment, we reproduced a subpixel target detection scenario with controlled fill factors by means of the *matched-pair method for target implanting* [30], [31]. This allows the algorithms to be tested on a real hyperspectral image but in a controlled environment that enables a large number of target pixels to be tested and, thus, statistically reliable ROC curves to be drawn. As shown in Fig. 1, the original hyperspectral image is used as “background image,” producing N instances of $\mathbf{x}|H_0$; then, a second image is obtained by implanting a given target spectrum in each of the N image pixels with a user-specified fill factor α_{ff} , thus obtaining N instances of



(a)



(b)



(c)

Fig. 2. (a) RGB representation of the hyperspectral image used in the matched-pair experiments. (b) Reflectance spectrum used for target implanting. (c) Target spectrum superimposed to spectra extracted from randomly selected image pixels.

the test pixel under the alternative “target-present” hypothesis. By running the desired set of algorithms on both images and thresholding the histograms of the corresponding test statistics, ROC curves can be drawn.

The hyperspectral image employed in this experiment is shown in Fig. 2(a), and the spectrum used for implanting, which was measured over a green wooden panel [not present in the image of Fig. 2(a)], is plotted in Fig. 2(b). Both image and target spectrum were taken from the SHARE 2012 collection campaign data [32], [33] acquired over Avon, NY, USA, the characteristics of which are summarized in Table I. Fig. 2(c) plots the target spectrum superimposed to spectra extracted from randomly selected image pixels.

The image consists of 250×419 pixels and, based on recommendations from previous works extensively examining these data [34], [35], a subset of 50 bands in the 218–643-nm range was employed for processing.

Target implanting was performed with $\alpha_{ff} = \{0.05, 0.075, 0.10, 0.0125, 0.15\}$, producing five different target detection scenarios.

As noted earlier, the uniform prior was here employed in Bayesian algorithms. Numerical integration was conducted

TABLE I
AVON DATA CHARACTERISTICS

SENSOR	ProSPECTIR- VS2
SPECTRAL RANGE [nm]	400 - 2450
Average FWHM [nm]	~5
# SPECTRAL BANDS	360
NOMINAL IFOV [mrad]	1.3
SCENARIO	Rural with man- made structures
ELEVATION a.g.l [Km]	~1.5-2.5

with the GL rule with $n_{\text{pts}} = 6$. The number of NNs employed for VKDE was $k = 100$, fully satisfying the practical recommendation mentioned above.

b) Real subpixel target detection scenario: The SHARE2012 subpixel target detection experiment [32], [34] was leveraged for further validation of the methods. We make specific reference to the hyperspectral image acquired over Avon, NY, USA, where about 100 wooden panels (each 20" by 12") were deployed in two sets of 50 panels—each set arranged with panels placed in such a way that the target fill factor resulted in (much) lower than 20%. The first set of targets consisted of green wooden panels arranged on a patch of grass, whereas the other set featured yellow wooden panels arranged on a basketball court. The hyperspectral image with the two sets of targets is shown in Fig. 3(a), and ground-truth photographs of the green and yellow targets are shown in Fig. 3(b) and (c), respectively, while Fig. 3(d) and (e) shows the target reflectance spectra. The hyperspectral image consisted of 150×400 pixels, and the same band subset as in the previous experiment [34], [35] was employed for processing. The image characteristics are summarized in Table I.

All three priors described above were employed in this experiment. Context-based a priori information was leveraged to select prior parameters. Specifically, the beta pdf shape parameters were selected as $a = 0.5$ and $b = d$, whereas $m = d$ was taken for the power-law prior—both choices were made to craft pdf functions strongly favoring subpixel scenarios. Use of these priors allowed us to inject our *context-based* a priori knowledge of target size into the Bayesian approach.

Numerical integration was performed with the GL rule applied with $n_{\text{pts}} = 6$ (see Section III-B3).

The number of NNs in VKDE was taken as in the previous experiment ($k = 100$), which totally fits with the aforementioned practical recommendation. In this article, the performance will be examined by means of ROC curves, FAR at given DRs, and AUC metrics.

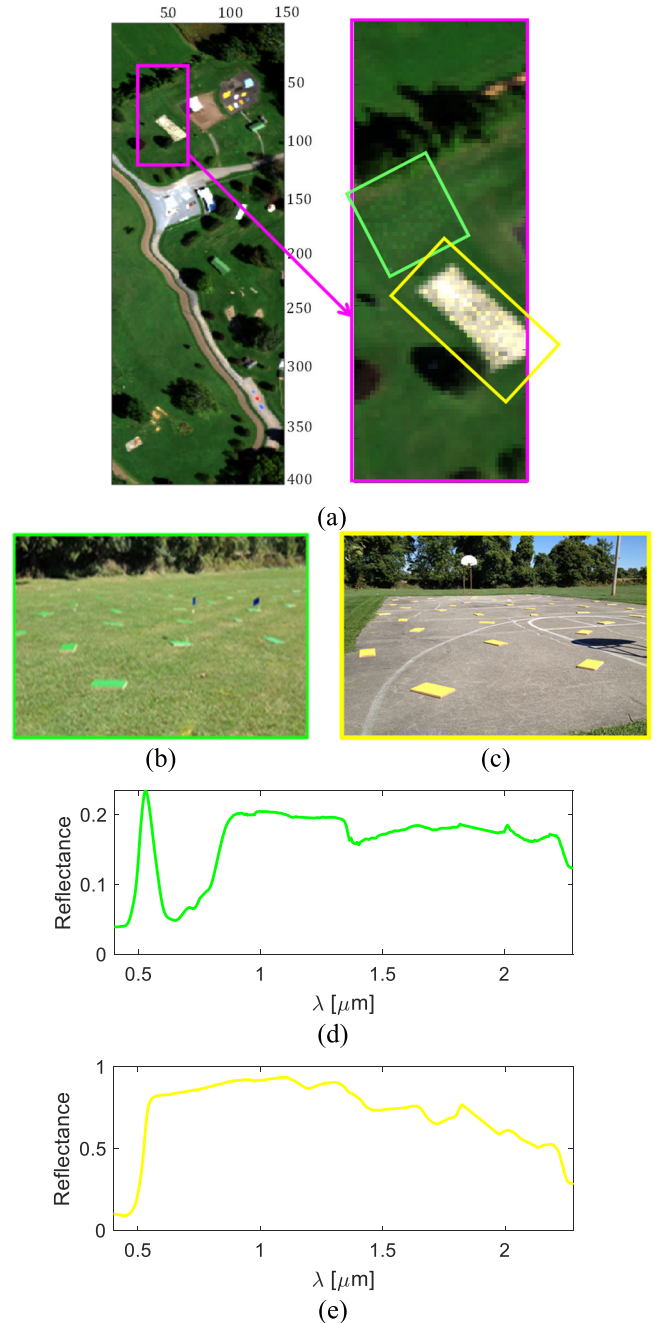


Fig. 3. (a) RGB representation of the hyperspectral image used. The target area is highlighted in magenta—an enlarged view is displayed where the areas including the two types of green and yellow wooden panels are marked in green and yellow. (b) and (c) Ground-truth photographs (taken from [33]) of the green and yellow panels. (d) and (e) Reflectance spectra of the green and yellow panels.

