Robust adaptive target detection in hyperspectral imaging

François Vincent*, Olivier Besson

ISAE-SUPAERO, 10 Avenue Edouard Belin, Toulouse 31055, France

Keywords:
Hyperspectral imagery
Robust Detection
GLRT
AMF

A B S T R A C T

One of the main issue in detecting a target from an hyperspectral image relies on properly identifying the background. Many assumptions about its distribution can be advocated, even if the Gaussian hypothesis prevails. Nevertheless, the huge majority of the resulting detection schemes assume that the background distribution remains the same whether the target is present or not. In practice, because of the spectral variability of the target and the non-linear mixing with the background radiance, this hypothesis is not strictly true. In this paper, we consider that an unknown background mismatch exists between the two hypotheses. Under the assumption that this mismatch is small, we derive an approximation of the Likelihood Ratio for the problem at hand. This general formulation is then applied to the case of Gaussian distributed background, leading to a robust Adaptive Matched Filter. The behaviour of this new detector is analysed and compared to popular detectors. Numerical simulations, based on real data, show the possible improvement in case of target signature mismatch.

1. Introduction

An hyperspectral image is a two-dimensional map, where each pixel is composed of hundreds of spectral bands. This spectral information allows to characterize the different materials present in the image and hence hyperspectral imaging has encountered a large field of applications, ranging from remote sensing to medicine [1–8]. One of these applications consists in detecting a target whose spectral signature could be known from laboratory experiments, or unknown. In this latter case, one talks of anomaly detection. In this paper, we will focus on the first case, namely target detection.

The difficulty of this problem lies in the fact that the signature of interest (SoI) $\mathbf{t}$ is buried in a background $\mathbf{b}$ with unknown statistics. Even if its distribution were known, the parameters describing it (for instance mean and covariance matrix) are not known and must be estimated from the available data. Consequently, detection of the SoI in a pixel under test (PUT) $\mathbf{y}$ requires using other pixels (so called training samples) to learn the background present in the PUT. It has to be noticed that in the hyperspectral domain, we deal with real positive data leading to non zero mean signals.

Certainly due to its easy handling properties, the Gaussian assumption prevails to model the background. In this context, many algorithms have been developed, such as the Adaptive Matched Filter (AMF) [9], Kelly's Generalized Likelihood Ratio Test (GLRT) [10], the Orthogonal Subspace Projection (OSP) [11], or the Adaptive Coherent/Cosine Estimator (ACE) [12], to name the most popular ones. However, with real hyperspectral data, Gaussian distributions rarely occur [13–16], and more representative background distributions have been considered. One of the most common is the Elliptically Contoured (EC) t-distributed model that allows to extend the Gaussian distribution to a broader class of probability density functions (p.d.f.). Different detectors have been derived under such an hypothesis, as the EC-GLRT [17] for instance.

Nevertheless, although considering a broad class of distribution to model the background, the huge majority of the decision tests rely on the assumption that the background remains the same under the two hypotheses.

$$\begin{aligned} H_0 &: \mathbf{y} = \mathbf{b} \\ H_1 &: \mathbf{y} = \alpha\mathbf{t} + \mathbf{b} \end{aligned} \tag{1}$$

where α is the so-called fill-factor.

Hence, if considering the SoI and α as deterministic, which is a widespread assumption, this model amounts to consider the same distribution under the two hypotheses. The only difference between the two hypotheses is a mean shift.

Unfortunately, this strong assumption is usually not met in real hyperspectral detection schemes. Indeed, the first reason is related to the data acquisition itself. The radiance resulting from the mixture of a background and a target are driven by non-linear effects [18]. These effects are due to multi-reflections or masking effects. Thereby, the additive model from Eq. (1), is over simpli-

* Corresponding author.
E-mail addresses: francois.vincent@isae-supaero.fr, francois.vincent@isae.fr (F. Vincent), olivier.besson@isae-supaero.fr (O. Besson).

https://doi.org/10.1016/j.sigpro.2020.107905

fied. A more representative model, namely the so-called replacement model, can be used [1]. This model assumes that the target, if present, replaces a part of the background, as it will mask an area proportional to its fill-factor α. Some detectors have been derived under such a model, as the FTMF [19] or ACUTE [20]. A more general model, relaxing the unitary constraints on the amplitudes, only assumes a linear mixing between the background and the target [21]. But even if these two models are more representative, they do not take into account the multiple light reflections conducting to non-linear mixing, and consequently a more complicated background behaviour when the target exists.

The second main reason conducting to a change in the distribution between the two hypotheses is related to mismatches in the target SoI. Indeed, this so-called spectral variability is due to changes in the conditions between the lab measurements and the data acquisition [9,22]. In fact, the image is acquired by a different instrument and at a larger distance from the scene. Then, the raw radiance measurements have to be converted into reflectance data using a so-called radiometric compensation [23]. Different techniques can be used [24,25], nevertheless all these compensation techniques are based on models inducing possible SoI mismatches. Then, when assuming the SoI as deterministic and perfectly known, this spectral variability results in a noise measurement increase. Again, the consequence is that the PUT exhibits a different distribution beyond a simple mean shift.

Thereby, in this paper, we consider a more general detection model that may include a large class of mismatches, including those stated here-above. In this model, the background distribution may vary between the two hypotheses. More precisely we assume that when the target is present, the background undergoes a possible perturbation with unknown statistics:

$$\begin{aligned} H_0 &: \mathbf{y} = \mathbf{b} \\ H_1 &: \mathbf{y} = \alpha\mathbf{t} + \mathbf{b} + \Delta\mathbf{b} \end{aligned} \tag{2}$$

where $\Delta\mathbf{b}$ is the mismatch.

Conventionally, the distribution of $\mathbf{b}$ can be estimated from training samples $\mathbf{x}_k$. On the other hand, $\Delta\mathbf{b}$ has an unknown distribution, possibly linked to α. This unknown perturbation can model a large class of possible mismatches induced by non-linear mixtures between the target and the background or from target signature errors.

Under the hypothesis that $\Delta\mathbf{b}$ is small compared to $\mathbf{b}$, we derive a robust Likelihood Ratio (LR), and show that it can be written as the sum of the standard LR and a correction factor. When applied to the popular case where $\mathbf{b}$ is supposed to be Gaussian distributed, we derive an approximation of the GLRT. This robust detector is shown to be the popular AMF, corrected by a simple additive term.

