Anomaly detection for replacement model in hyperspectral imaging

Francois Vincent, Olivier Besson, Stefania Matteoli

Anomaly detection for replacement model in hyperspectral imaging

Francois Vincent [a,*], Olivier Besson [a], Stefania Matteoli [b]

[a] ISAE-SUPAERO, 10 Avenue Edouard Belin, 31055 Toulouse, France
[b] CNR-IEIIT, via Girolamo Caruso 16, Pisa, Italy

A B S T R A C T

In this paper we consider Anomaly Detection in the hyperspectral context, and we extend the popular RX detector, initially designed under the standard additive model, to the replacement model case. Indeed, in this more realistic framework, the target, if present, is supposed to replace a part of the background. We show how to estimate this background power variation to improve the standard RX scheme. The obtained Replacement RX (RRX) is shown to be closed-form and outperforms the standard RX on a real data benchmark experiment.

Keywords:
Hyperspectral imagery
Replacement model
GLRT
Anomaly detection

1. Introduction

Increasing the number of spectral bands of a 2-D image allows to recover a rich information from the environment. Thereby, HyperSpectral Imaging (HSI) has encountered a large field of applications, ranging from remote sensing to medicine [1–8]. One of the main processing step consists in separating a target $\mathbf{t} \in \mathcal{R}^N$, rarely present in the map, from the background $\mathbf{b} \in \mathcal{R}^N$ predominant in the image. This detection step is subject to one main difficulty, namely characterizing properly both the target and the background signatures.

Concerning the target, two main detection schemes are usually considered whether the target's signature is supposed to be known or not. In this last case, known as Anomaly Detection (AD), the target's signature is usually assumed to be deterministic and unknown. In the other case, known as target detection, the target's signature is usually supposed to be perfectly known, even if in practice there exists a mismatch between the real target's signature and the presumed one. This so-called spectral variability may cause significant losses in the detection precessing. The difficulty to model this mismatch has led to a two-stages processing, where one conducts first an AD step, relaxing the hypotheses on the target in considering $\mathbf{t}$ as unknown, followed by a classification step to sort out the different targets. In this paper we focus on the first step, namely AD.

Now, concerning the background, it is usually modelled as a random vector, whose parameters can be estimated from the so-called training samples. This training samples consist in pixels surrounding the Pixel Under Test (PUT) and hopefully free of targets.

The popular way to describe the AD problem is a two-hypotheses test, based on the additive model:

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

where the background is usually assumed to follow a Gaussian distribution $\mathbf{b} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$). In this case, the Generalized Likelihood Ratio Test (GLRT) leads to the popular RX detector [9]. This benchmark AD scheme simply consists in comparing the log likelihood under H_0 to a threshold.

Many versions of the RX detector exist. For instance, when choosing a local and limited training window around the PUT, one refers to Local RX (LRX), and when using the whole image to estimate the background mean and covariance matrix, one refers to Global RX (GRX). On the one hand, the training samples are surely more representative of the background in the PUT but the small number of samples can conduct to ill-conditioned covariance matrix estimations. On the other hand, the covariance matrix does not suffer from conditioning problems, but could be less representative of the background to be described, even if some techniques exist for "cleaning" the training samples from possible outliers, such as the BACON technique for instance [10].

This difficulty to describe real life spatially moving background [11–14] has led to the development of many modified RX techniques [13], the majority are based on the decomposition of the background into more homogeneous clusters [15,16], such as Gaus-

* Corresponding author.
E-mail addresses: olivier.besson@isae-supaero.fr (O. Besson), stefania.matteoli@ieiit.cnr.it (S. Matteoli).

https://doi.org/10.1016/j.sigpro.2021.108079

sian Markov Random Field (GMRF) AD [17,18], or Cluster Based (CB) AD [11,19]. To tackle this problem, other authors also proposed a non-linear version of RX, the Kernel RX [4].

Nevertheless, all these algorithms are based on the simple additive model described in eq. (1). Yet, in the HSI context, when dealing with reflectance data, more representative models exist, such as the popular replacement model [20,21] or the modified replacement model [22]. Indeed, this last model assumes that the target, if present replaces a part of the background, due to a masking effect: $\mathbf{y} = \eta\mathbf{t} + \beta\mathbf{b}$, where η is the target abundance, with $\eta = (1-\beta)$ in the standard replacement case, and $\beta < 1$ represents the background reduction. The over-simplified additive model from eq. (1), being an approximation when the target size is small with respect to the pixel's area ($\eta << 1$). Detection schemes based on this more representative model have proven their superiority in the target detection context [21–24]. But, it seems more difficult to use it in the present case of AD, as the number of unknowns would be bigger than the size of the data. The goal of this paper is to show how to bypass this issue to derive an AD based on the replacement model to improve the performances of the standard RX detector.