c) Examining different rules for the integration scheme: In this set of experiments, we examined two strategies for numerically estimating the Bayesian integral [i.e., the numerator in (4)]. To this end, we used the same data as in the first experiment, specifically the matched image pair obtained with a fill factor of $\alpha_{\text{ff}} = 0.05$. Bayesian algorithms were applied with the uniform prior and $k = 100$ was selected for the VKDE.

Both MP and GL rules were examined in terms of integral approximation capability and impact on the detection performance. Specifically, we evaluate numerical integration with

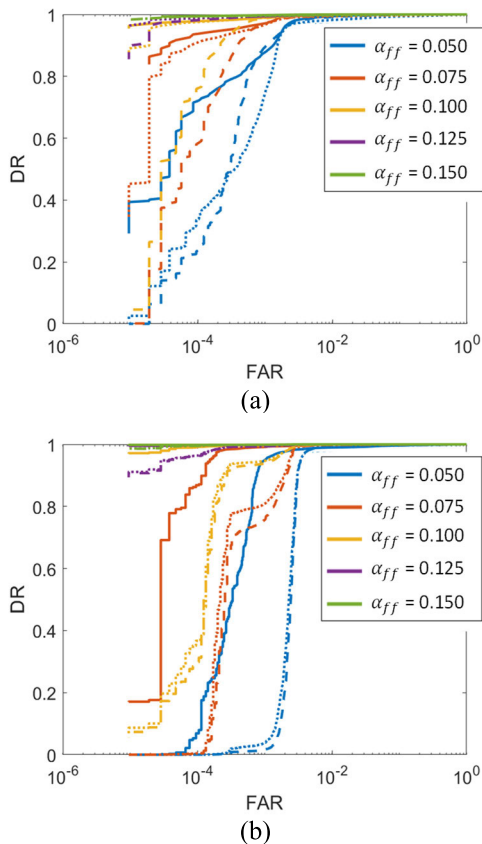


Fig. 4. ROC curves for (a) parametric and (b) nonparametric algorithms, parameterized with respect to the target fill factor. Solid curves represent clairvoyant, dotted curves indicate GLRT, and dashed curves denote Bayes.

both rules and for both parametric and nonparametric methods for $n_{\text{pts}} = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 50\}$. The approximation capability performance was evaluated by examining log-integral curves, plotted with respect to n_{pts} , and evaluated over 150 random pixels extracted from both images of the pair. Here, we do not seek the most accurate approximation possible but an approximation good enough to enable computational efficiency (low n_{pts}) and good detection performance. For the latter, we used the FAR obtained for a DR of 0.5 (FAR@DR=0.5) evaluated over the image pair and plotted against n_{pts} .

B. Results Discussion

Here, we present and discuss the results for the three sets of experiments.

1) *Results for the Experiments With Controlled Fill Factor:* ROC curves for the matched-pair experiments are shown in Fig. 4(a) and (b) for parametric and nonparametric algorithms, respectively, where clairvoyant versions of the algorithms are also displayed. ROC curves are here parameterized with respect to the target fill factor α_{ff} .

The graphs show that the parametric methods exhibit—in the low FAR region of greatest interest—a larger difference in performance between Bayes (with uniform prior) and GLRT, in favor of the latter (e.g., for $\alpha_{\text{ff}} = 0.075$ and DR = 0.8, the GLRT provided a FAR of around one order of magnitude lower

than that of Bayes). In contrast, the nonparametric approach shows very similar performance for the Bayes (with uniform prior) and GLRT methods.

As expected, the clairvoyant¹ detector acted as a reference, by mostly representing an upper bound to the performance for a given α_{ff} value and background modeling approach. Although comparing parametric versus nonparametric approaches is not the scope of the work, clairvoyant curves show that for the smallest α_{ff} values, the parametric background model was better than the nonparametric model at low FAR, whereas the nonparametric performed better at high DR. However, for operational GLRT and Bayes algorithms, parametric methods obtained on these data generally gave better performance.

Summary metrics of FAR values obtained at some DR values, specifically FAR@DR={0.7, 0.8, 0.9}, are reported in Table II. FAR@DR metrics provide snapshots of the ROC curves for values of the detection threshold such that DR = {0.7, 0.8, 0.9} and confirm the behavior observed in the curves. Specifically, boldface FAR@DR values in the table represent those FAR@DR values for which (for the same background model and α_{ff}) the difference between GLRT and Bayes is at least an order of magnitude. This occurs for the parametric methods in favor of GLRT, mostly for the central α_{ff} values, whereas for the nonparametric case, GLRT and Bayes provide more similar results. AUC statistics are also reported in Table II, which summarizes the overall algorithm performance throughout the entire ROC curve. Boldface AUC values represent the highest AUC values between GLRT and Bayes for a given background model and target fill factor α_{ff} . While ROC curves and FAR@DR metrics reveal a generally better behavior of GLRT in the low FAR region, AUC values show that, in most cases and especially for the parametric case, Bayesian algorithms exhibit better overall performance, though the difference is small.

From this first set of experiments, the Bayes algorithm with uniform prior provided generally good performance, comparable to the GLRT. It should be noted that for this experiment, we have only employed the uniform prior and, thus, we have not fully explored the potential for leveraging a priori knowledge through the prior pdfs. In the following section, prior pdfs are adopted.