The remaining of this paper is organized as follows. We first derive the Likelihood Ratio for the general case of any background distribution in section 2. Then, we focus on the popular Gaussian case in Section 3, where we propose a simple additive correction for the AMF detector to improve robustness. In Section 4, we compare the behaviour of the proposed robust AMF to popular detectors, thanks to their threshold limit plotted in a Matched Filter Residual (MFR) diagram. To finish, we compare these detectors using a Monte-Carlo simulation based on real data, in Section 5. Finally concluding remarks end this paper in Section 6.

2. Robust likelihood ratio

As already said in the introduction, we assume that the background under H_1 writes $\mathbf{b}_1 = \mathbf{b} + \Delta\mathbf{b}$, where the probability density function (pdf) of $\mathbf{b}$, $p_b(.)$ is supposed to be known, and its parameters can be estimated from the training samples, that is the standard hypothesis. Hence, for any $\Delta\mathbf{b}$ pdf, $p_\Delta(.)$, the pdf of $\mathbf{b}_1$ can be written as

$$p_{b_1}(\mathbf{u}) = (p_b * p_\Delta)(\mathbf{u}) = \int p_\Delta(\mathbf{z}) p_b(\mathbf{u}-\mathbf{z}) d\mathbf{z} = \mathbb{E}_\Delta[p_b(\mathbf{u}-\mathbf{z})] \tag{3}$$

Now, using the approximation derived in the Appendix, for any given $\mathbf{u}$, we have

$$\mathbb{E}_\Delta[p_b(\mathbf{u}-\mathbf{z})] \simeq p_b(\mathbf{u}-\boldsymbol{\mu}_\Delta) + \frac{1}{2}\mathrm{Tr}\left\{ \frac{\partial^2 p_b(\mathbf{z})}{\partial\mathbf{z}\partial\mathbf{z}^T}|_{\mathbf{u}-\boldsymbol{\mu}_\Delta} \mathbf{C}_\Delta \right\} \tag{4}$$

where Tr{} stands for the trace of the matrix between braces, $\boldsymbol{\mu}_\Delta$ and $\mathbf{C}_\Delta$ denote respectively the mean and covariance matrix of the unknown background mismatch $\Delta\mathbf{b}$, and $[\frac{\partial^2 p_b(\mathbf{z})}{\partial\mathbf{z}\partial\mathbf{z}^T}|_{\mathbf{u}-\boldsymbol{\mu}_\Delta}]$ denotes the Hessian of $p_b(\mathbf{z})$ evaluated at $(\mathbf{u}-\boldsymbol{\mu}_\Delta)$. It should be pointed out that this approximation does not require complete knowledge of the pdf of $\Delta\mathbf{b}$, but only of its mean and its covariance matrix, which is an appealing feature.

Hence, for any PUT measurement $\mathbf{y}$, the Likelihood Ratio (LR) associated with the detection problem from Eq. (2) can be written in the following form, as soon as the background mismatch $\Delta\mathbf{b}$ is small:

$$LR_r = \frac{p_{\mathbf{b}_1}(\mathbf{y}-\alpha\mathbf{t})}{p_\mathbf{b}(\mathbf{y})} \simeq \frac{p_\mathbf{b}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu}_\Delta)}{p_\mathbf{b}(\mathbf{y})} + \frac{\mathrm{Tr}\{\frac{\partial^2 p_b(\mathbf{z})}{\partial\mathbf{z}\partial\mathbf{z}^T}|_{\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu}_\Delta}\mathbf{C}_\Delta\}}{2p_\mathbf{b}(\mathbf{y})} \tag{5}$$

$$= \frac{p_\mathbf{b}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu}_\Delta)}{p_\mathbf{b}(\mathbf{y})}\left[1 + \frac{1}{2}\frac{\mathrm{Tr}\{\frac{\partial^2 p_b(\mathbf{z})}{\partial\mathbf{z}\partial\mathbf{z}^T}|_{\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu}_\Delta}\mathbf{C}_\Delta\}}{p_\mathbf{b}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu}_\Delta)}\right] \tag{6}$$

Moreover, $\Delta\mathbf{b}$ being some perturbation around $\mathbf{b}$, we assume that $\boldsymbol{\mu}_\Delta = 0$, so that we simply have

$$LR_r \simeq LR\left[1 + \frac{1}{2}\mathrm{Tr}\{\frac{\partial^2 p_b(\mathbf{z})}{\partial\mathbf{z}\partial\mathbf{z}^T}|_{\mathbf{y}-\alpha\mathbf{t}}\mathbf{C}_\Delta\}\right] \tag{7}$$

where $LR = \frac{p_\mathbf{b}(\mathbf{y}-\alpha\mathbf{t})}{p_\mathbf{b}(\mathbf{y})}$ is the LR in the standard case, i.e. when there is no background mismatch ($\Delta\mathbf{b} = 0$). Hence, we can see that for any zero-mean mismatch distribution, we can approximate the modified LR by the sum of the LR without mismatch and a corrective term, depending on the pdf of $\mathbf{b}$ and the covariance matrix of $\Delta\mathbf{b}$.

3. Robust adaptive matched filter

In this section, we now investigate the case where the background distribution under H_0 is Gaussian, $\mathbf{b} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$. Let us start with the nominal case, to recall the AMF derivation.

If no mismatch is present ($\mathbf{b}_1 = \mathbf{b}$), the standard LR writes

$$LR^G = \frac{p_b(\mathbf{y}-\alpha\mathbf{t})}{p_b(\mathbf{y})} = \frac{e^{-\frac{1}{2}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu})}}{e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu})}} \tag{8}$$

Assuming that the background parameters are known from the training samples, the only unknown parameter is α, whose Maximum Likelihood (ML) writes $\hat{\alpha} = \frac{\mathbf{t}^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{t}}$. Then, the logarithm of the Generalized LR (GLR) is shown to be:

$$\log GLR^G = \log\left[\frac{p_b(\mathbf{y}-\hat{\alpha}\mathbf{t})}{p_b(\mathbf{y})}\right] = \frac{1}{2}\frac{|\mathbf{t}^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu})|^2}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{t}} = \frac{1}{2}AMF \tag{9}$$

where we recognize the popular $AMF = \frac{|\mathbf{t}^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu})|^2}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{t}}$.

