ROBUST DETECTION USING *M*-ESTIMATORS FOR HYPERSPECTRAL IMAGING

J. Frontera-Pons[1], *M. Mahot*[1], *J.P. Ovarlez*[1,2], *F. Pascal*[1], *J. Chanussot*[3]

ABSTRACT

Hyperspectral data have been proved not to be multivariate normal but long tailed distributed. In order to take into account these features, the family of elliptical contoured distributions is proposed to describe noise statistical behavior. Although non-Gaussian models are assumed for background modeling and detectors design, the parameters estimation is still performed using classical Gaussian based estimators; as for the covariance matrix, generally determined according to the SCM approach. We discuss here the class of *M*-estimators as a robust alternative for background statistical characterization and highlight their outcome when used in an adaptive GLRT-LQ detector.

Index Terms— hypespectral imaging, target detection, elliptical distributions, M-estimators

1. INTRODUCTION

The basic idea in Hyperspectral imaging (HSI) extends from the fact that for any given material, the amount of radiation emitted varies with wavelength. Hyperspectral imaging sensors measure the radiance of the materials within each pixel area at a very large number of contiguous spectral bands and provide image data containing both spatial and spectral information. Hyperspectral target detection and anomaly detection may be used to locate targets that generally cannot be resolved by spatial resolution [1]. A broad variety of applications take advantage of these techniques for detection purposes.

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 binary hypothesis testing procedure. However, such widespread techniques are sub-optimal when the noise 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. It has to be stressed that the grade on the detector performance relies on the fit of the background into the assumed statistical model. As stated in [4] and [5], 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 investigated models for background statistics characterization is the family of elliptical contoured distributions. Indeed, these processes encompass a large number of non-Gaussian distributions, included of course multivariate normal distribution.

Generally, the statistical parameters (covariance matrix, mean vector) are unknown and need to be estimated from the data. Even when clutter is assumed to be non-Gaussian, the classical Sample Covariance Matrix (SCM) is used in several adaptive detection methods. But it does not correspond to the Maximum Likelihood estimator and results in poor performance of the detectors. These problems have been discussed in [6]. While different proposals for matrix estimates in non-Gaussian environment can be found in [7, 8].

For data in $\mathbb{R}^k$ robust alternatives for the sample covariance estimate are the *M*-estimators of multivariate location vector and scatter matrix. We study in this paper an adaptive non-Gaussian detector built with these improved estimators. Constant False Alarm Rate (CFAR) is pursued to allow the detector independence of nuisance parameters and false alarm regulation.

2. PROBLEM FORMULATION

In this section, we present the class of elliptically contoured (EC) distributions. They provide a multivariate location-scatter family of distributions that primarily serve as long tailed alternatives to the multivariate normal model. They are proved to represent a more accurate characterization of HSI data than models based on the multivariate Gaussian assumption.

2.1. Statistical Framework

The detection problem is typically formulated as a binary hypothesis test with two competing hypotheses: background only or target and background. We attempt to determine the occurrence of a target, i.e. the presence of a complex signal $\mathbf{s}$ corrupted by an additive noise $\mathbf{c}$. Symbolically, we aim to

[1]SONDRA Research Alliance, Supélec, France, [2]French Aerospace Lab, ONERA DEMR/TSI, France, [3]GIPSA-Lab, Grenoble Institute of Technology, France

distinguish between:

$$\begin{cases} H_0 : \mathbf{y} = \mathbf{c} & \mathbf{y_i} = \mathbf{c_i} \quad i = 1, \ldots, K \\ H_1 : \mathbf{y} = \mathbf{s} + \mathbf{c} & \mathbf{y_i} = \mathbf{c_i} \quad i = 1, \ldots, K \end{cases}$$

where $\mathbf{y}$ is the cell under test and $\mathbf{y_i}$ are K signal free independent measurements, usually referred as secondary data, used to estimate the mean and the covariance of the background. Generally these K secondary data are collected using a spatial sliding window (mask) centered on the cell under test. The size of the mask is chosen large enough to ensure the invertibility of the covariance matrix and small enough to justify both spatial and spectral homogeneity. Under hypothesis H_1, it is assumed that the observed data is composed by a signal $\mathbf{s} = \alpha\mathbf{p}$ and clutter $\mathbf{c}$, combined in an additive manner; where $\mathbf{p}$ is a complex steering vector (supposed perfectly known), characterizing the spectral signature of the material intended to detect and α the signal amplitude.