2) *Detection Performance Validation Results for the Real Subpixel Target Detection Scenario:* Before examining quantitative performance for the real subpixel target detection experiments, we first illustrate the algorithm performance by showing detection statistic images for both green (see Fig. 5) and yellow (see Fig. 6) target panels for (a) nonparametric methods and (b) parametric methods. In order to improve visual comparison and interpretation of the images, we rescaled the detection statistics using a log-ranked transform. Specifically, each rescaled statistic shows, for each image pixel z , the transformed value $-\log_{10}\{(\mathcal{R}[\Lambda(z)]/N)\}$,

¹In a small number of cases, the clairvoyant detector was slightly outperformed by another detector, even though the clairvoyant is in principle the optimal detector. But this principle strictly applies only if the data distribution exactly follows the background model. For real data, which we used here, this model may only approximate the true background distribution.

TABLE II
 FAR@DR = {0.7, 0.8, 0.9}* AND AUC** METRICS

FAR@DR = 0.7						
α_{ff}	NP-GLRT	NP-BAYES	NP-CLAIRVOYANT	P-GLRT	P-BAYES	P-CLAIRVOYANT
0.050	$2.63 \cdot 10^{-3}$	$2.71 \cdot 10^{-3}$	$5.35 \cdot 10^{-4}$	$8.88 \cdot 10^{-4}$	$4.39 \cdot 10^{-4}$	$8.59 \cdot 10^{-5}$
0.075	$2.77E \cdot 10^{-4}$	$3.25 \cdot 10^{-4}$	$3.82 \cdot 10^{-5}$	$1.91 \cdot 10^{-5}$	$1.43 \cdot 10^{-4}$	0
0.100	$1.53 \cdot 10^{-4}$	$1.62 \cdot 10^{-4}$	0	0	$5.73 \cdot 10^{-5}$	0
0.125	0	0	0	0	0	0
0.150	0	0	0	0	0	0
FAR@DR = 0.8						
α_{ff}	NP-GLRT	NP-BAYES	NP-CLAIRVOYANT	P-GLRT	P-BAYES	P-CLAIRVOYANT
0.050	$2.82 \cdot 10^{-3}$	$2.89 \cdot 10^{-3}$	$6.49 \cdot 10^{-4}$	$1.19 \cdot 10^{-3}$	$5.54 \cdot 10^{-4}$	$3.72 \cdot 10^{-4}$
0.075	$8.21 \cdot 10^{-4}$	$1.37 \cdot 10^{-3}$	$6.68 \cdot 10^{-5}$	$1.91 \cdot 10^{-5}$	$2.20 \cdot 10^{-4}$	0
0.100	$1.72 \cdot 10^{-4}$	$1.91 \cdot 10^{-4}$	0	0	$1.24 \cdot 10^{-4}$	0
0.125	0	0	0	0	0	0
0.150	0	0	0	0	0	0
FAR@DR = 0.9						
α_{ff}	NP-GLRT	NP-BAYES	NP-CLAIRVOYANT	P-GLRT	P-BAYES	P-CLAIRVOYANT
0.050	$3.07 \cdot 10^{-3}$	$3.12 \cdot 10^{-3}$	$7.64 \cdot 10^{-4}$	$1.52 \cdot 10^{-3}$	$1.04 \cdot 10^{-3}$	$1.30 \cdot 10^{-3}$
0.075	$1.97 \cdot 10^{-3}$	$2.13 \cdot 10^{-3}$	$1.34 \cdot 10^{-4}$	$7.64 \cdot 10^{-5}$	$3.82 \cdot 10^{-4}$	$3.82 \cdot 10^{-5}$
0.100	$2.39 \cdot 10^{-4}$	$2.77 \cdot 10^{-4}$	0	$1.91 \cdot 10^{-5}$	$2.20 \cdot 10^{-4}$	0
0.125	$9.55 \cdot 10^{-6}$	$9.55 \cdot 10^{-6}$	0	0	$9.55 \cdot 10^{-6}$	0
0.150	0	0	0	0	0	0
AUC						
α_{ff}	NP-GLRT	NP-BAYES	NP-CLAIRVOYANT	P-GLRT	P-BAYES	P-CLAIRVOYANT
0.050	0.996119	0.995974	0.998313	0.996243	0.999142	0.999506
0.075	0.999089	0.998997	0.999889	0.998340	0.999689	0.999911
0.100	0.999604	0.999604	0.999974	0.999006	0.999832	0.999970
0.125	0.999909	0.999917	0.999976	0.999395	0.999944	0.999986
0.150	0.999971	0.999976	0.999978	0.999592	0.999967	0.999992
Convex – AUC						
α_{ff}	NP-GLRT	NP-BAYES	NP-CLAIRVOYANT	P-GLRT	P-BAYES	P-CLAIRVOYANT
0.050	0.996129	0.995983	0.998322	0.996253	0.999149	0.999511
0.075	0.999099	0.999007	0.999895	0.998345	0.999694	0.999912
0.100	0.999613	0.999613	0.999974	0.999009	0.999837	0.999970
0.125	0.999915	0.999924	0.999976	0.999398	0.999945	0.999986
0.150	0.999974	0.999979	0.999978	0.999595	0.999968	0.999992

* boldface indicates values for which the difference between GLRT and Bayes (for same background model and α_{ff}) is around or higher than one order of magnitude.

** AUC values were evaluated under the ‘‘stair-step’’ experimental ROC curves (as represented in Fig. 4), whereas Convex – AUC values were evaluated under the ‘convexified’ ROC curves [29]. Boldface indicates the highest value between GLRT and Bayes (for same background model and α_{ff}).

where $\mathcal{R}[\Lambda(\mathbf{z})]$ is the ranking that the generic detector $\Lambda(\cdot)$ achieved at pixel \mathbf{z} compared to all other image pixels. More directly, a pixel value σ of the rescaled statistic means that 10^σ other pixels in the detection statistic image are weaker than it is. Because the transform is monotonic, it has no effect on the performance of the statistic (the ROC curves associated with a statistic and a *log-ranked* transform of the statistic are identical), but it makes it easier to visualize performance (since only the strongest pixels will be dark) and to compare the performance (since the two statistics being compared are scaled the same way). *Log-ranked* detection statistics for the case of green targets for nonparametric methods [see Fig. 5(a)] provide immediate visual evidence of the benefits of leveraging

the full potential of the Bayesian approach by crafting a suitable prior—while the targets stand out well for all the detection statistic images, more false alarms can be seen in the statistics of GLRT and uniform Bayes, compared to those of nonuniform Bayes. For parametric methods on green targets, on the other hand, *log-ranked* statistics exhibit fewer visual differences [see Fig. 5(b)]. For the yellow targets detected with nonparametric methods shown in Fig. 6(a), where the targets stand out comparably well in all statistics, fewer false alarms are evident in nonuniform Bayes statistics. Also, for the yellow targets detected with parametric methods shown in Fig. 6(b), Bayesian statistics are very similar one to another, whereas the GLRT statistic is apparently less sensitive to background