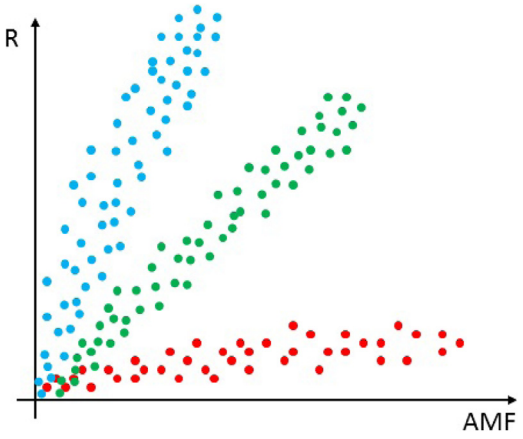

Fig. 1. Typical MFR plot.

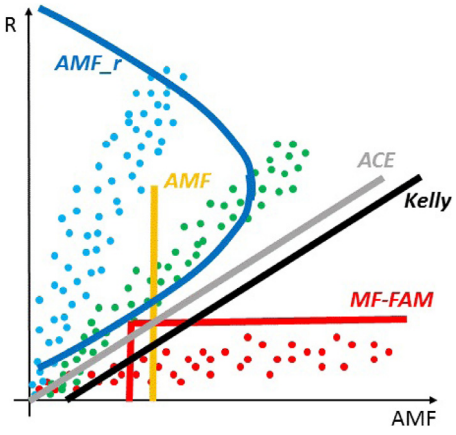

Fig. 2. Typical MFR plot.

Now, in the mismatched case ($\mathbf{b}_1 \neq \mathbf{b}$), the Hessian in Eq. (7) writes

$$\frac{\partial^2 p_b(\mathbf{z})}{\partial \mathbf{z} \partial \mathbf{z}^T}|_{\mathbf{y}-\alpha\mathbf{t}} = \mathbf{C}^{-\frac{1}{2}}\left[\mathbf{C}^{-\frac{1}{2}}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu})(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu})^T\mathbf{C}^{-\frac{1}{2}} - \mathbf{I}\right]\mathbf{C}^{-\frac{1}{2}} \times p_b(\mathbf{y}-\alpha\mathbf{t}) \tag{10}$$

so that the modified LR becomes

$$LR_r^G \simeq LR^G \cdot\left[1+\frac{1}{2}\mathrm{Tr}\left\{\left(\mathbf{C}^{-\frac{1}{2}}(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu})(\mathbf{y}-\alpha\mathbf{t}-\boldsymbol{\mu})^T\mathbf{C}^{-\frac{1}{2}} - \mathbf{I}\right)\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}\right\}\right] \tag{11}$$

First, we have to estimate α. As the background mismatch under H_1 is supposed to be small, we can use the ML estimation of α obtained here above in the nominal case, as a first order approximation, namely $\hat{\alpha} = \frac{\mathbf{t}^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{t}}$. Hence, taking the logarithm and substituting α, the GLR writes:

$$\log GLR_r^G \simeq \log GLR^G + \log \cdot\left[1+\frac{1}{2}\mathrm{Tr}\left\{\left(\mathbf{C}^{-\frac{1}{2}}(\mathbf{y}-\hat{\alpha}\mathbf{t}-\boldsymbol{\mu})(\mathbf{y}-\hat{\alpha}\mathbf{t}-\boldsymbol{\mu})^T\mathbf{C}^{-\frac{1}{2}} - \mathbf{I}\right)\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}\right\}\right] \tag{12}$$

or, thanks to Eq. (9)

$$2\log GLR_r^G \simeq AMF + 2\log\left[1+\frac{1}{2}\mathrm{Tr}\left\{\left(\mathbf{u}\mathbf{u}^T - \mathbf{I}\right)\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}\right\}\right] \tag{13}$$

where $\mathbf{u} = \mathbf{P}^{\perp}_{\mathbf{t}_c}\mathbf{y}_c$, with $\mathbf{t}_c = \mathbf{C}^{-\frac{1}{2}}\mathbf{t}$ and $\mathbf{y}_c = \mathbf{C}^{-\frac{1}{2}}(\mathbf{y}-\boldsymbol{\mu})$, the whitened versions of $\mathbf{t}$ and $(\mathbf{y}-\boldsymbol{\mu})$ respectively.

Therefore, we can see that assuming a Gaussian background under H_0, and for small mismatches under H_1, the GLR is the popular AMF, plus a correction. This additive term tends to measure the gap between the supposed background covariance matrix $\mathbf{C}$ and the actual one under H_1. Indeed, $\mathbf{u}$ is an estimation of the background under H_1, whitened by the supposed covariance matrix $\mathbf{C}$, so that $(\mathbf{u}\mathbf{u}^T - \mathbf{I})$ is a kind of mismatch measurement. In the same way, $\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}$ is also a normalized measurement of this gap.

Unfortunately, this expression of the correction is difficult to use in practice, except if one has an idea of $\mathbf{C}_\Delta$. In order to address this issue, we now derive an approximation of this correction in order to get a more convenient expression.