2.2. Elliptical distribution

A m-dimensional random complex vector $\mathbf{y}$ is said to have a complex elliptical distribution if its probability density function (PDF) has the form

$$f_{\mathbf{y}}(\mathbf{y}) = |\boldsymbol{\Sigma}|^{-1} h_{\mathbf{y}}((\mathbf{y} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})), \qquad (1)$$

where H denotes the conjugate transpose operator and $h_{\mathbf{y}} : [0, \infty) \to [0, \infty)$ is a positive, monotonically decreasing function, $\boldsymbol{\mu}$ is the mean vector and $\boldsymbol{\Sigma}$ is the scatter matrix. The function $h_{\mathbf{y}}$,usually called density generator, is assumed to be only approximately known. Note that it produces density contours corresponding to elliptical surfaces. If the second moments exist, then $\boldsymbol{\Sigma}$ reflects the structure of the covariance matrix of the elliptically distributed random vector $\mathbf{y}$, i.e. the covariance matrix equates the scatter matrix up to a scaling constant. We shall denote this complex elliptical distribution by $EC(\boldsymbol{\mu}, \boldsymbol{\Sigma}, h)$. It is worth pointing that the EC class includes an infinity of distributions, notably the Gaussian one, multivariate t distribution or multivariate Cauchy.

3. M- ESTIMATORS

Let $(\mathbf{c_1}, \ldots, \mathbf{c_K})$ be a K-sample of m-dimensional complex independent vectors with $\mathbf{c_i} \sim EC(\boldsymbol{\mu}, \boldsymbol{\Sigma}, h)$, $i = 1, \ldots, K$. The complex *M*-estimators of location and scatter are defined as the joint solutions to the estimating equations:

$$\begin{aligned} \widehat{\boldsymbol{\mu}} &= \frac{\sum_{n=1}^{K} u_1(t_n)\mathbf{c_n}}{\sum_{n=1}^{K} u_1(t_n)} \\ \hat{\mathbf{M}} &= \frac{1}{K} \sum_{n=1}^{K} u_2(t_n^2)(\mathbf{c_n} - \widehat{\boldsymbol{\mu}})(\mathbf{c_n} - \widehat{\boldsymbol{\mu}})^H \end{aligned} \qquad (2)$$

where $t_n = \left((\mathbf{c_n} - \widehat{\boldsymbol{\mu}})^H \hat{\mathbf{M}}^{-1} (\mathbf{c_n} - \widehat{\boldsymbol{\mu}})\right)^{1/2}$ and u_1, u_2 are two weighting functions on the quadratic form t_n. Note t_n^2 is in fact, the widely used Mahalanobis distance. *M*-estimators have first been studied in the real case, defined as solution of (2) with real samples. Existence and uniqueness have been proved in the real case, provided functions u_1, u_2 satisfy a set of general assumptions stated by Maronna [9] . Olilla has shown in [10] that these conditions hold also in the complex case. *M*-estimators are particularly suited for estimating the mean vector and the scatter matrix of an elliptical population. When dealing with heavy tailed clutter models, as in HSI, the use of robust estimates decreases the impact of highly impulsive samples and possible outliers.
Remark that if u_1 and u_2 are chosen to be constant and equal to one, the arising estimators correspond to the Sample Mean Vector and Sample Covariance Matrix respectively. They are indeed the the Maximum Likelihood estimators when Gaussian distributions are considered.