To this end, we will consider two steps. First of all, we will consider that the scaling factor on the background, namely β is known. Then, in a second step, we will estimate β using a low-rank assumption on the background. Indeed, it is widely accepted that the background is composed of a few number of representative materials, known as endmembers. In fact, many AD schemes exploit this situation, such as the Robust Principal Components Analysis (RPCA)AD [9] or the Low-Rank and Sparse Matrix Decomposition (LRaSMD) AD [25,26] and their improvements [27,28]. Hence, under such a widespread hypothesis, we will show in Section 2 how to get a relevant estimation of β. Then, this parameter will be plugged into the GLRT derived from the replacement model, assuming β known. To finish, we will explain why this new AD scheme should significantly improve the performance compared to the standard RX.

The remaining of this paper is organized as follows. We first derive the GLRT for the replacement detection problem from eq. (2) and give some insights to understand the differences with the popular RX detector. Then, we compare this detector to the standard RX, using a Monte-Carlo simulation based on real data, in Section 3. Finally concluding remarks end this paper in Section 4.

2. GLRT Derivation for the replacement model

In the case of AD, as the target signature $\mathbf{t}$ is unknown, the two replacement models described in the introduction amount to consider the same following test, where the target amplitude η has been integrated in the unknown signature $\mathbf{t}$:

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

As $\mathbf{t}$ is assumed to be deterministic, this more realistic model exhibits two differences between the distributions under H_0 and H_1, namely a change on the mean, as for the additive model (1), but also a scaling factor on the covariance matrix. We will show how to exploit this last information to improve performances. But let's first recall the GLRT formulation in the standard additive case, namely RX detector.

2.1. GLRT For the additive model (RX)

In the additive case, it is straight-forward to derive the GLRT, that simply consists in comparing the log-likelihood to a threshold:

$$RX = (\mathbf{y}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu}) = \left\|\mathbf{C}^{-\frac{1}{2}}\mathbf{y} - \mathbf{C}^{-\frac{1}{2}}\boldsymbol{\mu}\right\|^2 \tag{3}$$

where the mean and the covariance matrix of the background, namely $\boldsymbol{\mu}$ and $\mathbf{C}$ should be estimated from the training samples. Hence, RX amounts to measuring the difference between the data and the mean spectral signature estimated from the secondary data, after a whitening step. This gap being estimated by the quadratic norm of this difference. Hence, it only exploits the distance from the assumed mean, and not any possible changes in terms of background power. Nevertheless, in real life, if a target is present in the PUT, it usually tackles a background power decrease, as stated in model (2). The goal of this section is to derive a detector that both exploits the gap to the mean, but also the gap to the secondary data background power, two pieces of information allowing to separate the two hypotheses of the detection problem. To this end, we first derive the GLRT assuming that the scaling factor on the background β is known, and then we show how to estimate β based on the low-rank property of the background.

2.2. GLRT For the replacement model with β known

First of all, like for the standard RX detector, we consider a so-called two-step approach to derive the GLRT [21], in this paper. That is to say, we assume in a first step that the background parameters $(\boldsymbol{\mu}, \mathbf{C})$ are known, then we replace them with their estimations from the training samples. Then, in order to simplify the derivations, we can simplify model (2) by a whitening step, so that the detection problem can be written as follows.

$$\begin{aligned} H_0 &: \mathbf{x} = \mathbf{u} \\ H_1 &: \mathbf{x} = \mathbf{s} + \beta\mathbf{u} \end{aligned} \tag{4}$$

where $\mathbf{x} = \mathbf{C}^{-\frac{1}{2}}\mathbf{y}$ and $\mathbf{u} = \mathbf{C}^{-\frac{1}{2}}\mathbf{b}$ are the whitened counterparts of $\mathbf{y}$ and $\mathbf{b}$, so that $\mathbf{u} \sim \mathcal{N}(\mathbf{m}, \mathbf{I}))$, with $\mathbf{m} = \mathbf{C}^{-\frac{1}{2}}\boldsymbol{\mu}$, and $\mathbf{s} = \mathbf{C}^{-\frac{1}{2}}\mathbf{t}$ the unknown signal signature.

