# ROBUST DETECTION USING THE SIRV BACKGROUND MODELLING FOR HYPERSPECTRAL IMAGING

*J.P. Ovarlez*[1,3], *S.K. Pang*[2], *F. Pascal*[3], *V. Achard*[1] *and T.K. Ng*[2]

[1] : French Aerospace Lab, ONERA, France, *jean-philippe.ovarlez@onera.fr*, *veronique.achard@onera.fr*

[2] : DSO National Laboratories, Singapore, *pszekim@dso.org.sg*, *nteckkhi@dso.org.sg*

[3] : SONDRA, Supélec, France, *frederic.pascal@supelec.fr*

## ABSTRACT

This paper deals with hyperspectral detection in impulsive and/or non homogeneous background contexts. In hyperspectral imaging applications, the detection performance of the detectors (target detection or anomaly detection like Mahalanobis distance) is typically evaluated on Gaussian assumption. However, it is well known that hyperspectral imaging data exhibit spatial heterogeneity and non-Gaussian behavior leading to a poor performance for all the conventional Gaussian detectors. Many works have been already derived in the context of radar detection in non-homogeneous and non-Gaussian clutter. These works can be easily extended in the context of hyperspectral detection. The aim of this paper is twofold. In the context of Spherically Invariant Random Vectors (SIRV) modeling for the background, we recall some properties of different non-Gaussian detectors built with a nice and robust estimate of the background Covariance Matrix. Secondly, we present some results on regulation of false alarm obtained on experimental background hyperspectral data. These results demonstrate the interest of the proposed detection scheme, and show an excellent correspondence between experimental and theoretical results.

***Index Terms*—** Hyperspectral Imaging, Anomaly Detection, SIRV, non-Gaussian process.

## 1. INTRODUCTION

Anomaly detection and detection of targets or activity such as chemical plumes, aerosols, vehicles, anomalous targets, arise in many military and civilian applications [1]. Hyperspectral imaging sensors provide 2D spatial image data containing spectral information. This information can be used to address such detection tasks. Hyperspectral imaging sensors measure the radiance of the materials within each pixel area at a very large number spectral bands.

It is often assumed that signals, interferences, noises, background are Gaussian stochastic processes. Indeed, this assumption makes sense in many applications. In these contexts, Gaussian models have been widely investigated in the framework of Statistical Estimation and Detection Theory. They have led to appealing and well known algorithms such as the Matched Filter and its adaptive variants in radar detection [2, 3]. The mathematical framework for the design and evaluation of detection algorithms is provided by the well known binary hypothesis testing procedure. The detection problem is typically formulated as a binary hypothesis test with two competing hypotheses: background only or target and background. In practice the background statistics are unknown and have to be estimated from the data. Generally, the statistical parameters (covariance matrix, mean, ...) of the background can be estimated by using all pixels within an area of interest. The size of the area has to be chosen large enough to ensure the invertibility of the covariance matrix and small enough to justify both spectral homogeneity (stationarity) and spatial homogeneity. Since the two hypotheses contain unknown parameters (for example, the covariance matrix of the background) that have to be estimated from the data, the detector has to be adaptive, and it is usually designed by using the Generalized-Likelihood-Ratio Test (GLRT) approach.

However, such widespread techniques are sub-optimal when the noise process is a non-Gaussian stochastic process. Therefore, non-Gaussian noise modeling has gained much interest these last decades and presently leads to active researches in the literature. In radar applications, experimental clutter measurements, made by institutions like MIT [4], showed that these data are not correctly described by Gaussian statistical models. In hyperspectral imaging, the actual response of a detector to the background pixels differs from the theoretically predicted distribution for Gaussian backgrounds. In fact, as stated in [5, 6], the empirical distribution usually has heavier tails compared to the theoretical distribution, and these tails strongly influence the observed false-alarm rate of the detector.

One of the most general and elegant impulsive noise model is provided by the so-called *Spherically Invariant Random Vectors* (SIRV). Indeed, these processes encompass a large number of non Gaussian distributions, included of course Gaussian processes. SIRV and their variants have been used in various problems of radar clutter echoes. A SIRV is a compound process, it is the product of a Gaussian multivariate process and the square root of a non-negative random scalar variable called the *texture*. Thus, the SIRV is fully characterized by the texture (representing an unknown intensity) and the unknown covariance matrix of the Gaussian vector. One of the major challenging difficulties in SIRV modeling, is to estimate these two unknown quantities. For example, the classical Sample Covariance Matrix used in adaptive detection in Gaussian noise is not at all the best estimate and does not correspond to the Maximum Likelihood estimator. These problems have been investigated in [7] for the texture estimation while [8] and [9] have proposed different estimates for the covariance matrix. A complete statistical analysis of these covariance matrix estimates has been realized in [10].