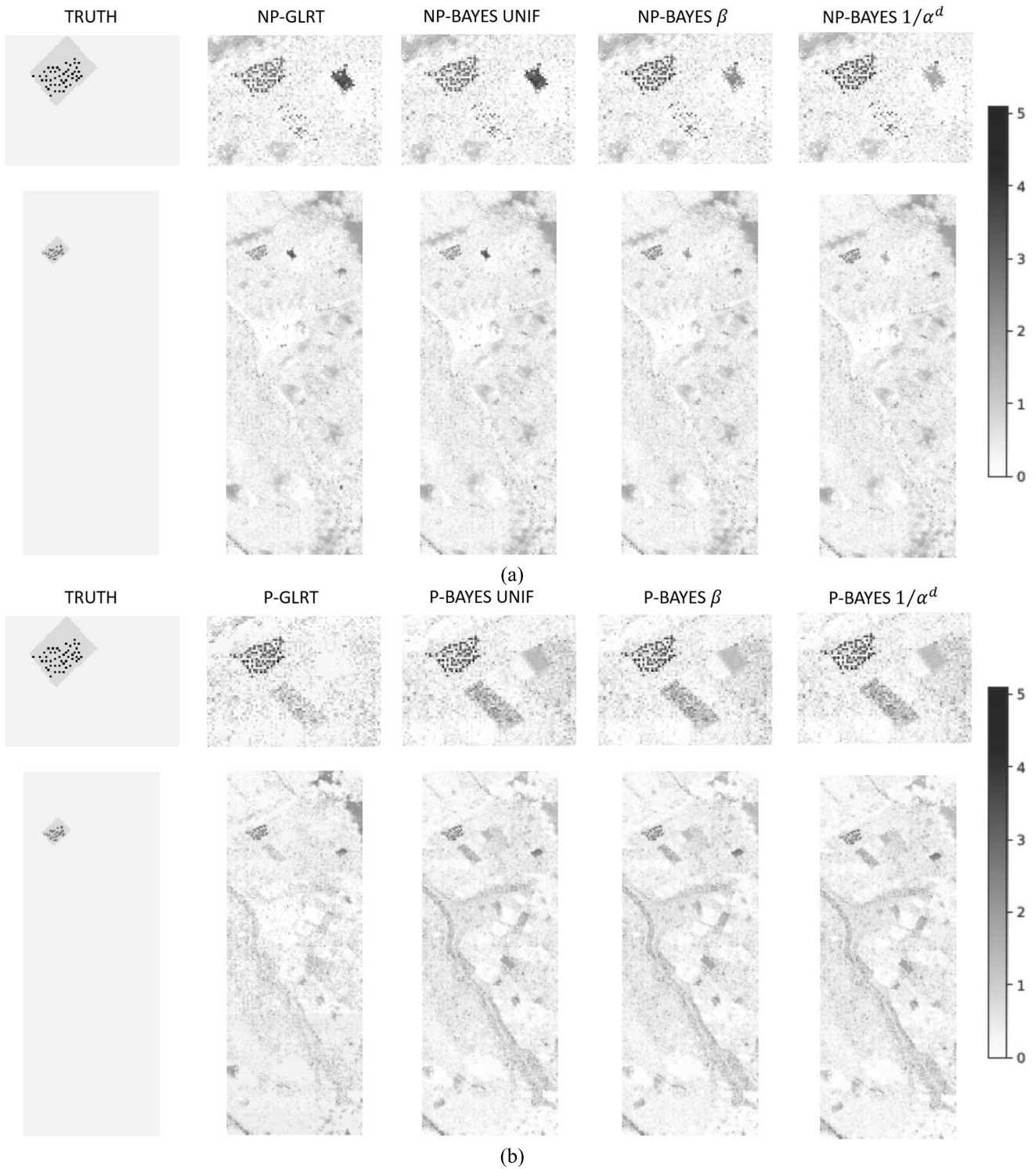


Fig. 5. Log-ranked statistics for the green target panels (both local target area and entire image are shown) for (a) nonparametric and (b) parametric algorithms. Ground truth is depicted on the left; target pixels are black, background pixels are light, and agnostic pixels (which are not counted as target or background because they are so close to the targets that the ground truth for these pixels is less certain) are gray. Log-ranked statistics correspond to a rescaling of detection statistics to a common distribution. The highest pixel in the detection statistic (ranking = 1) maps into $\log_{10}(N/1) \approx 5.02$, where $N = 250 \times 419$ is the number of pixels in the image. The tenth highest pixel in the detection statistic maps into $\log_{10}(N/10) \approx 4.02$, and the 100th pixel maps into $\log_{10}(N/100) \approx 3.02$. The weakest pixel (with the smallest detection statistic: ranking = N) maps into $\log_{10}(N/N) = 0$.

structures but also shows the targets that do not appear to stand out as well as for the Bayesian statistics.

The quantitative performance can be observed in terms of the ROC curves plotted in Fig. 7 for the green [(a) and (b)]

and yellow targets [(c) and (d)]. Results for the parametric algorithms are shown in Fig. 7(a) and (c), while those for nonparametric algorithms are shown in Fig. 7(b) and (d). As most of the figures clearly show, using a nonuniform prior

