ROBUST ANOMALY DETECTION IN HYPERSONTICAL IMAGING

J. Frontera-Pons¹, M.A. Veganzones², S. Velasco-Forero³, F. Pascal¹, J.P. Ovarlez¹,⁴ and J. Chanussot²

1: SONDRA Lab, SUPELEC, Gif sur Yvette, France,
2: GIPSA-lab, Grenoble-INP, Saint Martin d’Hyères, France
3: National University of Singapore, Singapore
4: ONERA, the French Aerospace Lab, Palaiseau, France,

IEEE International Geoscience And Remote Sensing Symposium, July 2014, Quebec, Canada
OUTLINE OF THE TALK

- Problems description and motivation,
- The Elliptical Distribution modeling for Hyperspectral Imaging,
- Estimation of the Elliptical Distribution background parameters,
- Adaptive and Anomaly Detection in the Elliptical Distribution background,
- Detection and False Alarm regulation results on experimental data,
- Conclusions and Perspectives.
PROBLEMS DESCRIPTION

- **ANOMALY DETECTION IN HYPERSPECTRAL IMAGES**
  To detect all that is « different » from the background (Mahalanobis distance) - Regulation of False Alarm. Application to radiance images.

- **DETECTION OF TARGETS IN HYPERSPECTRAL IMAGES**
  To detect (GLRT) targets (characterized by a given spectral signature \( p \)) - Regulation of False Alarm. Application to reflectance images (after some atmospheric corrections or others).
CONVENTIONAL METHODS OF DETECTION

- Many methodologies for detection and classification in hyperspectral images can be found in radar detection community. We can retrieve all the detectors family commonly used in radar detection (AMF (intensity detector), ACE (angle detector), sub-spaces detectors, ...).

- Almost all the conventional techniques for anomaly detection and targets detection are based on Gaussian assumption and on spatial homogeneity in hyperspectral images.

All these adaptive techniques need to estimate the data covariance matrix (whitening process).
ADAPTIVE DETECTION IN HOMOGENEOUS GAUSSIAN BACKGROUND

Binary Hypotheses test

\[
\begin{cases}
H_0 : \quad \mathbf{c} = \mathbf{b}, \quad \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K \\
H_1 : \quad \mathbf{c} = \mathbf{b} + \mathbf{A} \mathbf{p}, \quad \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K
\end{cases}
\]

MLE:

\[
\hat{\mathbf{M}}_{SCM} = \frac{1}{K} \sum_{k=1}^{K} (\mathbf{c}_k - \hat{\mu}_{SCM}) (\mathbf{c}_k - \hat{\mu}_{SCM})^H
\]

Complex Multivariate Gaussian PDF:

\[
f_{\mathbf{c}}(\mathbf{c}) = \frac{1}{\pi^m |\mathbf{M}|} \exp \left( - (\mathbf{c} - \mu)^H \mathbf{M}^{-1} (\mathbf{c} - \mu) \right)
\]

TWO-STEP ADAPTIVE MATCHED FILTER

\[
\Lambda(\mathbf{c}) = \frac{\mathbf{p}^H \hat{\mathbf{M}}_{SCM}^{-1} (\mathbf{c} - \hat{\mu}_{SCM})^2}{\mathbf{p}^H \hat{\mathbf{M}}_{SCM}^{-1} \mathbf{p}} \quad \overset{H_1}{\underset{H_0}{\gtrless}} \lambda
\]

\[
P_{fa} = 2F_1 \left( K - m, K - m + 1; K; -\frac{\lambda}{K + 1} \right)
\]

TWO-STEP GLRT KELLEY DETECTOR

\[
\hat{\mathbf{M}}_{SCM} = \frac{1}{K} \sum_{k=1}^{K} (\mathbf{c}_k - \hat{\mu}_{SCM}) (\mathbf{c}_k - \hat{\mu}_{SCM})^H
\]

\[
\Lambda(\mathbf{c}) = \frac{\left| \mathbf{p}^H \hat{\mathbf{M}}_{SCM}^{-1} (\mathbf{c} - \hat{\mu}_{SCM}) \right|^2}{\left( \mathbf{p}^H \hat{\mathbf{M}}_{SCM}^{-1} \mathbf{p} \right) \left( N + (\mathbf{c} - \hat{\mu}_{SCM})^H \hat{\mathbf{M}}_{SCM}^{-1} (\mathbf{c} - \hat{\mu}_{SCM}) \right)} \overset{H_1}{\underset{H_0}{\gtrless}} \lambda
\]

\[
\hat{\mathbf{M}}_{SCM} = \frac{1}{K} \sum_{k=1}^{K} (\mathbf{c}_k - \hat{\mu}_{SCM}) (\mathbf{c}_k - \hat{\mu}_{SCM})^H
\]