Then, the logarithm of the Likelihood Ratio (LR) is shown to be

$$L = -N\log\beta + \frac{1}{2}\left[\|\mathbf{x}-\mathbf{m}\|^2 - \frac{\|\mathbf{x}-\beta\mathbf{m}-\mathbf{s}\|^2}{\beta^2}\right] \tag{5}$$

Assuming, in this first step, that β is known and $\mathbf{s}$ is unknown, the Maximum Likelihood (ML) of $\mathbf{s}$ leads to removing the last term in the previous equation, so that the GLRT, up to a scaling factor, becomes:

$$GLRT = \|\mathbf{x}-\mathbf{m}\|^2 - 2N\log\beta = RX - 2N\log\beta \tag{6}$$

Hence, considering the replacement model instead of the additive one, simply conducts to a modified version of the popular RX detector, including a correction factor : $-2N\log\beta$, based on the background reduction β. It has to be noticed that this additive term is still positive, as $\beta \leqslant 1$ and null when $\beta = 1$, so that we recover the standard RX detector in such a case. Then, we can feel that this term can improve detection, by increasing the test when a power reduction appears under H_1.

As stated in the introduction, the main issue is that β is not known and could not be estimated as an extra parameter under H_1, because we have more unknowns that measurements, in such a case. Nevertheless, in the following subsection, we show how to obtain a closed-form estimation of β.

2.3. β Estimation

As stated in the introduction, it is a standard assumption to consider that the background is composed of a low-rank main part plus an additive white noise, and can be written as follows.

$$\mathbf{b} = \mathbf{U}\boldsymbol{\alpha} + \mathbf{n} \tag{7}$$

where $\mathbf{U} \in \mathcal{R}^{N\times K}$ represents the signal subspace of the covariance matrix, $\boldsymbol{\alpha}$ is assumed to be Gaussian distributed with a mean $\boldsymbol{\mu}_\alpha$

Table 1
RRX algorithm.

Background parameters estimation	Compute $\hat{\mathbf{C}} = \frac{1}{M}\sum_{k=0}^{k=M-1}\mathbf{z}_k\mathbf{z}_k^T$ and $\hat{\mu} = \bar{\mathbf{z}}$
from secondary data $\mathbf{z}_k \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$	Identify subspace parameters: $\hat{\mathbf{C}} = \mathbf{U}\mathbf{R}_\alpha\mathbf{U}^T + \sigma^2\mathbf{I}$
for $k = 0, ., (M-1)$	Compute $\boldsymbol{\mu}_\alpha = \left(\mathbf{U}^T\hat{\mathbf{C}}^{-1}\mathbf{U}\right)^{-1}\mathbf{U}^T\hat{\mathbf{C}}^{-1}\bar{\mathbf{z}}$
Projection on $<\mathbf{U}>$ to maximize	$\hat{\beta} = Min\left[\frac{1}{2K}\left(\sqrt{(\boldsymbol{\mu}_\alpha^T\mathbf{C}_U^{-1}\mathbf{y}_U)^2 + 4K\mathbf{y}_U^T\mathbf{C}_U^{-1}\mathbf{y}_U} - (\boldsymbol{\mu}_\alpha^T\mathbf{C}_U^{-1}\mathbf{y}_U)\right), 1\right]$
the background and estimation of β	where $\mathbf{y}_U = \mathbf{U}^T\mathbf{y}$ and $\mathbf{C}_U = \mathbf{R}_\alpha + \sigma^2\mathbf{I}$
Test computation	$RRX = (\mathbf{y}-\boldsymbol{\mu})^T\hat{\mathbf{C}}^{-1}(\mathbf{y}-\boldsymbol{\mu}) - 2N\log\hat{\beta}$
and comparison to a threshold	$= RX - 2N\log\hat{\beta} \lessgtr \delta$

and a covariance matrix $\mathbf{R}_\alpha$, namely $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \mathbf{R}_\alpha)$, and $\mathbf{n}$ is a zero mean white Gaussian noise with power σ^2 $\left(\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})\right)$, independent from $\boldsymbol{\alpha}$. In such a case, the background covariance matrix writes $\mathbf{C} = \mathbf{U}\mathbf{R}_\alpha\mathbf{U}^T + \sigma^2\mathbf{I}$, with $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$, to ensure uniqueness of the decomposition. Hence, according to eq. (2), under H_1, the signal can be written as follows

$$\mathbf{y} = \mathbf{t}_\perp + \mathbf{t}_U + \beta(\mathbf{U}\boldsymbol{\alpha} + \mathbf{n}) \tag{8}$$

where we have decomposed $\mathbf{t}$ in its part belonging to the background subspace $<\mathbf{U}>$, namely $\mathbf{t}_U = \mathbf{U}\boldsymbol{\gamma}$, and an orthogonal part $\mathbf{t}_\perp$. It has to be noticed that this last part is the main contribution in the detection process, as the background-like part, namely $\mathbf{t}_U = \mathbf{U}\boldsymbol{\gamma}$ will be mostly rejected during the whitening step present in most processing. Thereby, projecting the data onto $<\mathbf{U}>$, we can maximize the background power while minimizing the effects of the unknown target signature $\mathbf{t}$, to obtain an estimation of β, as shown hereafter.

