Performance Analysis of Robust Detectors for Hyperspectral Imaging

Joana Frontera-Pons, Jean-Philippe Ovarlez, Frédéric Pascal, Jocelyn
Chanussot

PERFORMANCE ANALYSIS OF ROBUST DETECTORS FOR HYPERSPECTRAL IMAGING

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

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

ABSTRACT

When accounting for heterogeneity and non-Gaussianity of real hyperspectral data, elliptical distributions provide reliable models for background characterization. Through these assumptions, this paper highlights the fact that robust estimation procedures are an interesting alternative to classical methods and can bring some great improvement to the detection process. The goal of this paper is then not only to recall well-known methodologies of target detection but also to propose ways to extend them for taking into account the heterogeneity and non-Gaussianity of the hyperspectral images.

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

1. INTRODUCTION

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.

Target detection tasks arise in many different military and civilian applications and have been widely investigated in several signal processing domains such as radar, sonar, communications, etc. When the spectral signature of the desired target is known, it can be used as steering vector in *Target Detection* techniques [1]. Most of these techniques can also be found in the radar detection domain. In this context, Statistical Detection Theory [2] has led to several well-known algorithms, for instance the Matched Filter, the Kelly Detector and its adaptive versions [3]. Other interesting approaches are based on subspace projection methods [4]. But all of these different detection algorithms are based on the statistical background characterization, i.e. on the estimation of the mean vector and the covariance matrix of the background.

It has been often assumed for detector design that signals, interferences, noise and background are modeled as Gaussian stochastic processes. Indeed, this assumption makes sense in many applications. However, 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], 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 acknowledged models for background statistics characterization is the family of Elliptical Contoured Distributions (ECD). They account for non-Gaussianity providing a long tailed alternative to multivariate normal model. They are proven to represent a more accurate characterization of HSI than models based on Gaussian assumption [5].

Although non-Gaussian distributions are assumed for background modeling, the parameters estimation is still performed using classical Gaussian based estimators; as for the covariance matrix, generally determined according to the SCM approach. These classical estimators correspond to the Maximum Likelihood Estimators for Gaussian assumption. However, they lead to sub-optimal detection schemes when the noise is a non-Gaussian process. When working on ECD framework the model can be used to assess the robustness of statistical procedures and to derive alternative robust estimators of the parameters, the mean vector and the covariance matrix [6]. These can then be used as plug-in estimators in place of the unknown mean vector or/and of the covariance matrix. This is a simple but often efficient method to obtain robust properties for signal processors derived under the Gaussian assumption. We note that the use of robust signal processors is of fundamental interest due to the occurrence of impulsive environments and outliers in practical hyperspectral images [7, 8].

2. ELLIPTICALLY CONTOURED DISTRIBUTIONS

A *m*-dimensional random complex vector $\mathbf{y} = [y_1 y_2 ... y_m]^T$ with mean $\boldsymbol{\mu}$ and scatter matrix $\boldsymbol{\Sigma}$ has an elliptical distribution if its probability density function (PDF) has the form

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

where H denotes the conjugate transpose operator and $h_m(.)$ is any function such as (1) defines a PDF in $\mathbb{C}^m$. The function h_m is usually called density generator and it is assumed to be only approximately known. Note that it produces density contours corresponding to elliptical surfaces. If the second-order moment exists, then $\boldsymbol{\Sigma}$ reflects the structure of the covariance matrix of the elliptically distributed random vector $\mathbf{y}$, i.e. the covariance matrix is equal to the scatter matrix up to a scalar constant. It serves to characterize the correlation structure existing within the spectral bands. It is worth pointing out that the ECD class includes a large number of distributions, notably the Gaussian distribution, multivariate t distribution, K-distribution or multivariate Cauchy. Thus it allows for heterogeneity of the background power with the texture.