3.1. The Fixed Point estimates

According to the Fixed point approach, the joint estimation of $\mathbf{M}$ and $\boldsymbol{\mu}$ leads to [11]:

$$\hat{\mathbf{M_{FP}}} = \frac{m}{K} \sum_{k=1}^{K} \frac{(\mathbf{c_k} - \widehat{\boldsymbol{\mu}})(\mathbf{c_k} - \widehat{\boldsymbol{\mu}})^H}{(\mathbf{c_k} - \widehat{\boldsymbol{\mu}})^H \hat{\mathbf{M}}_{FP}^{-1} (\mathbf{c_k} - \widehat{\boldsymbol{\mu}})} \qquad (3)$$

and

$$\widehat{\boldsymbol{\mu}} = \frac{\sum_{k=1}^{K} \frac{\mathbf{c_k}}{(\mathbf{c_k} - \widehat{\boldsymbol{\mu}})^H \hat{\mathbf{M}}_{FP}^{-1} (\mathbf{c_k} - \widehat{\boldsymbol{\mu}})}}{\sum_{k=1}^{K} \frac{1}{(\mathbf{c_k} - \widehat{\boldsymbol{\mu}})^H \hat{\mathbf{M}}_{FP}^{-1} (\mathbf{c_k} - \widehat{\boldsymbol{\mu}})}} \qquad (4)$$

Obtained when choosing $u_1(s) = s^{-1}$ and $u_2(s) = ms^{-1}$. For the matrix estimate, existence and uniqueness have been established in [12]. Although the proof for simultaneous scatter and location estimates is still an open question, they have been found to be useful and reliable for elliptical contours estimation parameters because of its easy implementation. They are specified by implicit equations and can be easily computed using a recursive algorithm. We refer to [13] for a detailed performance analysis of the Fixed Point covariance matrix estimate. The main results of the statistical properties of the $\hat{\mathbf{M_{FP}}}$ are summarized: $\hat{\mathbf{M_{FP}}}$ is a consistent and unbiased estimate of $\mathbf{M}$; its asymptotic distribution is Gaussian and its covariance matrix is fully characterized in [14]; its asymptotic distribution is the same as the asymptotic distribution of a Wishart matrix with $mN/(m+1)$ degrees of freedom.

3.2. The Huber's *M*-estimates

Using the well-known Huber's ψ function [15] defined as,

$$\psi_k(s) = min(s, k) \qquad (5)$$

with $k > 0$. One can obtain Huber's *M*-estimator by taking $u_1(s) = \psi_k(s)/s$ and $u_2(s) = \psi_{k^2}/s$. We remark that the Huber function can be seen as a mix between the Fixed Point estimate and the conventional SCM estimate. Extreme values of t_n^2 outside $[0, k^2]$ are strongly attenuated by the $1/s$ decreasing function (as for the Fixed Point) while normal values below k^2 are uniformly kept (SCM behavior).

4. THE ANMF BUILT WITH THE M-ESTIMATORS

Different types of adaptive non-Gaussian detectors may be derived for target enhancement purposes. We focus here on the study of the GLRT-Linear Quadratic [16], also known as Adaptive Cosine Estimate built with the differents *M*-estimators presented above,

$$\Lambda(\hat{\mathbf{M}}, \widehat{\boldsymbol{\mu}}) = \frac{|\mathbf{p^H}\hat{\mathbf{M}}^{-1}(\mathbf{y}-\widehat{\boldsymbol{\mu}})|^2}{(\mathbf{p^H}\hat{\mathbf{M}}^{-1}\mathbf{p})((\mathbf{y}-\widehat{\boldsymbol{\mu}})^H\hat{\mathbf{M}}^{-1}(\mathbf{y}-\widehat{\boldsymbol{\mu}}))} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \tag{6}$$