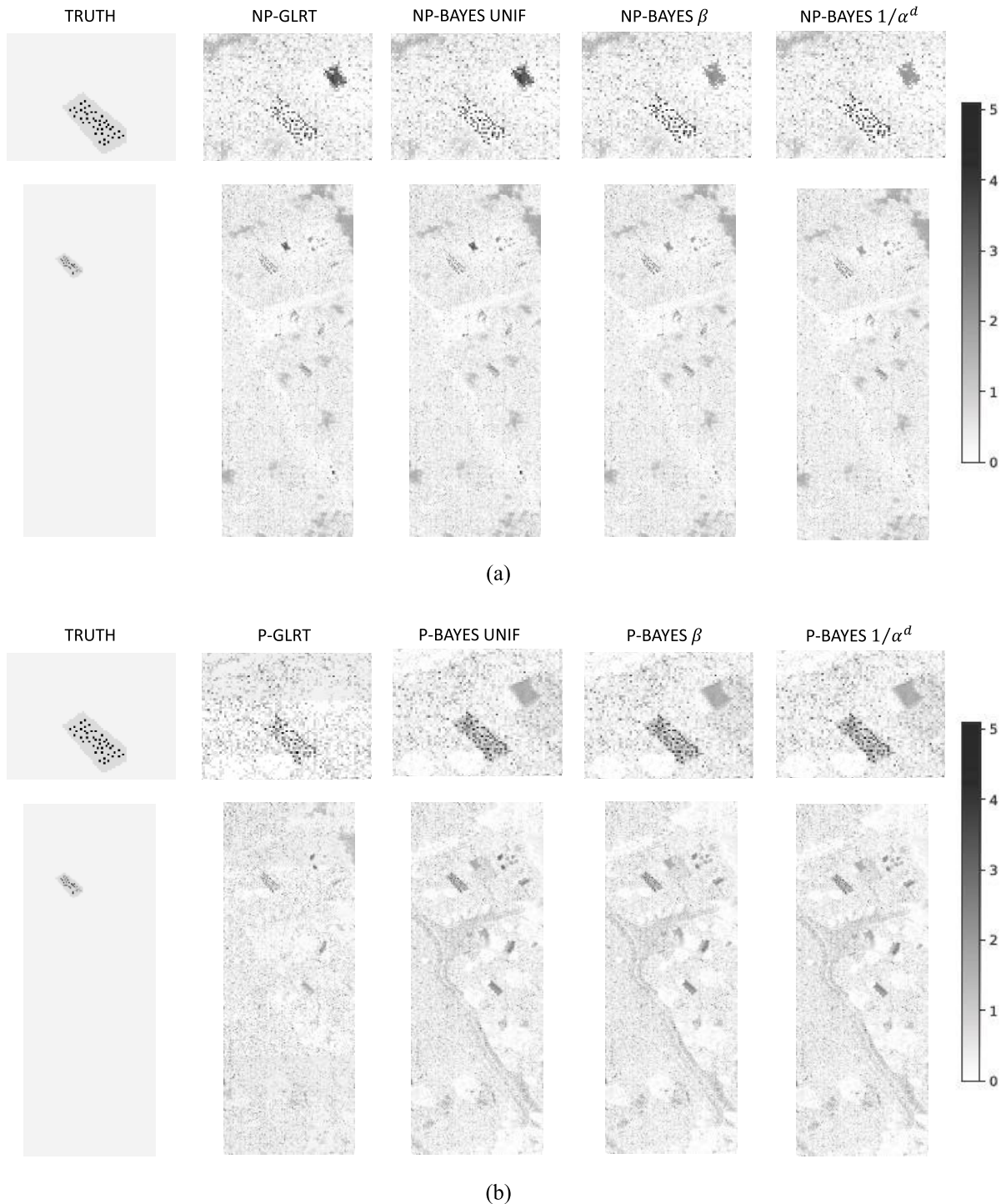


Fig. 6. Same as Fig. 5, but with the spectrum of the yellow panels (instead of the green) used as the target signature. As in Fig. 5, results are shown for (a) nonparametric and (b) parametric background models.

with the Bayesian algorithm does improve the performance, regardless of the background modeling approach. Except for the detection of the green panels with parametric algorithms [see Fig. 7(a)]—where all methods provide similarly good performance—the Bayesian method with nonuniform priors provided the best performance across the two types of targets and background modeling approaches. Usage of the uniform prior led to poorer performance than either of the nonuniform

priors. These results are confirmed by AUC and convex-AUC values reported in Table III for various algorithms and the two kinds of targets. In most cases, AUC values (of both kinds) for the Bayesian algorithms with nonuniform priors are notably higher than those of GLRT and Bayes with uniform prior.

It should be noted that the two different priors provided nearly identical performance. This was not surprising for two main reasons: first, because the same type of context-based

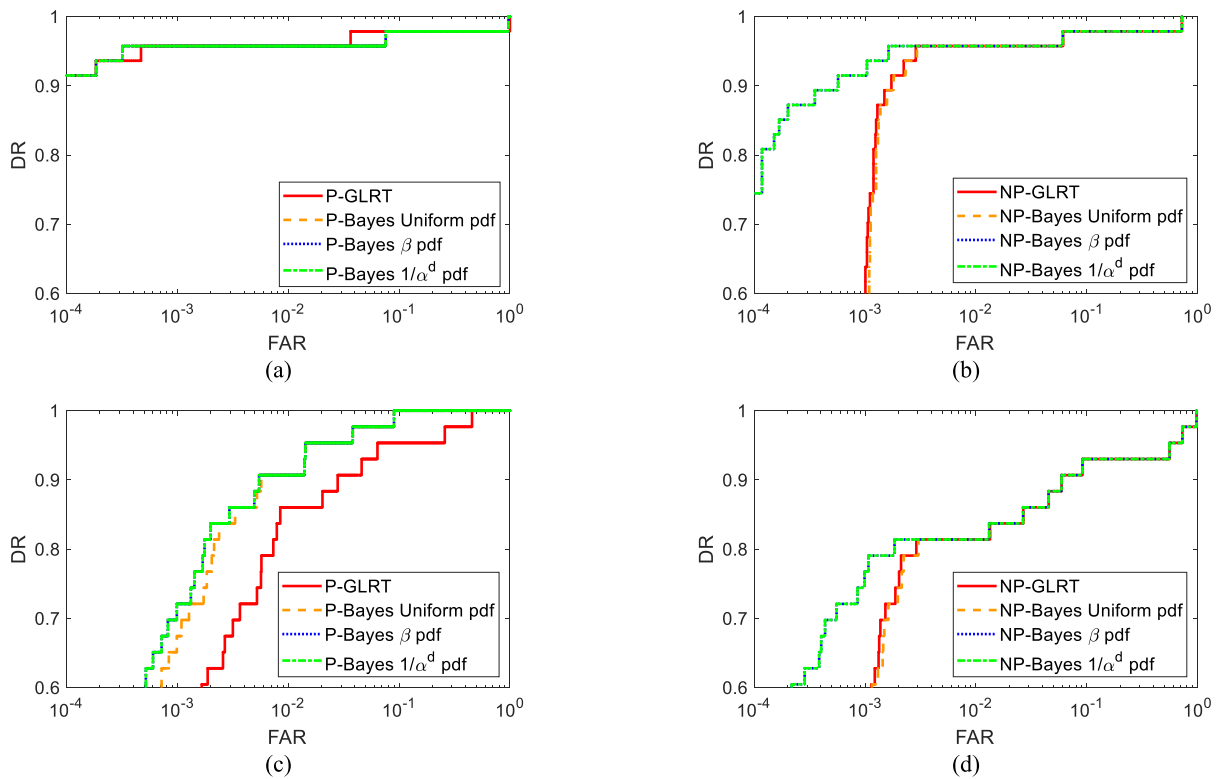


Fig. 7. ROC curves for (a) and (c) parametric and (b) and (d) nonparametric algorithms for detection of (a) and (b) green and (c) and (d) yellow panels, using four detectors. GLRT is shown as a red solid line, Bayesian with uniform prior is dashed orange, Bayesian with beta pdf is dotted blue, and Bayesian with $1/\alpha^d$ prior is dotted-dashed green.