The correction depends on $T = \mathrm{Tr}\{\left(\mathbf{u}\mathbf{u}^T - \mathbf{I}\right)\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}\}$ that can be written as follows

$$T = \mathrm{Tr}\{\tilde{\mathbf{C}}_\Delta\}\left[\frac{\mathbf{u}^T\tilde{\mathbf{C}}_\Delta\mathbf{u}}{\mathrm{Tr}\{\tilde{\mathbf{C}}_\Delta\}} - 1\right] \tag{14}$$

Fig. 3. Complete RGB view of the Viareggio test scene.

where $\tilde{\mathbf{C}}_\Delta = \mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}$ is the covariance matrix of the background mismatch whitened by $\mathbf{C}$. Diagonalising this covariance matrix, $\tilde{\mathbf{C}}_\Delta = \sum \epsilon_k \mathbf{u}_k\mathbf{u}_k^T$, the quadratic term here above writes $\frac{\mathbf{u}^T\tilde{\mathbf{C}}_\Delta\mathbf{u}}{\mathrm{Tr}\{\tilde{\mathbf{C}}_\Delta\}} = \sum(\frac{\epsilon_k}{\sum\epsilon_k})(\mathbf{u}_k^T\mathbf{u})^2$. Thereby, it corresponds to a weighted sum of the energy of $\mathbf{u}$. Unfortunately, in most cases, we do not have access to $\mathbf{C}_\Delta$ and the weights are unknown. To circumvent this problem, we simply propose to use a non-weighted sum to estimate this energy, namely $\frac{\mathbf{u}^T\tilde{\mathbf{C}}_\Delta\mathbf{u}}{\mathrm{Tr}\{\tilde{\mathbf{C}}_\Delta\}} \simeq \frac{\mathbf{u}^T\mathbf{u}}{N}$. This approximation amounts to consider that all the eigenvalues of $\tilde{\mathbf{C}}_\Delta$ are equal, namely that $\mathbf{C}_\Delta = \epsilon\mathbf{C}$. In other words, there is no approximation when the covariance matrix of the mismatch is proportional to the covariance matrix of the background $\mathbf{C}$. This hypothesis sometimes appears in the literature when considering target signature mismatches, even if the rationale behind it is not obvious [26]. Nevertheless, this hypothesis includes the popular case where both background and the mismatch are supposed to be white.

On the other hand, $\mathrm{Tr}\{\tilde{\mathbf{C}}_\Delta\} = \mathrm{Tr}\{\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_\Delta\mathbf{C}^{-\frac{1}{2}}\} = \mathrm{Tr}\{\mathbf{C}_\Delta\mathbf{C}^{-1}\} = \mathrm{Tr}\{\mathbb{E}\{\Delta\mathbf{b}\Delta\mathbf{b}^T\mathbf{C}^{-1}\}\} = \mathbb{E}\{\Delta\mathbf{b}^T\mathbf{C}^{-1}\Delta\mathbf{b}\}$ represents the expectation of the energy of the mismatch, whitened by $\mathbf{C}$. Assuming that $\Delta\mathbf{b}$ and $\mathbf{b}$ are decorrelated, we can estimate this energy as the difference between the energy of $\mathbf{b}_1$ and the energy of $\mathbf{b}$, both whitened by $\mathbf{C}$. The first one can be estimated by $\mathbf{u}^T\mathbf{u}$, as $\mathbf{u}$ is an estimation of

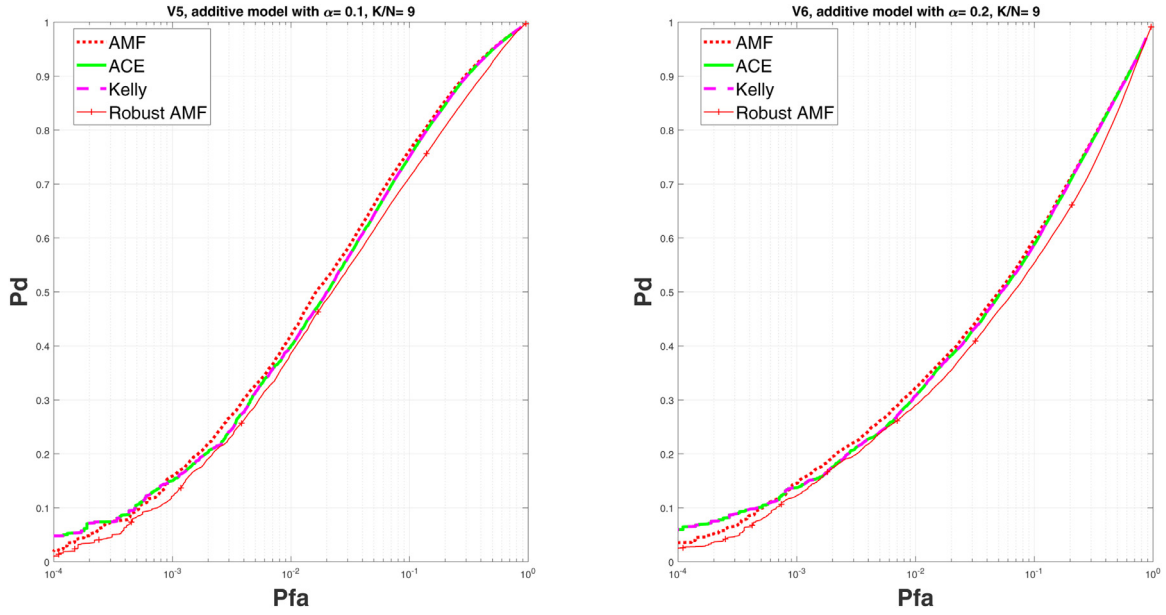

Fig. 4. Receiver Operating Characteristic for V_5 and V_6 targets, for $r = 0$ (no mismatch).