Since the noise parameters estimates are fully characterized, it is now a major issue to use them in the hyperspectral detection process. The first contribution of this paper is the study of adaptive non-Gaussian detector built with an improved covariance matrix estimate, the Fixed Point Covariance Matrix. Constant False Alarm Rate (CFAR) properties allow this detector to be independent of nuisance parameters. Then, a theoretical relationship for false alarm regulation is introduced and validated on real hyperspectral data. This application allows to study the detector behavior in realistic scenario.

## 2. PROBLEM FORMULATION

In this section, we introduce the SIRV noise model under study and the associated adaptive detector built with the Fixed Point estimate. In the following, $H$ denotes the conjugate transpose operator, $E(.)$ stands for the statistical mean.

### 2.1. Statistical Framework

The basic problem of detecting a complex signal corrupted by an additive SIRV noise $\mathbf{c}$ in a $m$-dimensional complex vector $\mathbf{y}$ can be stated as the following binary hypothesis test:

$$\begin{cases} H_0 : \mathbf{y} = \mathbf{c} & \mathbf{y}_i = \mathbf{c}_i \quad i = 1, \dots, K \\ H_1 : \mathbf{y} = \mathbf{s} + \mathbf{c} & \mathbf{y}_i = \mathbf{c}_i \quad i = 1, \dots, K \end{cases} \tag{1}$$

where $\mathbf{y}$ is the cell under test and where the $\mathbf{y}_i$'s are $K$ signal-free independent measurements, traditionally called the secondary data, used to estimate the background covariance matrix. Generally, these $K$ secondary data are collected using a 2D sliding window (mask) centered on the cell under test. Under hypothesis $H_1$, it is assumed that the observed data consists in the sum of a signal $\mathbf{s} = \alpha\,\mathbf{p}$ and clutter $\mathbf{c}$, where $\mathbf{p}$ is a perfectly known complex steering vector (characterizing for example the spectral material to detect) and $\alpha$ is the signal amplitude.

Let us recap some SIRV theory results. A noise modeled as a SIRV is a non-homogeneous Gaussian process with random power. More precisely, a SIRV [11] is the product $\mathbf{c} = \sqrt{\tau}\,\mathbf{x}$ of the square root of a positive random variable $\tau$ (*texture*) and a $m$-dimensional independent complex circular Gaussian vector $\mathbf{x}$ (*speckle*) with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{M} = \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu})\,(\mathbf{x} - \boldsymbol{\mu})^H]$ with normalization $\mathrm{Tr}(\mathbf{M}) = m$.The SIRV Probability Density Function (PDF) expression is given by:

$$p_m(\mathbf{c}) = \int_0^{+\infty} \frac{1}{(\pi\,\tau)^m\,|\mathbf{M}|}\,\exp\left(-\frac{(\mathbf{c} - \boldsymbol{\mu})^H\,\mathbf{M}^{-1}\,(\mathbf{c} - \boldsymbol{\mu})}{\tau}\right) p(\tau)\,d\tau\,.$$

The SIRV family encompasses an infinity of distributions, notably the Gaussian one, the Weibull distribution, the K-distribution or the T-distribution.

The SIRV modelling has been used to derive several Generalized Likelihood Ratio Tests like the GLRT-Linear Quadratic (GLRT-LQ) in [12, 13] defined by

$$\Lambda(\mathbf{M}, \boldsymbol{\mu}) = \frac{|\mathbf{p}^H\,\mathbf{M}^{-1}\,(\mathbf{y} - \boldsymbol{\mu})|^2}{(\mathbf{p}^H\,\mathbf{M}^{-1}\,\mathbf{p})((\mathbf{y} - \boldsymbol{\mu})^H\,\mathbf{M}^{-1}\,(\mathbf{y} - \boldsymbol{\mu}))} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda\,, \tag{2}$$

where $\mathbf{p}$ is the spectral steering vector, $\mathbf{y}$ the observed vector and $\lambda$ the detection threshold associated to this detector. Note that the mean $\boldsymbol{\mu}$ is generally omitted in radar detection (and therefore not estimated) as the noise is always zero mean. So, in hyperspectral Imaging context, as the data represent intensity and are positive, we have to estimate it, jointly with the covariance matrix $\mathbf{M}$.