TABLE III
AUC* METRICS

	AUC			
	NON-PARAMETRIC		PARAMETRIC	
	GREEN	YELLOW	GREEN	YELLOW
GLRT	0.982372	0.941160	0.983458	0.978236
BAYES UNIF	0.982283	0.941065	0.978542	0.995525
BAYES β	0.983254	0.941725	0.977849	0.995681
BAYES $1/\alpha^d$	0.983255	0.941727	0.977849	0.995681
	Convex – AUC			
	NON-PARAMETRIC		PARAMETRIC	
	GREEN	YELLOW	GREEN	YELLOW
GLRT	0.990284	0.957656	0.988569	0.983649
BAYES UNIF	0.990238	0.957622	0.987747	0.996699
BAYES β	0.990933	0.958178	0.988117	0.996854
BAYES $1/\alpha^d$	0.990933	0.958179	0.988117	0.996854

*AUC values were evaluated under the “stair-step” experimental ROC curves (as represented in Fig. 5), whereas *Convex* – AUC values were evaluated under the ‘convexified’ ROC curves [29]. Boldface indicates the highest values (for a same background model and target kind).

a priori information (namely, the small fill factors were expected) was exploited to craft both prior pdfs; and second, because a higher number of points than $n_{\text{pts}} = 6$ for numerical integration is likely to be needed to appreciate the differences in two pdfs with such strong weights toward low α values.

More importantly, this outcome shows that Bayes is worth applying even if the “optimal” prior cannot be derived—as long as the available a priori information is injected into the process, even “rough” priors expressing the importance associated with small fill factors can benefit the final performance.

3) *Insights Into the Integration Scheme*: Results of experiments exploring insights into the integration scheme for Bayesian algorithms are shown in Fig. 8(a)–(f).

Average log-integral curves are shown in Fig. 8(a)–(d) where the average natural logarithm of the integral at the numerator of $\Lambda_{\text{Bayes}}^{\text{RTM}}$ in (6) is plotted against the number of points n_{pts} . The integrals are averaged over 150 randomly selected pixels extracted from the background H_0 image [see Fig. 8(a) and (b)] and the target-implemented H_1 image [see Fig. 8(c) and (d)]. Results for parametric algorithms are shown in Fig. 8(a) and (c), while those for nonparametric algorithms are shown in Fig. 8(b) and (d). The MP and GL rules are identified by the blue and red colors, respectively.

As can be observed, the values of the integrals for parametric and nonparametric algorithms are different; this is not a serious concern, however, because it is the *ratio* of likelihoods that defines the detectors. What really matters here is that the trends of the curves are similar. More specifically, regardless of the background modeling approach, the MP and GL integration schemes both produced integrals that converged to the same value and did so both for background and target-implemented images. In general, though this happened to a greater extent for the parametric case, the GL strategy led to a faster convergence than MP. Also, regardless of the