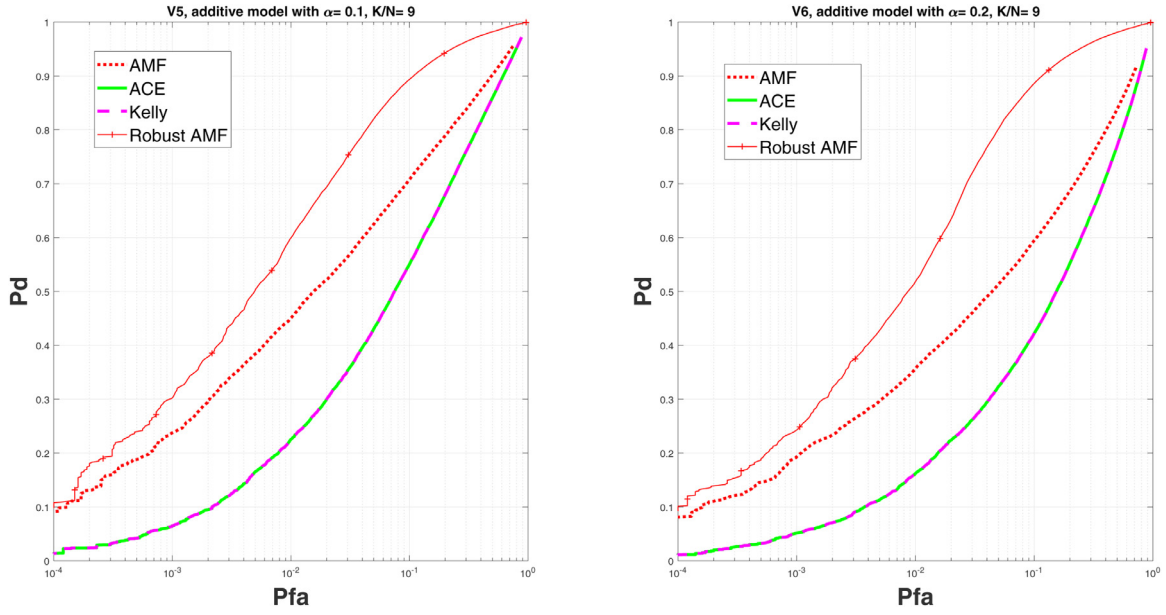

Fig. 5. Receiver Operating Characteristic for V_5 and V_6 target, for $r = 0.2$.

the background under H_1, whitened by $\mathbf{C}$, whereas the second one is simply equal to N, in average.

Substituting both $\tilde{\mathbf{C}}_\Delta$ and $\frac{\mathbf{u}^T\tilde{\mathbf{C}}_\Delta\mathbf{u}}{\mathrm{Tr}\{\tilde{\mathbf{C}}_\Delta\}}$ by their estimation in Eq. (14) we have

$$T \simeq N\left[\frac{\mathbf{u}^T\mathbf{u}}{N} - 1\right]^2 \tag{15}$$

Then, from Eq. (13), we can define a robust AMF as

$$AMF_r = AMF + 2\log\left[1 + \frac{N}{2}\left(\frac{\mathbf{u}^T\mathbf{u}}{N} - 1\right)^2\right] \tag{16}$$

At this stage, we can make a few comments on the proposed robust AMF. Whereas the standard AMF only measures the energy of the data projected on the target subspace, the proposed scheme also measures the remaining energy, projected on the background subspace. It checks if this energy corresponds to the background energy estimated from the training samples, and measures the gap. Then, it corrects the AMF with respect to this estimated mismatch energy. Other popular detectors also use the data energy projected on the background subspace to improve performance, like ACE or Kelly's GLRT, as we will see hereafter. In the following part, we compare our robust AMF with these popular detectors regarding their behaviour with respect to this residual energy. To this end, we plot both the data and the detectors threshold on the so-called Matched Filter Residual (MFR) diagrams. These informative $2D$ plots show both the energy projected on the target subspace (MF) and the energy projected on the subspace orthogonal to the target (R).

4. Insights

Just like for the AMF, the majority of popular detectors are based on the whitened and centred versions of both the data and

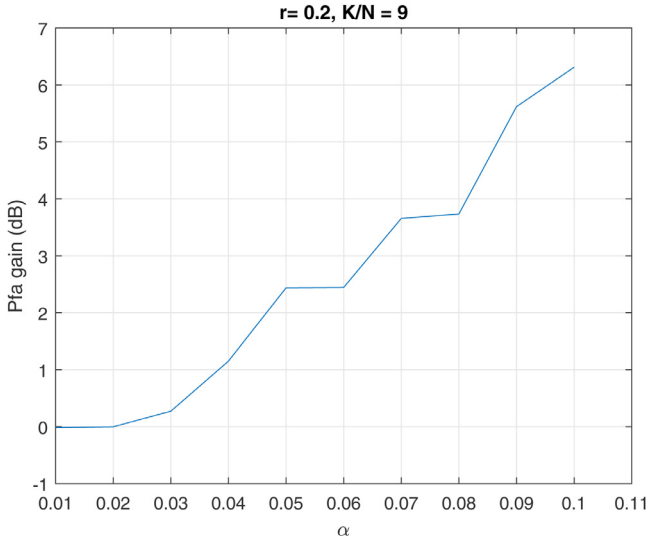

Fig. 6. P_{fa} gain for $P_d = 0.5$, for V5 target, with $\alpha = 0.1$ and $\frac{K}{N} = 9$.

the target signature, namely $\mathbf{y}_c$ and $\mathbf{t}_c$. As noticed previously, the standard AMF simply consists in projecting the data onto the $1D$ target subspace and in measuring the energy, as $AMF = \|\mathbf{P}_{\mathbf{t}_c}\mathbf{y}_c\|^2 = \mathbf{y}_c^T\mathbf{P}_{\mathbf{t}_c}\mathbf{y}_c$, where $\mathbf{P}_{\mathbf{t}_c} = \frac{\mathbf{t}_c\mathbf{t}_c^T}{\mathbf{t}_c^T\mathbf{t}_c}$ is the projection onto the target subspace. Nevertheless, the global energy, $\|\mathbf{y}_c\|^2$, or the residual energy, namely $R = \left\|\mathbf{P}_{\mathbf{t}_c}^{\perp}\mathbf{y}_c\right\|^2 = \mathbf{y}_c^T\mathbf{P}_{\mathbf{t}_c}^{\perp}\mathbf{y}_c = (\|\mathbf{y}_c\|^2 - AMF)$, where $\mathbf{P}_{\mathbf{t}_c}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{t}_c}$ is the projection onto the so-called noise subspace, are also informative. Indeed, when the target amplitude is unknown, this last quantity is an estimation of the Log likelihood of the data, up to a scaling factor and an additive constant. Hence plotting data in a $2D$ space, where the x-axis is the AMF and the y-axis is R, is convenient and valuable both for data and detectors analysis.

Such real data MFR analyses exist in the literature [27–29] and show that the scatter plots for all the pixels are distributed along lines converging to the origin of the plot, each line corresponding to a given material on the ground. Fig. 1 shows a representative situation with just 3 pixels classes (for instance, forest in green, roads in blue, and buildings in red), considering the buildings as the target for the AMF axis.