[J. Frontera-Pons et al., IEEE Trans. on SP, in review]

GLRT ACE DETECTOR

\[
\Lambda(\mathbf{c}) = \frac{\left| \mathbf{p}^H \hat{\mathbf{M}}_{SCM}^{-1} (\mathbf{c} - \hat{\mu}_{SCM}) \right|^2}{\left( \mathbf{p}^H \hat{\mathbf{M}}_{SCM}^{-1} \mathbf{p} \right) \left( \mathbf{c} - \hat{\mu}_{SCM} \right)^H \hat{\mathbf{M}}_{SCM}^{-1} (\mathbf{c} - \hat{\mu}_{SCM})} \overset{H_1}{\underset{H_0}{\gtrless}} \lambda
\]

\[
P_{fa} = (1 - \lambda)^{K - m + 1} 2F_1 \left( K - m + 1, K - m; K; \lambda \right)
\]

Derived and valid only under Gaussian hypotheses,
Their false alarm rate are independent of the covariance matrix: CFAR-matrix property in homogeneous Gaussian data.
ANOMALY DETECTION IN HOMOGENEOUS GAUSSIAN BACKGROUND

GLRT RX ANOMALY DETECTOR: Mahalanobis Distance

\[ f_c(c) = \frac{1}{\pi^m |M|} \exp \left(- (c - \mu)^H M^{-1} (c - \mu) \right) \]

Binary Hypotheses test

\[
\begin{align*}
H_0 : & \quad c = b, \quad c_1, c_2, \ldots, c_K \\
H_1 : & \quad c = b + A p, \quad c_1, c_2, \ldots, c_K
\end{align*}
\]

\[ \Lambda(c) = (c - \hat{\mu}_{SCM})^H \hat{M}_{SCM}^{-1} (c - \hat{\mu}_{SCM})^{H_1} \geq H_0 \lambda \]

(Hotelling's T-squared distributed)

\[ \frac{K - m}{m (K+1)} \Lambda(c) \sim F_{m,K-m} \]

(When K tends to infinity, this test becomes chi-squared distributed)

 Derived and valid only under Gaussian hypotheses,
 Its false alarm rate is independent of the covariance matrix: CFAR-matrix property in homogeneous Gaussian data.
Hyperspectral data are generally spatially heterogeneous in intensity and/or cannot be only characterized by Gaussian statistic:

Elliptical distribution models have started to be studied in the hyperspectral scientific community and can help in characterizing heterogeneous background but one generally uses .... Gaussian estimates!
is a \( m \)-dimensional random complex vector with \( \Sigma \) the scatter matrix and \( \mu \) the mean vector, \( h_c \) usually called density generator, is assumed to be known.

**Subclass of Spherically Invariant Random Vector:** Compound Gaussian Process [Yao 73]

\[
f_c(c) = \frac{1}{\pi^m |\Sigma|} \int_0^\infty \frac{1}{\tau^m} \exp \left( -\frac{(c - \mu)^H \Sigma^{-1} (c - \mu)}{\tau} \right) p(\tau) \, d\tau
\]

\( c = \sqrt{\tau} x + \mu \)

\( x \sim \mathcal{CN}(0, \Sigma) \) \( \tau \sim p(\tau) \)

\( \sqrt{\tau} \) is a random variable representing the heterogeneity of the spatial background.

**Powerful statistical model that allows:**
- to encompass the Gaussian model but also to extend it (K, Weibull, Fisher, Cauchy, Alpha-Stable, Generalized Gaussian, etc.),
- to take into account the heterogeneity of the background power with the texture,
- to take into account possible correlation existing within the \( m \)-channels of observation.
In the framework of ECD, complex \( M \)-estimators of location and scatter [P.J. Huber 64, R.A. Maronna 76] are defined as the joint solutions of two implicit equations:

\[
\hat{\mu} = \frac{1}{K} \sum_{n=1}^{K} u_1(t_n)c_n, \\
\hat{M} = \frac{1}{K} \sum_{n=1}^{K} u_2(t_n^2)(c_n - \hat{\mu})(c_n - \hat{\mu})^H
\]

where \( t_n = \left( (c_n - \hat{\mu})^H \hat{M}^{-1} (c_n - \hat{\mu}) \right)^{1/2} \).