As the dimension K of this subspace is supposed to be small compared to N, we consider in a first approximation that $\mathbf{t}_U$ is negligible compared to the background power. Then, starting from eq. (8) and projecting the data $\mathbf{y}$ onto the background subspace $<\mathbf{U}>$, we have

$$\mathbf{y}_U \triangleq \mathbf{U}^T\mathbf{y} = \boldsymbol{\gamma} + \beta(\boldsymbol{\alpha} + \mathbf{n}_U) \simeq \beta(\boldsymbol{\alpha} + \mathbf{n}_U) \tag{9}$$

where $\mathbf{n}_U = (\mathbf{U}^T\mathbf{n}) \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, and we recall that $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \mathbf{R}_\alpha)$.

Then, from eq. (9), considering that $\boldsymbol{\gamma}$ is negligible, the log-likelihood of $\mathbf{y}_U$ writes

$$-\frac{1}{2}\log\left[\beta^{2K}|\mathbf{C}_U|\right] - \frac{1}{2}\left[\frac{(\mathbf{y}_U - \beta\boldsymbol{\mu}_\alpha)^T\mathbf{C}_U^{-1}(\mathbf{y}_U - \beta\boldsymbol{\mu}_\alpha)}{\beta^2}\right] \tag{10}$$

where $\mathbf{C}_U = \mathbf{R}_\alpha + \sigma^2\mathbf{I}$ is the covariance matrix of $(\boldsymbol{\alpha} + \mathbf{n}_U)$. Then, deriving with respect to β and comparing to zero, it is straightforward to get the ML estimate of β:

$$\hat{\beta}_{ML} = \frac{1}{2K}\left[\sqrt{(\boldsymbol{\mu}_\alpha^T\mathbf{C}_U^{-1}\mathbf{y}_U)^2 + 4K\mathbf{y}_U^T\mathbf{C}_U^{-1}\mathbf{y}_U} - (\boldsymbol{\mu}_\alpha^T\mathbf{C}_U^{-1}\mathbf{y}_U)\right] \tag{11}$$

To ensure that β is lower than 1 and as a consequence, that the correction term in eq. (6) will be positive, we choose after all $\hat{\beta} = min[\hat{\beta}_{ML}, 1]$.

It has to be noticed that $\hat{\beta}$ depends on both $\mathbf{C}_U$ and $\boldsymbol{\mu}_\alpha$ that can be easily estimated from the training samples. Indeed, $\mathbf{R}_\alpha$ and σ^2 are obtained from the eigen-decomposition of $\mathbf{C} = \mathbf{U}\mathbf{R}_\alpha\mathbf{U}^T + \sigma^2\mathbf{I}$, and it is straight-forward to derive the ML estimation of $\boldsymbol{\mu}_\alpha$ as the secondary data writes $\mathbf{z}_k = \mathbf{U}\boldsymbol{\alpha} + \mathbf{n}_k \quad k = 0, ., (M-1)$ where $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ and $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \mathbf{R}_\alpha)$:

$$\boldsymbol{\mu}_\alpha = \left(\mathbf{U}^T\mathbf{C}^{-1}\mathbf{U}\right)^{-1}\mathbf{U}^T\mathbf{C}^{-1}\bar{\mathbf{z}} \tag{12}$$

where $\bar{\mathbf{z}} = \frac{1}{M}\sum_{k=0}^{k=M-1}\mathbf{z}_k$.

2.4. Insights

This expression of $\hat{\beta}$ is valid as soon as $\mathbf{t}_U$ is negligible with respect to the background power. In other words, this approximation is accurate as soon as $\mathbf{t}$ contains most of its power outside the background subspace $<\mathbf{U}>$, whose dimension is usually much smaller than N. It can be noticed that if it is not the case it will be very difficult to detect the target in any cases, as the target has almost the same signature than the background, and will be mostly rejected during the whitening step, always present in most detectors. Nevertheless, even in such a case, we should observe a background power change ($\hat{\beta} < 1$). Indeed, as we consider reflectance data $\mathbf{y}$, the power of the data remains the same whatever the hypothesis. Then, the power projected onto $<\mathbf{U}>$, under H_1, will be reduced by the part of $\mathbf{t}$ being outside the background subspace, in contrast to H_0. So even if β would be over-estimated in such a case, it will remain under 1, still improving the standard RX detector. Although, once again, in this situation none AD scheme will achieve good results because target and background get similar signatures.