In many problems, non-Gaussian noise can be characterized by SIRVs but the covariance matrix $\mathbf{M}$ is generally not known and an estimate $\widehat{\mathbf{M}}$ is required. It can be noted here that the classical Sample Covariance Matrix is not at all a good solution. The next section is devoted to an adaptive GLRT built from an Approximate Maximum Likelihood (AML) estimate of the SIRV background covariance matrix.

### 2.2. The ANMF with the Fixed Point Matrix Estimate

Let us now present the adaptive GLRT, i.e. the adaptive version of (2) and called the Adaptive Normalized Matched Filter (ANMF) or ACE (Adaptive Cosine Estimator), used for the detection problem under study:

$$\Lambda(\widehat{\mathbf{M}}, \widehat{\boldsymbol{\mu}}) = \frac{|\mathbf{p}^H\,\widehat{\mathbf{M}}^{-1}\,(\mathbf{y} - \widehat{\boldsymbol{\mu}})|^2}{(\mathbf{p}^H\,\widehat{\mathbf{M}}^{-1}\,\mathbf{p})((\mathbf{y} - \widehat{\boldsymbol{\mu}})^H\,\widehat{\mathbf{M}}^{-1}\,(\mathbf{y} - \widehat{\boldsymbol{\mu}}))} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda\,. \tag{3}$$

Moreover, for $\boldsymbol{\mu} = \mathbf{0}$, Conte and Gini in [8, 9] have shown that an Approximate Maximum Likelihood (AML) estimate $\widehat{\mathbf{M}}$ of $\mathbf{M}$ is a solution of the following equation:

$$\widehat{\mathbf{M}} = \frac{m}{K}\sum_{i=1}^{K} \frac{\mathbf{c}_i\,\mathbf{c}_i^H}{\mathbf{c}_i^H\,\widehat{\mathbf{M}}^{-1}\,\mathbf{c}_i}. \tag{4}$$

Existence and uniqueness of the solution to the above equation, denoted $\widehat{\mathbf{M}}_{FP}$ have already been investigated in [14] while its performance analysis has been studied by [10] in which it was shown that the Fixed Point estimate is the covariance matrix estimate that has a better match to the problem under study thanks to its statistical performance and its easy implementation and practical use. Eq. (4) obviously implies that $\widehat{\mathbf{M}}_{FP}$ is independent of the $\tau_i$'s. The main results of the statistical properties of $\widehat{\mathbf{M}}_{FP}$ are recapped: $\widehat{\mathbf{M}}_{FP}$ is a consistent and unbiased estimate of $\mathbf{M}$; its asymptotic distribution is Gaussian and its covariance matrix is fully characterized in [15]; its asymptotic distribution is the same as the asymptotic distribution of a Wishart matrix with $m\,N/(m+1)$ degrees of freedom.

When the noise is not centered, as in hyperspectral imaging, the joint estimation of $\mathbf{M}$ and $\boldsymbol{\mu}$ leads to [16]:

$$\hat{\mathbf{M}}_{FP} = \frac{1}{K} \sum_{k=1}^{K} \frac{(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})\,(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})^H}{(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})^H\, \hat{\mathbf{M}}_{FP}^{-1}\,(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})}\,, \tag{5}$$

and

$$\widehat{\boldsymbol{\mu}} = \frac{\displaystyle\sum_{k=1}^{K} \frac{\mathbf{c}_k}{(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})^H\, \hat{\mathbf{M}}_{FP}^{-1}\,(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})}}{\displaystyle\sum_{k=1}^{K} \frac{1}{(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})^H\, \hat{\mathbf{M}}_{FP}^{-1}\,(\mathbf{c}_k - \widehat{\boldsymbol{\mu}})}}\,. \tag{6}$$

These two estimates given by implicit equations (Fixed Point Equation) can be easily computed using a recursive approach. In the section dealing with applications to experimental hyperspectral data, we will use the GLRT-FP $\hat{\Lambda}(\widehat{\mathbf{M}}_{FP}, \widehat{\boldsymbol{\mu}})$ as detector. This detector has essential CFAR properties like texture-CFAR (independent of the distribution of $\tau$), matrix-CFAR (independent of $\mathbf{M}$) and mean-CFAR (independent of $\boldsymbol{\mu}$). One of the first deduction of previous results is that regardless of the SIRV used, for different distributions of the texture and for different covariance matrices, the resulting GLRT $\Lambda(\widehat{\mathbf{M}}_{FP}, \widehat{\boldsymbol{\mu}}))$ follows the same distribution. This is of a major interest in areas of background transition like for example, in coastal areas (ground and sea) or at the edge of forests (fields and trees) because the detector should be insensitive to the different clutter areas.