3. ROBUST PARAMETERS ESTIMATION

Along with their well-known properties and their simplicity of analysis, the sample covariance matrix (SCM) $\hat{\mathbf{M}}_{SCM} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i\mathbf{y}_i^H$ and the sample mean vector (SMV) $\hat{\boldsymbol{\mu}}_{SMV} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i$ are the most extended estimators since they are the Maximum Likelihood Estimators for Gaussian case. However, such widespread techniques are suboptimal when the noise is a non-Gaussian stochastic process. This article reviews some robust procedures particularly suited for estimating the covariance matrix and the mean vector of elliptical populations. The M-estimators were first introduced as a generalization of Maximum Likelihood Estimators (MLE) and have been used in several signal processing applications, such as radar detection. The complex M-estimators of location and scatter are defined as the joint solution of:

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{i=1}^{N} u_1(t_i)\,\mathbf{y}_i}{\sum_{i=1}^{N} u_1(t_i)}, \qquad (2)$$

$$\hat{\mathbf{M}} = \frac{1}{N}\sum_{i=1}^{N} u_2(t_i^2)\,(\mathbf{y}_i - \hat{\boldsymbol{\mu}})\,(\mathbf{y}_i - \hat{\boldsymbol{\mu}})^H \qquad (3)$$

where $t_i = \left((\mathbf{y}_i - \hat{\boldsymbol{\mu}})^H\hat{\mathbf{M}}^{-1}(\mathbf{y}_i - \hat{\boldsymbol{\mu}})\right)^{1/2}$, and where u_1, u_2 denote any real-valued *weighting functions* on the quadratic form t_i. Remark that the main purpose of u_1 and u_2 is to attenuate high contributions of this distance (outliers rejection). The choice of u_1 and u_2 do not need to be related to a particular elliptical distribution and therefore, M-estimators constitute a wide class of estimators that include the MLEs. Existence and uniqueness have been proven 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. We detail here, the particular cases of the *Fixed Point M*-estimator and *Huber*'s type M-estimator.

3.1. Fixed Point Estimators

The Fixed Point approach, according to the definition proposed by Tyler in [11], satisfy the following equations:

$$\hat{\boldsymbol{\mu}}_{FP} = \frac{\displaystyle\sum_{i=1}^{N} \frac{\mathbf{y}_i}{\left((\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})^H\hat{\mathbf{M}}_{FP}(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})\right)^{1/2}}}{\displaystyle\sum_{i=1}^{N} \frac{1}{\left((\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})^H\hat{\mathbf{M}}_{FP}(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})\right)^{1/2}}} \qquad (4)$$

and

$$\hat{\mathbf{M}}_{FP} = \frac{m}{N}\sum_{i=1}^{N} \frac{(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})\,(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})^H}{\left((\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})^H\hat{\mathbf{M}}_{FP}(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_{FP})\right)} \qquad (5)$$

which are the particular cases of (2) and (3) for $u_1(t) = t^{-1}$ and $u_2(t^2) = mt^{-2}$. For the scale-only problem, the Fixed Point covariance matrix estimate has been widely investigated in statistics and signal processing literature [12]. We refer to [13] for a detailed performance analysis.

The main results on the statistical properties of $\hat{\mathbf{M}}_{FP}$ are recalled for elliptical distribution framework (and $\boldsymbol{\mu}$ assumed to be known): $\hat{\mathbf{M}}_{FP}$ is a consistent and unbiased estimate of 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. Remark that the distribution of $\hat{\mathbf{M}}_{FP}$ does not depend on the specific elliptical distribution.

3.2. Huber type *M*-estimators

The *Huber's M*-estimators [15] are defined in the complex case [16] when taking in (2), (3) the following weighting functions:

$$u_1(t) = min\left(1, k/t\right) \quad u_2(t^2) = \frac{1}{\beta}min\left(1, k^2/t^2\right)$$

where $k > 0$ and β depend on an adjustable parameter $0 < q < 1$ according to

$$q = F_{2m}(2k^2)$$

$$\beta = F_{2m+2}(2k^2) + k^2\frac{1-q}{m}$$

where F_m is the cumulative distribution function of a χ^2 distribution with m degrees of freedom. The *Huber*-estimators might be interpreted as a mixture between the robust *Fixed Point* and the classical SCM. The values of the quadratic form below k^2 are kept and treated as in the SMV and SCM estimators; and the values higher than k^2 are attenuated by the weighting function similarly to the *Fixed Point* approach. In a complex Gaussian context, it can be shown that when N tens to infinity, the proportion of data processed with the SCM is equal to q.