2.5. Summary

To sum-up the derivation presented here-above, the proposed procedure to compute the AD adapted to the replacement model is summarize in Table 1. For convenience reasons and because of its formulation very close to the popular RX detector, we name it Replacement RX (RRX).

3. Performance evaluation

In order to compare the proposed RRX detector with its additive model AD counterpart, namely RX, we now conduct a Monte-Carlo simulation based on a real data experiment. To this end, we consider the airborne Viareggio 2013 trial [29] that took place in Viareggio (Italy), in May 2013, with an aircraft flying at 1200 meters so that the nominal spatial resolution of the image is about 0.6 meters. The image is composed of $[450 \times 375]$ pixels with 511 samples in the Visible Near InfraRed (VINR) band $(400 - 1000nm)$.

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. As can be seen on Fig. 1, the scene is composed of parking lots, roads, buildings, sport fields and pine woods.

In addition, calibration targets can be used to remove all atmospheric effects and non-uniform sun illumination, to convert the raw radiance measurements into reflectance. This so called radiometric compensation is conducted here using the Empirical Line Method (ELM) [30,31]. Then a spectral binning [32] is performed to reduce the vector size dimension to $N = 128$.

Based on this real life environment, we will conduct a synthetic target detection benchmark to obtain statistical results. To this end, we will randomly insert two different real life target initially not present in the image (the so-called V_4 and V_5 targets), thanks to model (2) and for different values of β. To estimate the Sample Covariance Matrix (SCM) $\hat{\mathbf{C}}$ as well as the mean of the background $\hat{\boldsymbol{\mu}}$, we use a 27×27 pixels window, corresponding to $M = 5.7 \times N$ samples. In order to estimate β, we need to define the main background subspace $<\mathbf{U}>$. In this experiment, we define it as the subspace containing 99% of the total background energy, so that

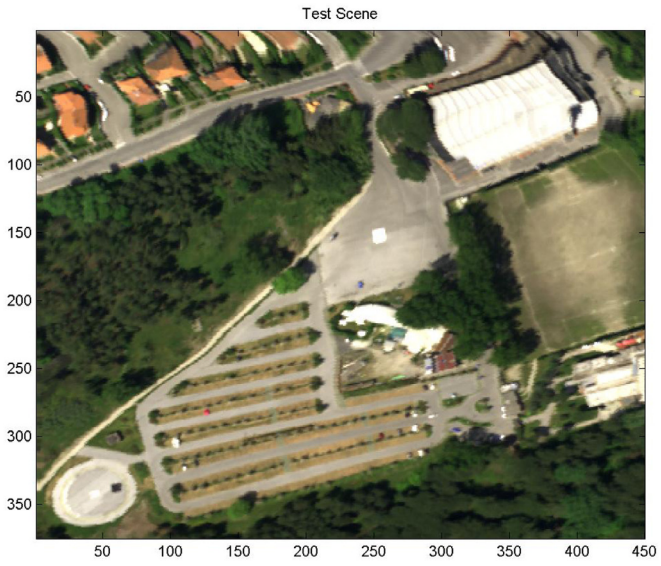

Fig. 1. Complete RGB view of the Viareggio test scene.

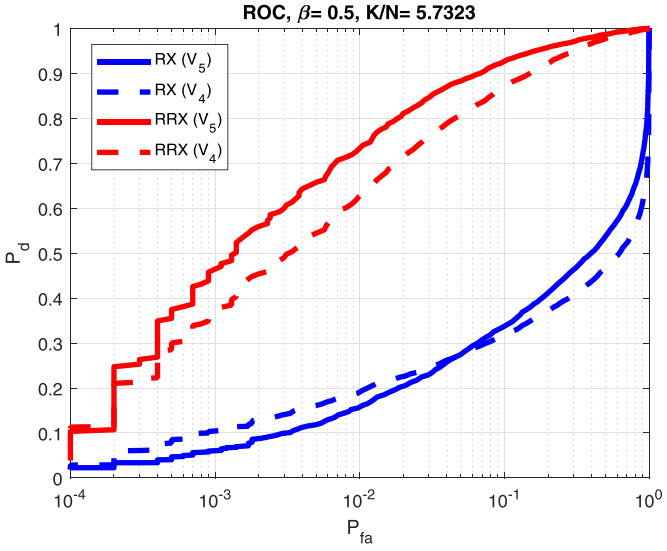

Fig. 3. Receiver Operating Characteristic comparison for $\beta = 0.5$.