A theoretical relationship between the detection threshold $\lambda$ and the Probability of False Alarm (PFA) $P_{fa} = \mathbb{P}(\Lambda > \lambda | H_0)$ has been established:

$$P_{fa} = (1-\lambda)^{a-1}\, {}_2F_1(a, a-1; b-1; \lambda)\,, \tag{7}$$

where $a = \dfrac{m}{m+1}K - m + 2$, $b = \dfrac{m}{m+1}K + 2$ and ${}_2F_1$ is the Hypergeometric function.

Note that the previous Pfa-threshold relationship (7) has been derived assuming radar data being complex and is not valid for real data. As the hyperspectral data are positive and real, they have been passed through an Hilbert filter to render them complex. The following section presents some results relative to the regulation of false alarm obtained on experimental data.

## 3. DETECTION RESULTS ON EXPERIMENTAL HYPERSPECTRAL DATA

The experimental set of data was provided by DSO National Laboratories (see figure 1). The figure 2 shows the regulation of the false alarm for the conventional Adaptive Matched Filter built with the classical Sample Covariance Matrix. The figure 3 shows the results obtained with the ANMF or ACE given in (3) built with the Fixed Point and the mean given respectively in (5) and (6). These preliminary results show a better regulation for the proposed detection scheme than the conventional one.

## 4. CONCLUSIONS

The SIRV modelling as pointed out in [5, 6] is shown to be very interesting when dealing with heterogeneity and/or non Gaussian data. All the previous works that have been done in the context of radar detection can be applied successfully on hyperspectral imagery for the purpose of detection. The ACE detector built not with the conventional SCM but with the proposed Fixed Point estimate is shown to be SIRV CFAR. These preliminary results have to be of course evaluated further. They have been shown to have a good potential for target detection in hyperspectral sensing. These works[1] can also be used for anomaly detection purpose.

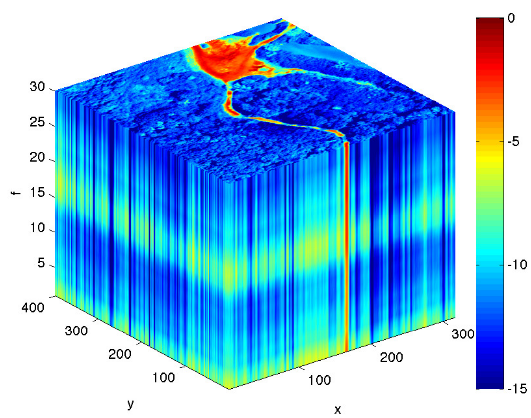


**Fig. 1**. Normalized hyperspectral data set

[1]The authors would like to thank french DGA for its financial support.

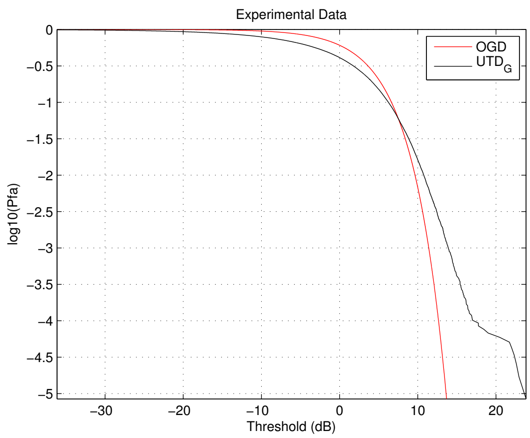


**Fig. 2**. Pfa-Threshold behaviors for the Adaptive Matched Filter applied on data (in black) and for the theoretical one (in red) for uniform spectral target. The spatial window used is here $13 \times 13$ with one guard cell around the cell under test ($K = 13^2 - 9$ for $m = 20$ wavelengths)