Observing such a MFR plot, it is clear that the standard AMF could not get good detection performance, as it only thresholds the x-axis, as shown on Fig. 2. To improve detection, one should also use the y-axis. That is what popular detection schemes do, such as Kelly's GLRT or ACE. Indeed, these two detectors share the same formulation, namely

$$\frac{AMF}{(a + \frac{1}{K}\mathbf{y}_c^T\mathbf{y}_c)} \tag{17}$$

where $a = 0$ for ACE, and $a = 1$ for Kelly, K being the number of training samples.

In both cases, comparing this quantity to a threshold γ amounts to comparing AMF to a limit depending on the square norm of the data, $\|\mathbf{y}_c\|^2$, so that the decision limit is defined by

$$AMF = \gamma\left(a + \frac{1}{K}\mathbf{y}_c^T\mathbf{y}_c\right) \tag{18}$$

Now, as $\|\mathbf{y}_c\|^2 = AMF + R$, the thresholds of these two detectors draw lines in the MFR plot, whose equation are

$$R = AMF\left(\frac{K}{\gamma} - 1\right) - Ka \tag{19}$$

As represented on Fig. 2, the threshold line corresponding to ACE goes through the origin of the plot, as $a = 0$. In this case, another intuitive way to see that the threshold region is a line, is to refer to the standard interpretation of ACE as the square cosine of the angle between $\mathbf{y}_c$ and $\mathbf{t}_c$. Thresholding ACE amounts to limit this angle as shown on Fig. 2. In the case of Kelly, the threshold is still a line, but shifted along the AMF axis, depending on the number of secondary data K.

Now, referring to the definition of $\mathbf{u}$ in Eq. (13), we have $R = \|\mathbf{u}\|^2$, so that the threshold limit of the robust AMF, AMF_r in the MFR plot is defined as

$$AMF = \gamma - 2\log\left[1 + \frac{N}{2}\left(\frac{R}{N} - 1\right)^2\right] \tag{20}$$

Then, this threshold limit is characterized by a second order polynomial equation in the MFR plot, as represented on Fig. 2. The threshold limit corresponding to the larger AMF being obtained when $R = N$. In the case of a white noise ($\mathbf{C} = \sigma^2\mathbf{I}$), this point corresponds to the case where the residual energy equals the noise energy estimated from the training samples, namely $\left\|\mathbf{P}_{\mathbf{t}}^{\perp}(\mathbf{y} - \boldsymbol{\mu})\right\|^2 = \sigma^2 N$. Hence, as can be seen on Fig. 2, the AMF_r has approximately the same behaviour as ACE or Kelly when the residual energy is smaller than that estimated from the training samples.

In the opposite case, the behaviour of AMF_r is not the same. Indeed, when the residual energy is larger than the training samples one, the threshold reduces to improve detection with respect to the data model from Eq. (2). In fact, according to this model, any noise energy variation with respect to the estimated one from the training samples is a clue for the H_1 hypothesis. Obviously, this possible detection improvement can also increase the detection of confuser elements. Indeed, improving detection according to model (2), goes against the so-called selectivity of the detector.

In situations where possible unwanted targets exit, one has to supplement the detection scheme by a so-called False Alarm Mitigation (FAM) step. These two-step procedures have been addressed in the hyperspectral literature [27–30], but also in the radar one [31]. The main idea behind this concept is to improve the selectivity issues by adding a second detection stage based on the residual energy. Indeed, as noticed previously, the residual energy is related to the likelihood of the data with respect to a given target type. Then, the possible false alarms existing after the first stage can be rejected using an appropriate threshold on the residual energy. This step amounts to adding a second threshold on the y-axis of the MFR plot. In the hyperspectral context, this procedure has been first developed to tackle the selectivity issues of the AMF, with the Matched Filter with False Alarm Mitigation (MF-FAM) [27] as represented on Fig. 2. This algorithm is very similar to the popular Mixture Tuned Matched Filter (MTMF) approach from ENVI®software environment. In this last case, the residual energy is replaced by an equivalent criterion, called infeasibility [30].

5. Performance evaluation

In order to assess the validity of the proposed robust detector, we now conduct a Monte-Carlo simulation based on a real experiment, namely the airborne Viareggio 2013 trial [32]. This benchmarking hyperspectral detection campaign took place in Viareggio (Italy), in May 2013, with an aircraft flying at 1200 meters. The open data consist in a $[450 \times 375]$ pixels map composed of 511 samples in the Visible Near InfraRed (VINR) band $(400 - 1000nm)$. The spatial resolution of the image is about 0.6 meters.

Different kinds of vehicles as well as coloured panels served as known targets. For each of these targets, a spectral signature obtained from ground spectroradiometer measurements is available. Moreover, a black and a white cover, serving as calibration targets, were also deployed. As can be seen on Fig. 3, the scene is composed of parking lots, roads, buildings, sport fields and pine woods.

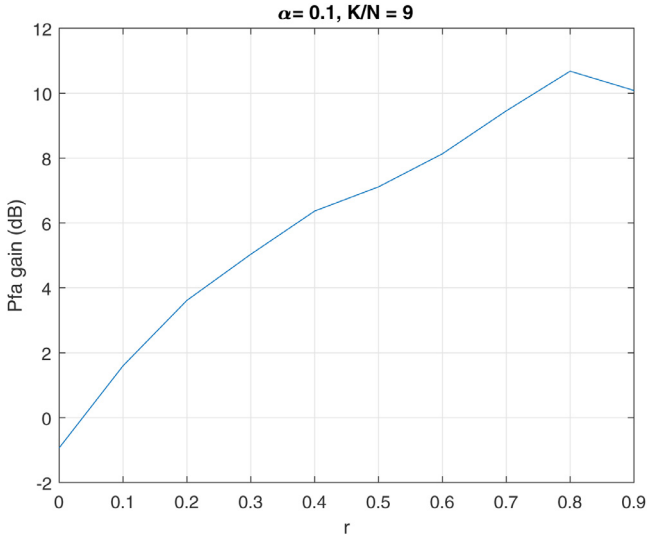

Fig. 7. P_{fa} gain for $P_d = 0.5$, for V5 target, with $\alpha = 0.1$ and $\frac{K}{N} = 9$.