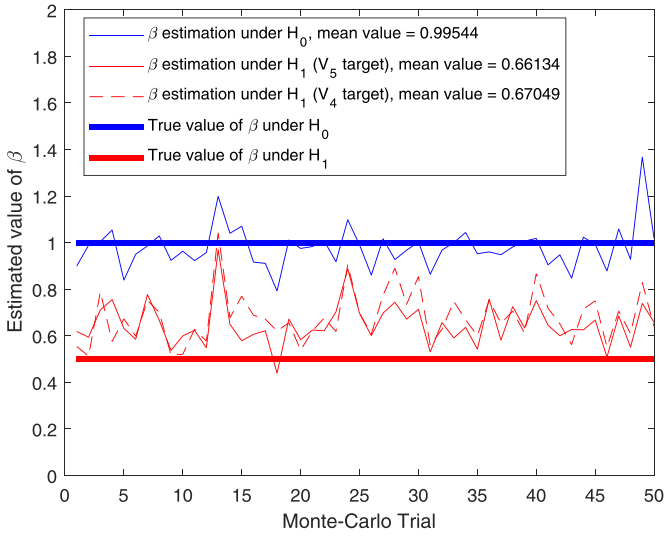

Fig. 2. Exemples of $\hat{\beta}$ estimation under H_0 and H_1.

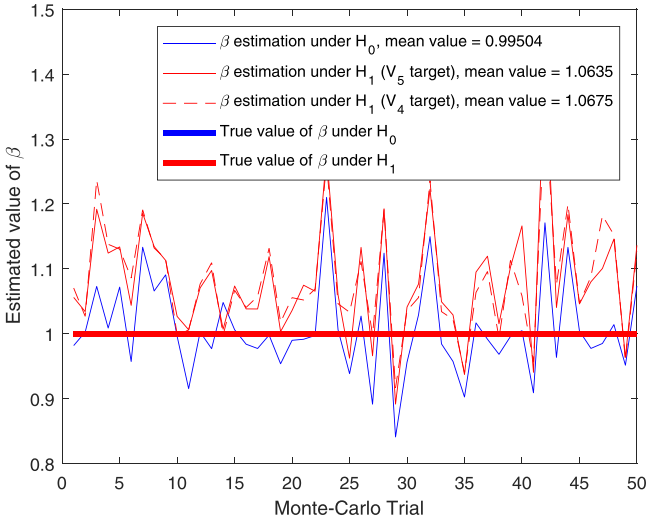

Fig. 4. Exemples of $\hat{\beta}$ estimation under H_0 and H_1.

the sum of the corresponding eigenvalues represents 99% of the trace of the covariance matrix $\mathbf{C}$.

First of all, Fig. 2 represents 50 out of 10000 Monte-Carlo estimation results for $\hat{\beta}$ (before saturation with 1), both under H_0 and H_1, as well as their estimated mean values. We can notice a rather good estimation for β under both hypotheses even if, as expected, the estimated values under H_1 are slightly bigger than the real value. Indeed, the mean is near $\bar{\hat{\beta}} \simeq 0.67$ for the V_4 target and near $\bar{\hat{\beta}} \simeq 0.66$ for the V_5 target respectively, whereas the real value is $\beta = 0.5$. We can notice that the estimation of β does not depend a lot on the target type. Indeed, the main part of the target's signature is eliminated by the projection onto the main background subspace $< \mathbf{U} >$. In the two cases, the estimations remain under 1 and increase the standard RX detector and improve detection. On the other hand, the estimated mean value under H_0 is near 1 as expected, producing almost no correction on RRX compared to RX, and preserving the false alarm rate. As a consequence, we can observe a detection improvement on the Receiver Operating Characteristic (ROC) plotted on Fig. 3. This gain is substantial as it reaches more than 2 decades of P_{fa} for a given P_d in such a case. Once again, we can notice that the performance's gains are almost the same whatever the target's type.

In the unlikely case where none power variation is induced by the presence of the target (pure additive model and $\beta = 1$), the performance of the proposed RRX is slightly lower than for the RX, as represented on Fig. 5. This is due to possible bad estimations for β as shown in Fig. 4, whereas it is known to be 1 using RX. This is somehow the price to be paid to get a more robust detector.