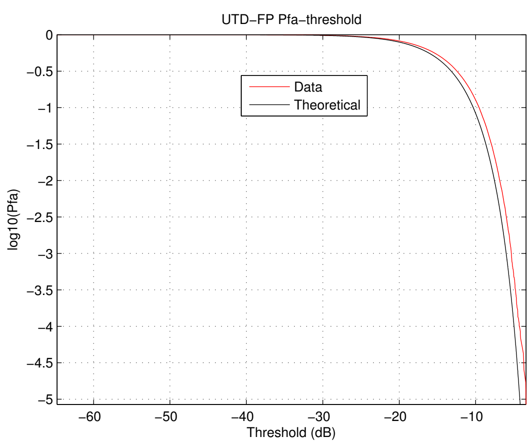


**Fig. 3**. Pfa-Threshold behaviors for the Adaptive Normalized Matched Filter applied on data (in red) and for the theoretical one (in black) given by (7) for uniform spectral shape target. The spatial window used is here $13 \times 13$ ($K = 13^2 - 9$ for $m = 20$ wavelengths)

## 5. REFERENCES


[1] S. Matteoli, M. Diani and G. Corsini, "A Tutorial Overview of Anomaly Detection in Hyperspectral Images", IEEE AES Magazine, Vol.25, No.7, July 2010.

[2] E.J. Kelly, "An Adaptive Detection Algorithm", *IEEE Trans. on AES*, vol. 23, no. 1, pp. 115-127, Nov. 1986.

[3] S. Kraut, L.L. Scharf and L.T. Mc Whorter, "Adaptive Subspace Detectors", *IEEE Trans. on SP*, vol. 49, no. 1, pp. 1-16, Jan. 2001.

[4] J.B. Billingsley, "Ground Clutter Measurements for Surface-Sited Radar", *Technical Report 780, MIT*, Feb. 1993.

[5] D. Manolakis, D. Marden, and G.A. Shaw, "Hyperspectral Image Processing for Automatic Target Detection Applications", *Lincoln Laboratory Journal*, Vol.14, No.1, 2003.

[6] D. Manolakis and G. Shaw, "Detection Algorithms for Hyperspectral Imaging Application," IEEE Signal Processing Magazine, pp.29-43, January 2002.

[7] E. Jay, J.P. Ovarlez, D. Declercq and P. Duvaut, BORD: Bayesian Optimum Radar Detector, *Signal Processing*, vol. 83, no. 6, pp. 1151–1162, June 2003.

[8] F. Gini and M. V Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter", *Signal Processing*, *Signal Processing*, vol. 82, no. 12, pp. 1847–1859, Dec. 2002.

[9] E. Conte, A. De Maio and G. Ricci, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection", *IEEE Trans. on SP*, vol. 50, no. 8, pp. 1908–1915, Aug. 2002.

[10] F. Pascal, P. Forster, J.P. Ovarlez and P. Larzabal, "Performance Analysis of Covariance Matrix Estimates in an Impulsive Noise", *IEEE Trans. on SP*, Vol.56, No.6, pp.2206-2217, Jun. 2008.

[11] K. Yao "A Representation Theorem and its Applications to Spherically Invariant Random Processes", *IEEE Trans. on IT*, vol. 19, no. 5, pp. 600–608, September 1973.

[12] E. Conte, M. Lops and G. Ricci, "Asymptotically Optimum Radar Detection in Compound-Gaussian Clutter", *IEEE Trans. on AES*, vol. 31, no. 2, pp. 617–625, April 1995.

[13] F. Gini, "Sub-Optimum Coherent Radar Detection in a Mixture of K-Distributed and Gaussian Clutter", *IEE Proc. Radar, Sonar Navig.*, vol. 144, no. 1, pp. 39–48, February 1997.

[14] F. Pascal, Y. Chitour, J.P. Ovarlez, P. Forster and P. Larzabal, "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise: Existence and Algorithm Analysis", *IEEE Trans. on SP*, Vol. 56, No.1, pp.34-48, Jan. 2008

[15] F. Pascal, P. Forster, J.P. Ovarlez and P. Larzabal, "Theoretical analysis of an improved covariance matrix estimator in non-Gaussian Noise", *IEEE-ICASSP*, Philadelphia, PA, USA, March 2005.

[16] M. Bilodeau and D. Brenner, Theory of Multivariate Statistics, Springer Verlag, New York, 1999