As for the majority of hyperspectral detection schemes, the first step of the processing aims at converting the raw measurements into a reflectance map, namely removing all atmospheric effects and non-uniform sun illumination. To this end, we use the Empirical Line Method (ELM) [23,25], considering the black and white calibration panels. Then a spectral binning [33] is performed to reduce the vector size dimension to $N = 32$.

Based on this map, we will conduct a Monte-Carlo simulation by inserting a synthetic target $\mathbf{t}$, initially not present in the image, into a randomly chosen pixel $\mathbf{y}$ of the map ($\mathbf{y} \rightarrow (\alpha\mathbf{t} + \mathbf{y})$, where α is the target amplitude). The target signature is subject to uncertainties and is composed of the sum of the assumed target signature $\bar{\mathbf{t}}$ and a random zero-mean normally distributed white vector $\mathbf{t}_\Delta$, whose energy is r times the energy of $\bar{\mathbf{t}}$:

$$\mathbf{t} = \bar{\mathbf{t}} + \mathbf{t}_\Delta \tag{21}$$

where $\mathbf{t}_\Delta \sim \mathcal{N}(\mathbf{0}, r\frac{\bar{\mathbf{t}}^T\bar{\mathbf{t}}}{N}\mathbf{I})$.

Figs. 4 and 5 represent the Receiver Operating Characteristic (ROC) of different detectors, for both the so-called V_5 and V_6 targets and for $r = 0$ (no mismatch) and $r = 0.2$, respectively. In both cases, we use a 17×17 window to estimate the background mean and the sample covariance matrix. The target's amplitude are $\alpha = 0.1$ for V_5 and $\alpha = 0.2$ for V_6, as this last one is more difficult to detect.

First of all, in the nominal case, where no mismatch exists ($r = 0$), all the detectors exhibit approximately the same performance, with a slight loss for the proposed AMF_r. On the other hand, when there is a mismatch on the target's signature, the robust AMF outperforms the other detectors. We can notice that both ACE and Kelly exhibit a larger loss than the standard AMF. Indeed, the existence of a mismatch on the target signature increases the residual part (R) of the data. Then, referring to Fig. 2, it is obvious that both ACE and Kelly will be the more affected in this situation. We highlight here the compromise between selectivity and robustness. That is why the FAM techniques, using first a robust scheme to avoid non-detection, followed by a selective scheme to sort the targets and reduce false alarm, are worthy of interest.

Now, we will successively analyse the influence of the fill factor α and of the level of uncertainties r. To this end, we compute the gain provided by AMF_r compared to AMF in terms of probability of false alarm (P_{fa}) for a given P_d, i.e., $10\log_{10}\frac{P_{fa}(AMF)}{P_{fa}(AMF_r)}$, for $P_d = 0.5$. This P_{fa} gain allows to measure the improvement (if positive) or the loss (when negative) of the robust AMF with respect to the standard one. We consider the so-called V_5 target for both these analysis. Fig. 6 represents this false alarm gain with respect to α. On this figure we keep the same configuration as in Fig. 4. We roughly observe a linear performance increase as the target amplitude increases.

To finish with, we now make vary the level of uncertainties on $\mathbf{t}$, namely r, in Fig. 7. As expected, we observe that the improvement proposed by the robust AMF improves as the level of uncertainties increases, with less than $1db$ loss when the target signature is perfectly known ($r = 0$). Therefore, as soon as the target signature is not perfectly known, the proposed robust AMF has to be considered.

6. Conclusions

In this paper, we considered the detection of a target in an hyperspectral image when the background distribution may slightly vary between the null hypothesis and the alternate one, due e.g., to spectral variability. Under such an assumption we derived the LR for any background distribution, and the GLRT under the popular Gaussian case. This last detector is shown to be a simple correction of the AMF. We provided insights into the behaviour of this robust scheme and showed its good performance through numerical simulations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Proof of (4)

In this appendix, we derive the proof to Eq. (4). For any random vector $\mathbf{u}$, we can write $\mathbf{u} = \mathbb{E}\{\mathbf{u}\} + \Delta\mathbf{u}$, where $\Delta\mathbf{u}$ is a zero mean vector with the same covariance matrix as $\mathbf{u}$, namely $\mathbf{C_u}$. Then, we can write the following Taylor series expansion around $\mathbb{E}\{\mathbf{u}\}$:

$$f(\mathbf{u}) = f(\mathbb{E}\{\mathbf{u}\} + \Delta\mathbf{u}) \simeq f(\mathbb{E}\{\mathbf{u}\}) + \frac{\partial f(\mathbf{u})}{\partial \mathbf{u}^T}|_{\mathbb{E}\{\mathbf{u}\}}\Delta\mathbf{u} + \frac{1}{2}\Delta\mathbf{u}^T\left[\frac{\partial^2 f(\mathbf{u})}{\partial\mathbf{u}\partial\mathbf{u}^T}|_{\mathbb{E}\{\mathbf{u}\}}\right]\Delta\mathbf{u} \tag{A.1}$$

Taking the expectation of this expression leads to

$$\mathbb{E}\{f(\mathbf{u})\} \simeq f(\mathbb{E}\{\mathbf{u}\}) + \frac{1}{2}\mathbb{E}\left\{\Delta\mathbf{u}^T[\frac{\partial^2 f(\mathbf{u})}{\partial\mathbf{u}\partial\mathbf{u}^T}|_{\mathbb{E}\{\mathbf{u}\}}]\Delta\mathbf{u}\right\} \tag{A.2}$$

$$= f(\mathbb{E}\{\mathbf{u}\}) + \frac{1}{2}\text{Tr}\{[\frac{\partial^2 f(\mathbf{u})}{\partial\mathbf{u}\partial\mathbf{u}^T}|_{\mathbb{E}\{\mathbf{u}\}}]\mathbf{C_u}\} \tag{A.3}$$

CRediT authorship contribution statement

François Vincent: Conceptualization, Methodology, Validation, Formal analysis, Software, Writing - original draft, Writing - review & editing. **Olivier Besson:** Methodology, Validation, Writing - review & editing.