In order to determine from which background attenuation under H_1 it is worth using the proposed RRX rather than the standard RX, we now make vary β, as shown on Fig. 6. This experiment has been conducted with the V_5 target and we compute the P_{fa} gain for $P_d = 0.5$, as a figure of merit. This P_{fa} gain is defined as $G = 10 \log \frac{Pfa_{RX}(P_d=0.5)}{Pfa_{RRX}(P_d=0.5)}$. Then, we can observe that it is worth using the proposed RRX as soon as $\beta < 0.85$, as shown by a positive P_{fa} gain. Moreover, this gain can reach more than 20 dB when $\beta = 0.5$, as already noticed on fig. (3), whereas the loss is limited to a few dB when the additive model is almost valid ($0.85 < \beta < 1$).

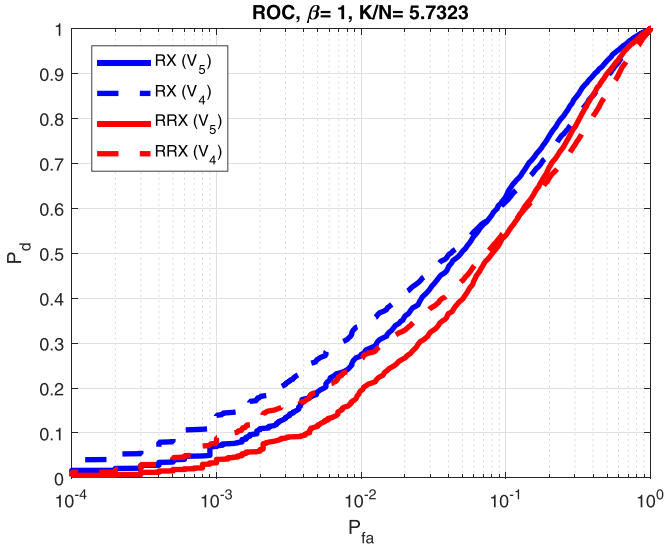

Fig. 5. Receiver Operating Characteristic comparison for $\beta = 1$.

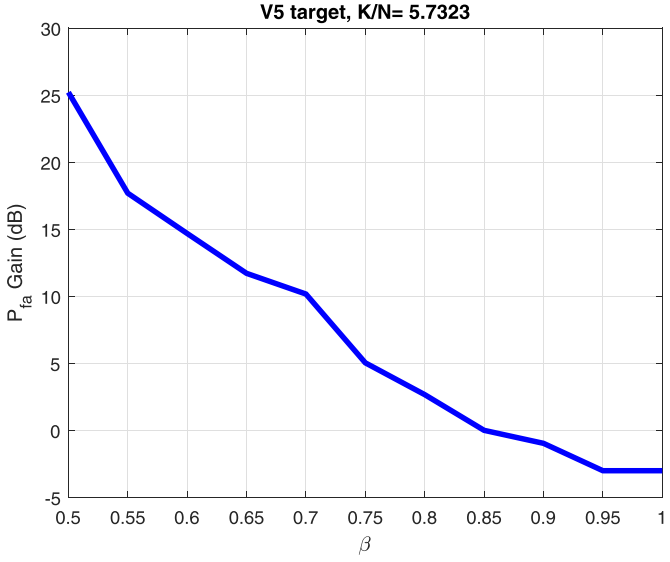

Fig. 6. P_{fa} Gain for $P_d = 0.5$.

4. Conclusions

In this paper, we considered Anomaly Detection using a more realistic model than most of the state of the art detectors, in the hyperspectral context. Indeed, we derived the GLRT based on the replacement model, and showed that it is a simple correction of the benchmark RX scheme. The main issue concerns the estimation of the background reduction that cannot be obtained through a standard ML approach. To circumvent this problem, we project the data onto the low-rank background subspace to minimizing the influence of the unknown target signature. Through numerical simulations, based on real life hyperspectral data, we show that this estimation is valid and that the improvement of the proposed RRX compared to the standard RX can be huge.

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.

CRediT authorship contribution statement

Francois Vincent: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Olivier Besson:** Methodology, Validation, Formal analysis, Writing - review & editing. **Stefania Matteoli:** 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, volume 5159, 2004.

[3] S. Michel, P. Gamet, M.-J. Lefevre-Fonollosa, Hypxim a hyperspectral satellite defined for science, security and defence users, Proceesings 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011.