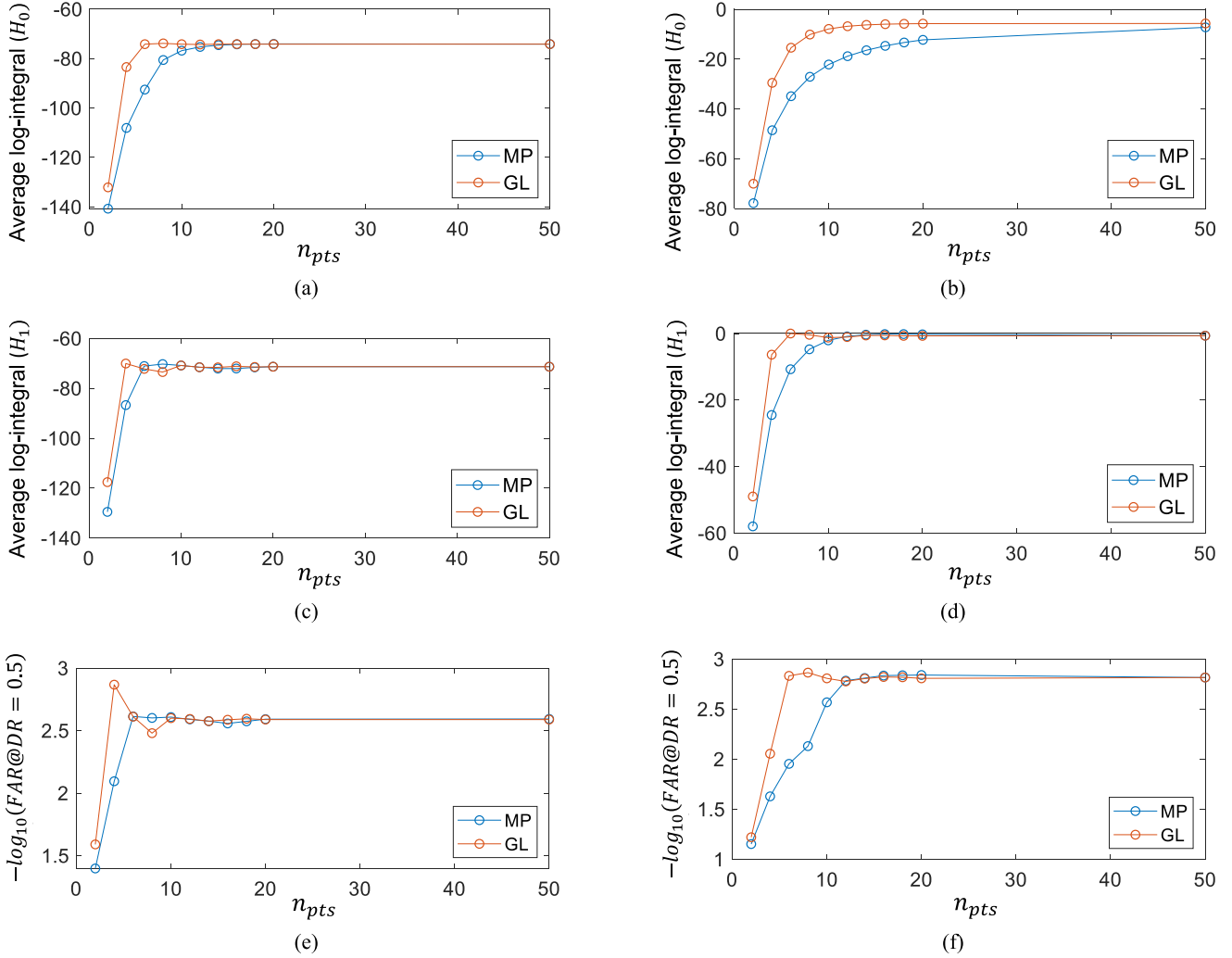


Fig. 8. (a)–(d) Logarithm of the integral in the numerator of $\Lambda_{\text{Bayes}}^{\text{RTM}}$, averaged over randomly extracted pixels from the background H_0 image in (a) and (b) and the target-implanted H_1 image in (c) and (d)—for (a) and (c) parametric and (b) and (d) nonparametric Bayesian algorithms with uniform prior, evaluated with both MP (in blue) and GL (in red) integration rules and plotted with respect to the number of points n_{pts} . (e) and (f) Negative logarithm of FAR for a DR of 0.5 is evaluated over all pixels of the matched image-pair, based on (e) parametric and (f) nonparametric Bayesian algorithms with uniform prior. Both MP (in blue) and GL (in red) integration rules are used, and results are plotted with respect to the number of points n_{pts} .

scheme, the convergence was faster for the nonparametric algorithms. In general, all log-integral curves exhibited a mostly monotonic increase with n_{pts} , although some slight oscillations can be observed for small n_{pts} in the GL-based schemes, due to the nonuniform point spacing that changes with n_{pts} .

As said before, log integrals should be observed coupled with some detection performance curves, as those shown in Fig. 8(e) and (f). Here, the negative base-ten logarithm of the FAR obtained for a DR of 0.5 ($-\log_{10}(\text{FAR}@DR=0.5)$), which is evaluated from all the pixels in the matched image pair, is plotted versus n_{pts} as summary performance measure (the higher the better) for parametric [see Fig. 8(e)] and nonparametric [see Fig. 8(f)] Bayesian algorithms. Also here, both MP- and GL-based integration schemes led to curves converging to the same $-\log_{10}(\text{FAR}@DR = 0.5)$ value. In general, the nonparametric algorithms exhibited a faster convergence than the parametric ones. Overall, the GL-based scheme provided

a faster convergence than MP, with $n_{\text{pts}} = 6$ almost providing the convergence value for both nonparametric and parametric cases. It should also be noted that the GL-based scheme allowed for higher $-\log_{10}(\text{FAR}@DR=0.5)$ values than those obtained with the MP-based scheme.

Accounting for both log-integral and performance curves, we decided to adopt throughout the experiments the GL-based scheme, which, with its faster convergence and generally lower FAR, allowed better performance to be obtained with smaller n_{pts} values than the MP-based scheme. We chose $n_{\text{pts}} = 6$ as a tradeoff between performance and computational efficiency.

IV. CONCLUSION

Bayesian algorithms were developed and explored for solid subpixel target detection in hyperspectral imagery. These algorithms provide admissible alternatives to GLRT-based solutions to the composite hypothesis testing problem. To fully

benefit from the flexibility offered by the Bayesian approach, we have employed numerical integration schemes that maintain the admissibility property but do not depend on simplifying or approximating the likelihood expression, and do not limit the choice of prior based on considerations of closed-form integrability. This framework gives the user freedom to choose the prior, based either upon available a priori contextual information or upon whatever relevance the user chooses to attach to given target sizes.

We have specifically developed Bayesian detectors based on the RTM for a parametric (elliptically contoured multivariate- t distribution) and a nonparametric (kernel-based data-driven distribution) background model. Two strategies for numerical integration have been analyzed and compared in terms of both their accuracy in approximating the integral as well as their performance in detecting the targets.

Experimental results over real hyperspectral images featuring both synthetic and real subpixel target detection scenarios have shown that usage of a suitable prior pdf indeed boosts the Bayesian detection performance. In fact, whereas usage of noninformative (i.e., uniform) priors has provided Bayesian performance similar to that of GLRT, injecting context-based information through the priors has been shown to considerably improve the performance and to a major extent with the nonparametric approach. Very similar performance improvements were obtained using quite different prior pdf functions (a proper beta-function prior and an “improper” power-law prior), showing that even rough expressions of prior information are still effective and can outperform a noninformative prior. This great flexibility shown by Bayesian algorithms from a target detection performance perspective provides impetus to further research. It is our view, in fact, that substantial benefits may be obtained by suitably engineering prior functions to optimize the desired performance criteria.