References

[1] D.G. Manolakis, R.B. Lockwood, T.W. Cooley, Hyperspectral Imaging Remote Sensing, Cambridge University Press, 2016.
[2] E. Keith, H. Dan, O. William, J. Shridhar, B. Eustace, L. Dereniak, Hyperspectral imaging for astronomy and space surveillance, in: Proc. SPIE, vol. 5159, 2004.
[3] S. Michel, P. Gamet, M.-J. Lefevre-Fonollosa, HYPXIM A hyperspectral satellite defined for science, security and defence users, Processing 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011.

[4] H. Kwon, N. Nasrabadi, Kernel rx-algorithm: a nonlinear anomaly detector for hyperspectral imagery, IEEE Trans. Geosci. Remote Sens. 43 (2) (2005) 388–397.
[5] E.M. Winter, M. Miller, C. Simi, A. Hill, T. Williams, D. Hampton, M. Wood, J. Zadnick, M. Sviland, Mine detection experiments using hyperspectral sensors, SPIE Int. Soc. Opt. Eng., Orlando, FL, United States, 2004.
[6] C.C. Funk, J. Theiler, D.A. Roberts, C.C. Borel, Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal imagery, IEEE Trans. Geosci. Remote Sens. 39 (7) (2001) 1410–1420.
[7] D.-W. Sun, Hyperspectral Imaging for Food Quality Analysis and Control, Elsevier, 2010.
[8] R. Koprowski, Processing of Hyperspectral Medical Images, Applications in Dermatology Using Matlabë, Springer International Publishing, 2017.
[9] D. Manolakis, G. Shaw, Detection algorithms for hyperspectral imaging applications, IEEE Signal Process. Mag. 19 (1) (2002) 29–43.
[10] E. Kelly, An adaptive detection algorithm, IEEE Trans. Aerosp. Electron.Syst. 22 (2) (1986) 115–127.
[11] C. Chang, Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis, IEEE Trans. Geosci. Remote Sens. 43 (3) (2005) 502–518.
[12] S. Kraut, L.L. Scharf, L.T. McWhorter, Adaptive subspace detectors, IEEE Trans. Signal Process. 49 (1) (2001) 1–16.
[13] M.J. Carlotto, A cluster-based approach for detecting man-made objects and changes in imagery, IEEE Trans. Geosci. Remote Sens. 43 (2) (2005) 374387.
[14] D.B. Marden, D. Manolakis, Using elliptically contoured distributions to model hyperspectral imaging data, in: Proceedings of SPIE - The International Society for Optical Engineering, vol. 5425, 2004, p. 558572.
[15] D.G. Manolakis, D. Marden, J.P. Kerekes, G.A. Shaw, Statistics of hyperspectral imaging data, in: Proceedings SPIE 4381 Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII, vol. 4381, 2001, pp. 308–316.
[16] S. Matteoli, M. Diani, G. Corsini, A tutorial overview of anomaly detection in hyperspectral images, IEEE Aerosp. Electron. Syst. Mag. 25 (7) (2010) 5–27.
[17] J. Theiler, B.R. Foy, EC-GLRT: Detecting weak plumes in non-Gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test, in: Proceedings IGARSS, vol. 1, 2008, pp. 221–224. Boston, MA.
[18] N. Dobigeon, J.-Y. Tourneret, C. Richard, J.C.M. Bermudez, S. McLaughlin, A.O. Hero, Nonlinear unmixing of hyperspectral images, IEEE Signal Process. Mag. (2014) 82–94.
[19] A. Schaum, A. Stocker, Spectrally-selective target detection, in: Proceedings of ISSSR, vol. 12, 1997, pp. 2015–2018.
[20] F. Vincent, O. Bessson, One-step generalized likelihood ratio test for subpixel target detection in hyperspectral imaging, IEEE Trans. Geosci. Remote Sens. 58 (6) (2020) 4479–4489.
[21] F. Vincent, O. Bessson, Generalized likelihood ratio test for modified replacement model in hyperspectral imaging detection, Signal Process. 174 (2020) 107643.
[22] R. DiPietro, D. Manolakis, R. Lockwood, T. Cooley, J. Jacobson, Performance evaluation of hyperspectral detection algorithms for sub-pixel objects, in: Proceedings of SPIE - The International Society for Optical Engineering, 2010.
[23] G. Ferrier, Evaluation of apparent surface reflectance estimation methodologies, Int. J. Remote Sens. 16 (1995) 2291–2297.
[24] T. Popp, Correcting atmospheric masking to retrieve the spectral albedo of land surfaces from satellite measurements, Int. J. Remote Sens. 16 (1995) 34833508.
[25] G.M. Smith, E.J. Milton, The use of the empirical line method to calibrate remotely sensed data to reflectance, Int. J. Remote Sens. 20 (1999) 2653–2662.
[26] D. Manolakis, R. Lockwood, T. Cooley, J. Jacobson, Is there a best hyperspectral detection algorithm? in: Proc. of SPIE, vol. 7334, 2009.
[27] R.S. DiPietro, D. Manolakis, R.B. Lockwood, T. Cooley, J. Jacobson, Hyperspectral matched filter with false-alarm mitigation, Opt. Eng. 51 (1) (2012) 1–8.
[28] E. Truslow, Performance evaluation of the adaptive cosine estimator detector for hyperspectral imaging applications, Northeastern University Boston, Massachusetts, 2012 Ph.D. thesis.
[29] R.S. DiPietro, The detection of sub-pixel objects and mitigation of false alarms in hyperspectral imagery, Northeastern University Boston, Massachusetts, 2010 Ph.D. thesis.
[30] J.W. Boardman, F.A. Kruse, Analysis of imaging spectrometer data using n-dimensional geometry and a mixture-tuned matched filtering approach, IEEE Trans. Geosci. Remote Sens. 49 (11) (2011) 4138–4152.
[31] N.B. Pulsone, C.M. Rader, Adaptive beamformer orthogonal rejection test, IEEE Trans. Signal Process. 49 (2001) 521–529.
[32] N. Acito, S. Matteoli, A. Rossi, M. Diani, G. Corsini, Hyperspectral airborne ăViareggio 2013 trialg data collection for detection algorithm assessment, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (6) (2016) 2356–2376.
[33] M. Shi, G. Healey, Hyperspectral texture recognition using a multiscale opponent representation, IEEE Trans. Geosci. Remote Sens. 41 (5) (2003) 1090–1095.