[4] H. Kwon, N.M. Nasrabadi, Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery, IEEE Transactions Geoscience Remote Sensing 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., 2004. Orlando, FL, United States

[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 Transactions Geoscience Remote Sensing 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] I. S. Reed, X. Yu, Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution, IEEE Transactions Acoustics Speech Signal Processing 38 (10) (1990) 1760–1770.

[10] N. Billor, A.S. Hadi, P.F. Velleman, Bacon: blocked adaptive computationally efficient outlier nominators, Comput. Stat. Data Anal. 34 (3) (2000) 279–298.

[11] M.J. Carlotto, A cluster-based approach for detecting man-made objects and changes in imagery, IEEE Transactions Geoscience Remote Sensing 43 (2) (2005) 374387.

[12] 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, volume 4381, 2001, pp. 308–316.

[13] S. Matteoli, M. Diani, G. Corsini, A tutorial overview of anomaly detection in hyperspectral images, IEEE Aerospace Electronics Systems Magazine 25 (7) (2010) 5–27.

[14] S. Matteoli, M. Diani, J. Theiler, An overview of background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6) (2014) 2317–2336.

[15] L. Ren, L. Zhao, Y. Wang, A superpixel-based dual window RX for hyperspectral anomaly detection, IEEE Geosci. Remote Sens. Lett. 17 (7) (2020) 1233–1237.

[16] Y. Liang, P.P. Markopoulos, E.S. Saber, Subpixel target detection in hyperspectral images with local matched filtering in slic superpixels, 8th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolutions in Remote Sensing (WHISPERS 2016), 2016.

[17] S.M. Schweizer, J.M.F. Moura, Hyperspectral imagery: clutter adaptation in anomaly detection, IEEE Trans. Inf. Theory 46 (5) (2000) 1855–1871.

[18] S.M. Schweizer, J.M.F. Moura, Efficient detection in hyperspectral imagery, IEEE Trans. Image Process. 10 (4) (2001) 584–597.

[19] E.A. Ashton, Detection of subpixel anomalies in multispectral infrared imagery using an adaptive bayesian classifier, IEEE Trans. Geosci. Remote Sens. 36 (2) (1998) 506–517.

[20] D. Manolakis, R. Lockwood, T. Cooley, J. Jacobson, Is there a best hyperspectral detection algorithm? in: Proc. of SPIE, volume 7334, 2009.

[21] F. Vincent, O. Bessson, One-step generalized likelihood ratio test for subpixel target detection in hyperspectral imaging, IEEE Transactions Geoscience Remote Sensing 58 (6) (2020) 4479–4489.

[22] F. Vincent, O. Bessson, Generalized likelihood ratio test for modified replacement model in hyperspectral imaging detection, Signal Processing 174 (2020) 107643.

[23] A. Schaum, A. Stocker, Spectrally-selective target detection, in: Proceedings of ISSSR, volume 12, 1997, pp. 2015–2018.

[24] O. Besson, F. Vincent, Sub-pixel detection in hyperspectral imaging with elliptically contoured t-distributed background, Signal Processing 175 (2020) 107662.

[25] S. Matteoli, M. Diani, G. Corsini, Hyperspectral anomaly detection with kurtosis-driven local covariance matrix corruption mitigation, IEEE Geosci. Remote Sens. Lett. 8 (3) (2011) 532–536.

[26] W. Sun, C. Liu, J. Li, Y.M. Lai, W. Li, Low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery, J. Appl. Remote Sens. 8 (1) (2014) 1–18.

[27] Y. Zhang, B. Du, L. Zhang, S. Wang, A low-rank and sparse matrix decomposition-based mahalanobis distance method for hyperspectral anomaly detection, IEEE Trans. Geosci. Remote Sens. 54 (3) (2016) 1376–1389.

[28] L. Li, W. Li, Q. Du, R. Tao, Low-rank and sparse decomposition with mixture of gaussian for hyperspectral anomaly detection, IEEE Trans. Cybern. (2020) 1–10.

[29] N. Acito, S. Matteoli, A. Rossi, M. Diani, G. Corsini, Hyperspectral airborne viareggio 2013 trial data collection for detection algorithm assessment, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (6) (2016) 2356–2376.
[30] G. Ferrier, Evaluation of apparent surface reflectance estimation methodologies, Int. J. Remote Sens. 16 (1995) 2291–2297.
[31] 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.
[32] M. Shi, G. Healey, Hyperspectral texture recognition using a multiscale opponent representation, IEEE Transactions Geoscience Remote Sensing 41 (5) (2003) 1090–1095.