- \( u_1, u_2 \) are two weighting functions acting on the quadratic form, i.e. Mahalanobis distance,
- The choice of \( u_1, u_2 \) results in different estimates for the covariance matrix and the mean vector,
- Existence and uniqueness of the solution have been proven provided that \( u_1, u_2 \) satisfy given conditions [Maronna 1976],
- Robust to outliers, to the presence of strong targets or high impulsive samples in the reference cells,
- Generalization of Maximum Likelihood Estimators if the density generator \( h_y(.) \) is known:

\[
u_2(t^2) = u_1(t) = -h'_y(t^2) / h_y(t^2)
\]
For an unknown but deterministic texture parameter, the Maximum Likelihood Estimate (MLE) of the Covariance \( \hat{M} \) (approached MLE in the SIRV context), called the Fixed Point \( \hat{M}_{FP} \) (FP), is the solution of the following implicit equations [Tyler 1987]:

**Fixed Point (FP):**

\[
\hat{u}_2(x) = \frac{m}{x}
\]

\[
\hat{M}_{FP} = \frac{m}{K} \sum_{k=1}^{K} \frac{(c_k - \hat{\mu}_{FP}) (c_k - \hat{\mu}_{FP})^H}{(c_k - \hat{\mu}_{FP})^H \hat{M}_{FP}^{-1} (c_k - \hat{\mu}_{FP})}
\]

\[
\hat{\mu}_{FP} = \left( \frac{1}{K} \sum_{k=1}^{K} (c_k - \hat{\mu}_{FP})^H \hat{M}_{FP}^{-1} (c_k - \hat{\mu}_{FP}) \right)^{-1/2} \frac{1}{K} \sum_{k=1}^{K} (c_k - \hat{\mu}_{FP})^H \hat{M}_{FP}^{-1} (c_k - \hat{\mu}_{FP})
\]

These two quantities can be jointly reached by a convergent iterative algorithm [Pascal et al. 2008]

- This estimator does not depend on the elliptical distribution density generator,
- Robust to outliers, strong targets or scatterers in the reference cells,
- The Fixed Point Matrix Estimator is consistent, unbiased, asymptotically Gaussian and is, for a fixed number \( K \) of secondary data, Wishart distributed with \( mK/(m+1) \) degrees of freedom.
The Huber’s M-Estimators [P.J. Huber 64] is defined with the following two functions $u_1(.)$ and $u_2(.)$

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

where $q = F_{2m} \left(2k^2\right)$ and $\beta = F_{2m+2} \left(2k^2\right) + k^2 \frac{1-q}{m}$ and where $F_m(.)$ is the Chi² CDF with $m$ degrees of freedom

**Huber’s M-Estimators is a mix between FP and SCM**

- Extreme values of $t_n^2$ outside the interval $[0, k^2]$ are attenuated (Fixed Point behavior),
- Normal values below $k^2$ are uniformly kept (SCM behavior),
- The parameter $k$ can be adjusted to choose the percentage of data treated as Gaussian,
- The Huber estimate is consistent, unbiased, asymptotically Gaussian and is, for a fixed number $K$, Whishart distributed with a slight increase $\sigma_1 K$ of degrees of freedom ($\sigma_1$ very close to 1).
This two-step GLRT test is homogeneous of degree 0: it is independent of any particular Elliptical distribution: CFAR texture and CFAR Matrix properties.

Under homogeneous Gaussian region, it reaches the same performance than those of the detector built with the SCM estimate.

\[
\Lambda(c) = \frac{|p^H \hat{M}^{-1} (c - \hat{\mu})|^2}{(p^H \hat{M}^{-1} p) \left((c - \hat{\mu})^H \hat{M}^{-1} (c - \hat{\mu})\right)} \frac{H_1}{H_0} \lambda.
\]

Where the parameter \( \sigma_1 \) is very close to 1 but depends on the M-estimator:

Ex: for the Fixed Point, \( \sigma_1 = (m+1)/m \)
The hyperspectral data are **real** and **positive** as they represent radiance or reflectance.

- A mean vector has to be included in the model and estimated jointly with the scatter matrix,
- In order to keep the analytic Pfa/threshold relationship derived for complex data, the real HS data has been transformed into complex ones by a linear Hilbert filter and then be decimated by a factor 2 (principle of analytic signals)

The required SNR is slightly higher for the ANMF test.