REFERENCES

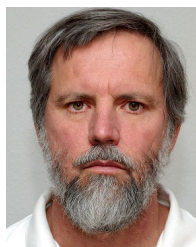
- [1] D. Manolakis, R. Lockwood, and T. Cooley, *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
- [3] J. Theiler, A. Ziemann, S. Matteoli, and M. Diani, “Spectral variability of remotely sensed target materials: Causes, models, and strategies for mitigation and robust exploitation,” *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 8–30, Jun. 2019.
- [4] J. Chen, “Penalized likelihood-ratio test for finite mixture models with multinomial observations,” *Can. J. Statist.*, vol. 26, no. 4, pp. 583–599, Dec. 1998.
- [5] A. Schaum, “Continuum fusion: A theory of inference, with applications to hyperspectral detection,” *Opt. Exp.*, vol. 18, no. 8, pp. 8171–8181, 2010.
- [6] J. Theiler, “Confusion and clairvoyance: Some remarks on the composite hypothesis testing problem,” *Proc. SPIE*, vol. 8390, May 2012, Art. no. 839003.
- [7] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. New York, NY, USA: Springer, 2005.
- [8] A. Schaum, “Hyperspectral target detection using a Bayesian likelihood ratio test,” in *Proc. IEEE Proc. Aerosp. Conf.*, Mar. 2002, pp. 1537–1540.
- [9] D. Manolakis et al., “On clairvoyant and universal tests for hyperspectral target detection,” *Proc. SPIE*, vol. 11130, Sep. 2019, Art. no. 1113004.
- [10] J. Theiler, S. Matteoli, and A. Ziemann, “Bayesian detection of solid subpixel targets,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2021, pp. 3213–3216.
- [11] S. Matteoli, M. Diani, and G. Corsini, “Bayesian non-parametric detector based on the replacement model,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2022, pp. 871–874.
- [12] J. Theiler, “Veritas: An admissible detector for targets of unknown strength,” *Proc. SPIE*, vol. 11727, Apr. 2021, Art. no. 117270B.
- [13] S. Matteoli, M. Diani, and J. Theiler, “An overview of background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2317–2336, Jun. 2014.
- [14] D. B. Marden and D. Manolakis, “Using elliptically contoured distributions to model hyperspectral imaging data and generate statistically similar synthetic data,” *Proc. SPIE*, vol. 5425, pp. 558–572, Aug. 2004.
- [15] J. Theiler and B. R. Foy, “EC-GLRT: Detecting weak plumes in non-Gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2008, pp. 221–224.
- [16] J. Theiler, B. Zimmer, and A. Ziemann, “Closed-form detector for solid sub-pixel targets in multivariate T-distributed background clutter,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 2773–2776.
- [17] W. Silverman, *Density Estimation for Statistics and Data Analysis*. London, U.K.: Chapman & Hall, 1986.
- [18] G. R. Terrell and D. W. Scott, “Variable kernel density estimation,” *Ann. Statist.*, vol. 20, no. 3, pp. 1236–1265, Sep. 1992.
- [19] S. Matteoli, M. Diani, and G. Corsini, “Closed-form nonparametric GLRT detector for subpixel targets in hyperspectral images,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1568–1581, Apr. 2020.
- [20] S. Matteoli, T. Veracini, M. Diani, and G. Corsini, “Background density nonparametric estimation with data-adaptive bandwidths for the detection of anomalies in multi-hyperspectral imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 163–167, Jan. 2014.
- [21] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. New York, NY, USA: Dover, 1972.
- [22] A. Schaum and A. Stocker, “Spectrally selective target detection,” in *Proc. ISSSR*, 1997, p. 23.
- [23] Y. Wang, X. Chen, F. Wang, M. Song, and C. Yu, “Meta-learning based hyperspectral target detection using Siamese network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5527913.
- [24] J. Zeng and Q. Wang, “Sparse tensor model-based spectral angle detector for hyperspectral target detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5539315.
- [25] X. Zhao, W. Li, C. Zhao, and R. Tao, “Hyperspectral target detection based on weighted Cauchy distance graph and local adaptive collaborative representation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5527313.
- [26] X. Wang et al., “A double dictionary-based nonlinear representation model for hyperspectral subpixel target detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5524516.
- [27] J. Theiler, C. Scovel, B. Wohlberg, and B. R. Foy, “Elliptically-contoured distributions for anomalous change detection in hyperspectral imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 271–275, Apr. 2010.
- [28] K. Fukunaga and L. D. Hostetler, “Optimization of k-nearest neighbor density estimates,” *IEEE Trans. Inf. Theory*, vol. MTT-19, no. 3, pp. 320–326, May 1973.
- [29] C. Chang, “An effective evaluation tool for hyperspectral target detection: 3D receiver operating characteristic curve analysis,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5131–5153, Jun. 2021.
- [30] M. Bar-Tal and S. R. Rotman, “Performance measurement in point target detection,” *Infr. Phys. Technol.*, vol. 37, pp. 231–238, Jan. 1996.
- [31] J. Theiler, “Matched-pair machine learning,” *Technometrics*, vol. 55, no. 4, pp. 536–547, Nov. 2013.
- [32] M. Giannandrea et al., “The SHARE 2012 data campaign,” *Proc. SPIE*, vol. 8743, May 2013, Art. no. 87430F.
- [33] *SHARE 2012: SpecTIR Hyperspectral Airborne Experiment 2012*. Accessed: Jul. 11, 2023. [Online]. Available: <https://www.rit.edu/cos/share2012/>
- [34] J. P. Kerekes, K. Ludgate, A. M. Giannandrea, N. G. Raqueno, and D. S. Goldberg, “SHARE 2012: Subpixel detection and unmixing experiments,” *Proc. SPIE*, vol. 8743, May 2013, Art. no. 87430H.
- [35] T. Sun and J. Kerekes, “An analysis task comparison of uncorrected vs. geo-registered airborne hyperspectral imagery,” *Proc. SPIE*, vol. 9472, Jan. 2015, Art. no. 94720I.



Stefania Matteoli (Senior Member, IEEE) received the B.S. and M.S. (cum laude) degrees in telecommunications engineering and the Ph.D. degree in remote sensing from the University of Pisa, Pisa, Italy, in 2003, 2006, and 2010, respectively.

From May 2008 to October 2008, she was a Visiting Student with the Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York. From January 2010 to December 2016, she was a first Post-Doctoral Fellow and then a temporary Researcher with the Department of Information Engineering, University of Pisa. She is currently a Senior Researcher with the Institute of Electronics, Information Engineering and Telecommunications, National Research Council of Italy, Pisa. Her main research interests include signal and image processing, anomaly and spectral signature-based target detection methodologies for hyperspectral remotely sensed images, radiation transfer modeling in the atmosphere and within sea water, spectroscopic signal processing in food safety applications, and fluorescence LIDAR signal processing for underwater object detection and recognition.

Dr. Matteoli is an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and the *Journal of Applied Remote Sensing* (SPIE). She is a referee for several international journals. In 2014, she was a recipient of the “SPIE Remote Sensing Europe Best Student Paper Award” within the “Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions” Conference, for the paper titled “Design and Validation of Object Recognition Methodologies for Underwater Fluorescence LIDAR Applications.”



James Theiler (Member, IEEE) received the Ph.D. degree in physics from the California Institute of Technology, Pasadena, CA, USA, in 1987. His thesis focused on the statistical and computational aspects of identifying chaos in time series data.

He subsequently held appointments with the University of California at San Diego, San Diego, CA, USA; the MIT Lincoln Laboratory, Lexington, MA, USA; the Los Alamos National Laboratory, Los Alamos, NM, USA; and the Santa Fe Institute, Santa Fe, NM, USA. In 1994, he joined the Technical Staff at Los Alamos, and in 2005, he was named a Laboratory Fellow. He formally retired in February 2020, but still holds a position as a Guest Scientist at Los Alamos and continues to cultivate interests in statistical modeling, machine learning, image processing, and remote sensing.