where $\mathbf{p}$ is the spectral steering vector, $\mathbf{y}$ the cell under test and λ the decision threshold.Note that the mean $\hat{\boldsymbol{\mu}}$ is generally omitted in radar detection (and therefore not estimated) as the noise is always zero mean. So, in hyperspectral imaging, as the data represent intensity values and are positive, we need to estimate it, jointly with the covariance matrix $\mathbf{M}$. Used with the Fixed Point estimate, this detector is particularly interesting because of its CFAR matrix properties. Hence, the detector GLRT $\Lambda(\hat{\mathbf{M}_{\mathbf{FP}}}, \hat{\boldsymbol{\mu}})$ behaves according to the same distribution for different covariance matrices.

4.1. Detector performance

The performance analysis has been realized over the data set provided by DSO National Laboratories, the normalized hypercube is shown in figure 1. The resulting ROC curves (Receiver Operating Characteristic) compare the output of the detector built with the Fixed Point estimates, the Huber *M*-estimators and the classical SCM. The test conducted consists in placing an artificial target with a fixed SNR through each pixel of the image. For all the possible threshold values, both probability of false alarm and probability of detection are computed. The outcome is illustrated in figure 2.
These preliminary results show the improvement in performance introduced by the use of *M*-estimators regarding the conventional SCM. The desired robustness properties previously mentioned lead to a higher P_d for small values of the P_{fa}.

4.2. False Alarm Regulation

A theoretical relationship between the detection threshold λ and the Probability of False Alarm $P_{fa} = \mathbb{P}(\Lambda > \lambda | H_0)$ has been established in [17]:

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

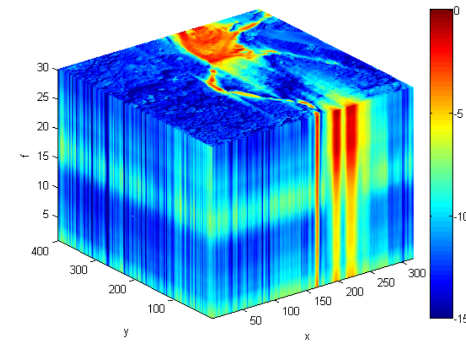

Fig. 1. Normalized data set.

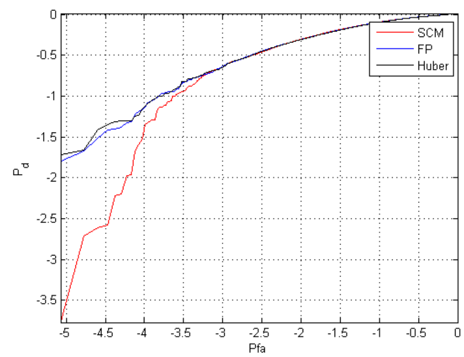

Fig. 2. ROC curves depicting the performance of the detector built with the SCM (in red), the Fixed Point (in blue) and the Huber type (in black) estimates. Probabilities are given in $log10$ scale.

where $a = \frac{m}{m+1}K - m + 2$, $b = \frac{m}{m+1}K + 2$ and ${}_2F_1$ is the Hypergeometric function. This expression holds when $\boldsymbol{\mu}$ is completely known and can be removed from the data. Figure 3 exhibits the regulation of the false alarm for the detector built with the classical SCM and figure 4 the corresponding results obtained with the Fixed Point estimates. The gap evidenced in the figure for the Fixed Point, is due to the joint estimation of $\mathbf{M}$ and $\boldsymbol{\mu}$, since $\Lambda(\hat{\mathbf{M}_{\mathbf{FP}}}, \hat{\boldsymbol{\mu}})$ is no longer unaffected by the distribution of $\hat{\boldsymbol{\mu}}$.
Note that the previous Pfa-threshold has been derived assuming radar data being complex and is not valid for real data. As the hyperspectral data are real and positive, they have been passed through an Hilbert filter to render them complex.