We have detailed the class of complex elliptical distributions as a general model for background characterization in Hyperspectral Imaging. Elliptical distributions account for heterogeneity and long tail distributions present in real hyperspectral data. Once established that hyperspectral data cannot fit a multivariate Normal distribution, the use of the Gaussian maximum likelihood estimates (SCM and SMV) do not provide the optimal parameter estimation. We propose the use of robust estimates for the mean vector and the covariance matrix. We have described the M-estimators, notably the FP and the Huber type approach. The joint estimation of both parameters is a new challenging problem that opens many unknowns and it will be further investigated. We introduce the use of these estimates on classical detection methods. For false alarm regulation purposes, we have derived the theoretical relationship to set the proper threshold for a fixed probability of false alarm. Finally, we have validated the theoretical analysis over simulations and given some results on a real hyperspectral image. We conclude that the robust estimation tools presented in this paper offer a versatile alternative to Gaussian estimates. We remark that proposed M-estimators in Gaussian environment are capable of reaching the same results as the SCM and SMV. On the other hand, they outperform the classical estimation methods in case of non-Gaussian impulsive noise. This adaptability and their robustness make them suitable estimates in most scenarios.
SECOND SET OF RESULTS FOR FALSE ALARM RATE REGULATION WITH ACE-FP

Original data set

Extracted region:
- 100 x 100 pixels,
- 5 bands,
- Sliding Window: 19x19

Non-Gaussian region

Selected bands
ANOMALY DETECTION IN HETEROGENEOUS AND NON-GAUSSIAN BACKGROUND

EXTENDED GLRT RX ANOMALY DETECTOR: Mahalanobis Distance

Binary Hypotheses test

\[
\begin{align*}
H_0 : & \quad \mathbf{c} = \mathbf{b}, \\
H_1 : & \quad \mathbf{c} = \mathbf{b} + A \mathbf{p},
\end{align*}
\]

\[\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K\]

\[f_c(\mathbf{c}) = |\Sigma|^{-1} h_c ((\mathbf{c} - \mu)^H \Sigma^{-1} (\mathbf{c} - \mu))\]

\[\mathbf{p} \text{ unknown}\]

\[\Lambda(\mathbf{c}) = (\mathbf{c} - \hat{\mu}_{FP})^H \hat{\Sigma}_{FP}^{-1} (\mathbf{c} - \hat{\mu}_{FP}) \overset{H_1}{\geq} \lambda \overset{H_0}{\leq} \lambda\]

Derived and valid for any Elliptical Contoured Distributions,
Its false alarm rate unfortunately depends on texture statistic of the data
FIRST SET OF RESULTS FOR ANOMALY DETECTION

Results obtained with artificial targets

Original image (Forest Region)

Target Spectrum

50 x 50 pixels, 126 spectral bands
FIRST SET OF RESULTS FOR ANOMALY DETECTION

\[ RXD_{SCM} = (\mathbf{c} - \hat{\mu}_{SCM})^H \hat{\mathbf{M}}_{SCM}^{-1} (\mathbf{c} - \hat{\mu}_{SCM}) \]

\[ RXD_{FP} = (\mathbf{c} - \hat{\mu}_{FP})^H \hat{\mathbf{M}}_{FP}^{-1} (\mathbf{c} - \hat{\mu}_{FP}) \]

Better detection for close targets with FP estimates (same Pfa).
SECOND SET OF RESULTS FOR ANOMALY DETECTION

Problem of detecting galaxies in HS MUSE (Multi Unit Spectroscopic Explorer) data (465-930 nm)

Classical RX Detector

MUSE images:
- 300 x 300 pixels,
- 3578 spectral bands

Enhanced RX Detector

\[
RXD_{SCM} = (c - \hat{\mu}_{SCM})^H \hat{M}_{SCM}^{-1} (c - \hat{\mu}_{SCM})
\]

\[
RXD_{FP} = (c - \hat{\mu}_{FP})^H \hat{M}_{FP}^{-1} (c - \hat{\mu}_{FP})
\]
CONCLUSIONS AND PERSPECTIVES

CONCLUSIONS

• Hyperspectral images like radar or SAR images can suffer from non-Gaussianity or heterogeneity that can reduce the performance of anomaly detectors (RXD) and target detectors (AMF, ANMF),

• Elliptically Contoured Distributions modeling is a very powerful theoretical tool in the hyperspectral context that can match and circumvent the heterogeneity and non-Gaussianity of the images,

• Jointly used with powerful and robust estimates, the proposed hyperspectral detectors may provide better performances and a more accurate false alarm regulation. Moreover, they keep the same performance than the conventional Gaussian detectors for homogeneous and Gaussian data.

SOME PERSPECTIVES BEING ADDRESSED

• Regularized Robust Covariance Matrix Estimation,
• Strong connexion with Random Matrix Theory,
• Change Detection problems,