4. TARGET DETECTION SCHEMES

We address the problem of the detection of a known signal vector (target response) $\mathbf{s}$ corrupted by an additive noise $\mathbf{b}$ in a m-dimensional complex vector $\mathbf{y}$. The signal absent vs. signal present problem can be stated as a binary hypothesis test H_0 and H_1. Under hypothesis H_1, it is assumed that the observed data consists on the sum of a signal and background noise $\mathbf{y} = \alpha\,\mathbf{p} + \mathbf{b}$ with $\mathbf{p}$ a perfectly known complex steering vector (characterizing for example the spectral material to detect) and α is the signal amplitude.

In practice, the background statistics are unknown and have to be estimated from N signal-free secondary data. The adaptive detector is obtained by replacing the unknown parameters by their estimates. For example, an estimator may be obtained from the range cells surrounding the cell under test. The size of the cell 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.

If some *a priori* knowledge of the noise statistics (e.g., K - distribution, t-distribution, etc.) is available, then $\mathbf{M}$ and $\boldsymbol{\mu}$ should be estimated by the MLE $\hat{\mathbf{M}}$ and $\hat{\boldsymbol{\mu}}$ of the covariance matrix and the mean vector of the assumed elliptical model. When there is no reliable statistical information on secondary data, they are assumed to be i.i.d. random samples from an unknown elliptical contoured distribution. Then practically any robust *M*-estimator could be used in the detector scheme. The *Adaptive Matched Filter* (AMF) is the optimal linear filter for maximizing the signal to noise ratio (SNR) in the presence of additive stochastic noise:

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

This detector hold the CFAR properties in the sense that its false alarm expression just depend on the dimension of the vector m and the number of secondary data used for the estimation N. However, its performance strongly relies on the good fit of the Gaussian model and the false alarm rate is highly increased when normal assumption is not attained.

The *Adaptive Normalized Matched Filter* (ANMF), also known as the linear quadratic Generalized Likelihood Ratio Test (GLRT) :

$$\Lambda_{ANMF}(\mathbf{y}) = \frac{|\mathbf{p}^H\,\hat{\mathbf{M}}^{-1}\,(\mathbf{y} - \hat{\boldsymbol{\mu}})|^2}{(\mathbf{p}^H\,\hat{\mathbf{M}}^{-1}\mathbf{p})\left((\mathbf{y} - \hat{\boldsymbol{\mu}})^H\,\hat{\mathbf{M}}^{-1}\,(\mathbf{y} - \hat{\boldsymbol{\mu}})\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

An important feature of the detector is its invariance under scalar multiples of $\mathbf{y}$. This implies, thanks to the properties of the elliptical distributions, that the detector behaves according to the same distribution regardless of the true elliptical distribution, i.e., it is distribution-free. This is of major interest in areas of background transition, like coastal areas (ground and sea) or the edge of forests (fields and trees) because the detector resulting distribution should be insensitive to the different clutter areas.

The *Kelly detector* was derived in [17]. It is based on the GLRT assuming in this case that the covariance matrix is unknown. Then, the adaptive Kelly detector scheme including the estimated mean takes the form:

$$\Lambda_{Kelly}(\mathbf{y}) = \tag{6}$$

$$\frac{|\mathbf{p}^H\,\hat{\mathbf{M}}^{-1}\,(\mathbf{y} - \hat{\boldsymbol{\mu}})|^2}{(\mathbf{p}^H\,\hat{\mathbf{M}}^{-1}\mathbf{p})\left(1 + \frac{1}{N}(\mathbf{y} - \hat{\boldsymbol{\mu}})^H\,\hat{\mathbf{M}}^{-1}\,(\mathbf{y} - \hat{\boldsymbol{\mu}})\right)} \underset{H_0}{\overset{H_1}{\gtrless}} N\lambda \tag{7}$$

The test in (7) is a CFAR detector, in which the probability of false alarm is independent of the true covariance matrix. However, it has no known optimality property. The AMF and the Kelly detector are based on the same assumptions about the nature of the observations. It is therefore interesting to compare their detection performance for a given P_{fa}. Note that for large values of N the performances are substantially the same.

5. SIMULATIONS

On a first stage, the theoretical analysis is validated on simulated data. We oppose the detection schemes proposed above for a heavy tail *K*-distribution, with shape parameter $\nu = 0.1$. The experiments were conducted on $m = 10$ dimensional vectors, for $N = 40$ secondary data and the computations have been made through 10^5 MonteCarlo trials.