5. CONCLUSIONS

Using elliptical distributions for background modeling allows for heterogeneity consideration in non-Gaussian environment. We have proposed different estimators for statistical

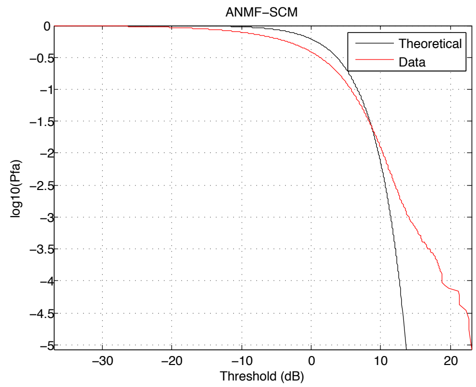

Fig. 3. Pfa-threshold. Adaptive Normalized Matched Filter built with the SCM

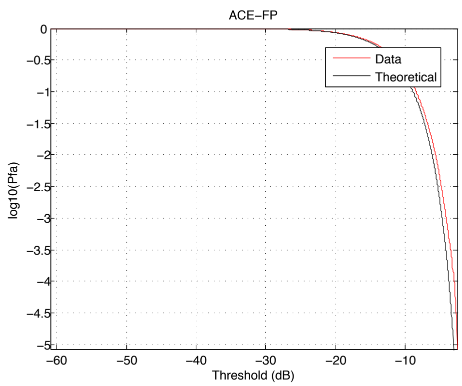

Fig. 4. Pfa-threshold. ACE detector built with Fixed Point Estimates

characterization of the clutter. In particular, the class of M-estimators, which are specially appropriate to the addressed problem. We have recalled Fixed Point estimates and introduced the Huber type M-estimators emphasizing their statistical properties. When building the ACE detector with these newfangled estimates, its capacity for a more accurate false alarm regulation is pointed, as well as a better performance in probability detection terms.

6. REFERENCES

[1] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 25, no. 7, pp. 5–28, 2010.

[2] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans.-AES*, vol. 23, no. 1, pp. 115–127, November 1986.

[3] S. Kraut, L. L. Scharf, and L. T. Mc Whorter, "Adaptive subspace detectors," *IEEE Trans.-SP*, vol. 49, no. 1, pp. 1–16, January 2001.

[4] D. Manolakis, D. Marden, and G.A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 79–116, 2003.

[5] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 29–43, 2002.

[6] 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.

[7] F. Gini and M. V. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Processing, special section on SP with Heavy Tailed Distributions*, vol. 82, no. 12, pp. 1847–1859, December 2002.

[8] 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.-SP*, vol. 50, no. 8, pp. 1908–1915, August 2002.

[9] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *Annals of Statistics*, vol. 4, no. 1, pp. 51–67, January 1976.

[10] E. Ollila and V. Koivunen, "Influence functions for array covariance matrix estimators," *Proc. IEEE Workshop on Statistical Signal Processing (SSP),ST Louis, MO*, pp. 445–448, October 2003.

[11] D.E. Tyler, "A distribution-free m-estimator of multivariate scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.

[12] 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.-SP*, vol. 56, no. 1, pp. 34–48, January 2008.

[13] F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Trans.-SP*, vol. 56, no. 6, pp. 2206–2217, June 2008.

[14] F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal, "Theoretical analysis of an improved covariance matrix estimator in non-gaussian noise," in *Proc. IEEE-ICASSP*, Philadelphia, March 2005, vol. IV, pp. 69–72.

[15] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.

[16] F. Gini, "Sub-optimum coherent radar detection in a mixture of k-distributed and gaussian clutter," *IEE Proc. Radar, Sonar and Navigation*, vol. 144, no. 1, pp. 39–48, February 1997.

[17] JP Ovarlez, SK Pang, F. Pascal, V. Achard, and TK Ng, "Robust detection using the sirv background modelling for hyperspectral imaging," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. IEEE, 2011, pp. 4316–4319.