We compare the three detection schemes in terms of probability of detection. We fixe the probability of false alarm to $P_{fa} = 10^{-3}$ and then we set the threshold to reach the desired P_{fa}, according to the false alarm regulation curves. In non-Gaussian case, see Fig. 1, the Gaussian based detectors do not present optimal results anymore. We observe that the AMF and the Kelly detector are the less performant in SNR requirements. The ANMF takes advantage of its invariance under a scale factor and indicates the best result, and so it is the detection test to retain. The use of robust estimators leads to a significant improvement when dealing with impulsive clutter; the decrease on the threshold allows to detect targets even for small SNR values.
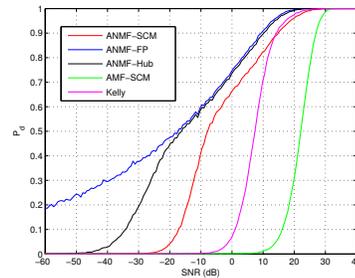

Fig. 1. Probability of detection for different SNR values and $P_{fa} = 10^{-3}$ in *K*-distribution case.

The same analysis has been performed on the hyperspectral image presented in Fig. 2 [18]. Hyperspectral data represent reflectance values, and hence, they are real and positive.

They have been passed through a Hilbert filter to render them complex. After sampling, the vector is of dimension $m = 10$ and the parameters estimation is performed locally using a sliding window of size seven, and $N = 48$. Once the P_{fa} has been fixed again at 10^{-3}, we place artificial targets over the image. We compare in Fig. 3 the outcome of the different detection tests. Note that the use of the *Fixed Point* estimators jointly with the ANMF lead to the best performance results, while AMF presents the highest SNR requirements.

Figure 1. Color rendering of self test hyperspectral image.

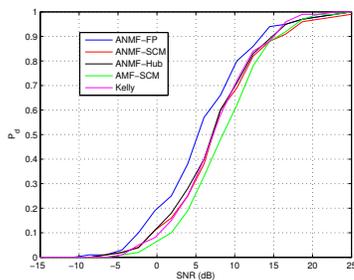**Fig. 2**. True color composition of the HyMap scene.

Fig. 3. Probability of detection for different SNR values and $P_{fa} = 10^{-3}$ in real HSI scene.

6. CONCLUSION

Using elliptical distributions for background modeling allows for heterogeneity consideration in non-Gaussian environment. We have proposed different estimators for statistical characterization of the noise. In particular, the class of M-estimators, which are specially appropriate to the addressed problem. We have recalled Fixed Point estimators and Huber type M-estimators emphasizing their statistical properties. Finally, we have presented three classical detection schemes and compared their performance in terms of probability of detection.

7. REFERENCES

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

[2] S. M. Kay, *Fundamentals of Statistical signal processing, Volume 2: Detection theory*. Prentice Hall PTR, 1998.

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

[4] 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.

[5] D. Manolakis and D. Marden, "Non gaussian models for hyperspectral algorithm design and assessment," in *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, vol. 3. IEEE, 2002, pp. 1664–1666.

[6] 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.

[7] J. Ovarlez, S. Pang, F. Pascal, V. Achard, and T. 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.

[8] J. Frontera-Pons, M. Mahot, J. Ovarlez, F. Pascal, S. Pang, and J. Chanussot, "A class of robust estimates for detection in hyperspectral images using elliptical distributions background," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*. IEEE, 2012, pp. 4166–4169.

[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. 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] ——, "Theoretical analysis of an improved covariance matrix estimator in non-gaussian noise," in *Proc. IEEE-ICASSP*, vol. IV, Philadelphia, March 2005, 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] E. Ollila and V. Koivunen, "Robust antenna array processing using m-estimators of pseudo-covariance," in *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, vol. 3. IEEE, 2003, pp. 2659–2663.

[17] E. J. Kelly, "An adaptive detection algorithm," *Aerospace and Electronic Systems, IEEE Transactions on*, no. 2, pp. 115–127, 1986.

[18] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a web-based application to evaluate target finding algorithms," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, vol. 2. IEEE, 2008, pp. II–915.