



Improvement of RX Algorithm Performance in Anomaly Detection Applied on Hyperspectral Imaging

A. El-Rewainy*, E. Farouk†

Abstract: Anomaly detection algorithms applied to hyperspectral imagery are able to identify objects from a natural environment without any prior knowledge about them based on the statistical analysis. RX detection algorithm is one of the most important detection algorithms in anomaly detection. It can detect targets with low probability of occurrence.

In this paper, we introduce pre-processing techniques for the hyperspectral image cube, Filtering and Whitening, to enhance the performance of the RX algorithm. These pre-processing techniques are done to ensure the requirements of the RX algorithm. Statistical analysis of the image cube is made before and after these pre-processing. The performance is investigated in terms of the detection probability, and the false alarm ratio (ROC curves).

Keywords: Hyperspectral imaging, Anomaly detection, Inner window region (IWR), Outer window region (OWR), Reed-Xiaoli (RX), Whitening.

1. Introduction

Hyperspectral sensors collect the spectral signature of a number of contiguous spatial locations (pixel) to form a hyperspectral data cube [1]. The basic task underlining many hyperspectral sensor imagery applications is to identify different materials based on their reflectance spectrum. In this respect, the concept of a spectral signature, which uniquely characterizes any given material, is highly attractive and widely used [2]. However, spectra observed from samples of the same material are never identical due to variations in the material surface, even in laboratory experiments. The amount of variability is more important in remote sensing applications due to the variations in atmospheric conditions, sensor noise, material composition, location, surrounding materials, and other factors. To make matters worse, totally different material types can have very similar spectra. Additional sources of spectral variability are calibration and illumination variations which are not currently handled by atmospheric correction codes. Under these conditions, it is sometimes preferable not to introduce target information when the application allows it [3, 4].

If we have no prior information about the target or if we wish to work with radiance, the most reasonable approach is to look for pixels whose spectral content is “significantly” different from those of the local background. This process is known in hyperspectral literature as anomaly detection [5, 6, 7]. Hyperspectral and multispectral imagery shows a great potential for this task because it provides both spectral and spatial features about the targets and backgrounds in the imagery.

* Egyptian Armed Forces, Egypt, eng_rewainy2007@yahoo.com

† Egyptian Armed Forces, Egypt, ezfarouk@yahoo.com

Most of the anomaly detection algorithms in the literature[2] assume that the HSI data can be represented by the multivariate normal (Gaussian) distribution and under the Gaussian assumption, the generalized likelihood ratio test (GLRT) is used to test the hypotheses to find the existence of a target in the image. The Gaussian assumption has been used mainly because of mathematical tractability that allows the formation of widely used detection models, such as GLRT. However, in reality the HSI data might not closely follow the Gaussian distribution.

One of the most famous anomaly detector algorithms is the Reed and Xiaoli Yu (RX) algorithm [5, 6], which is extensively used in multispectral and hyperspectral imagery to detect targets that are relatively small and only occur in the image scene with low probabilities, and depending on the Gaussian assumption of the image cube.

In this paper, we introduce two pre-processing techniques for the hyperspectral image cube, to enhance the performance of the RX algorithm (Gaussian assumption). The first one is filtering the image cube using a Gaussian low pass filter, and the second one is whitening the image cube to be uncorrelated data.

The paper is organized as follows, in section 2, we present an overview of Anomaly Detector algorithms. In section 3, we introduce Reed and Xiaoli Yu (RX) algorithm. In section 4, we introduce the two pre-processing techniques. In section 5, we show the statistical analysis and the ROC curves. Finally, we conclude on the performance of the algorithm.

2. Anomaly Detection

In several applications, we do not have any a priori information about the desired target. In such cases, it is possible to design algorithms searching for spectra which deviate from the local background (anomaly detection). This problem is typically formulated as a binary hypothesis test with two competing hypotheses: background only (\mathbf{H}_0) or target and background (\mathbf{H}_1). Since the two hypotheses depend on unknown parameters (for example background covariance matrix) which have to be estimated from the data, the detector has to be adaptive and is usually designed using the generalized likelihood ratio test (GLRT) approach. The type of statistical model used for the background leads to different anomaly detection algorithms.

3. The RX-Algorithm

Reed and Yu in [5] developed a GLRT, so called the RX anomaly detection, for multidimensional image data assuming that the spectrum of the received signal (spectral pixel) and the covariance of the background clutter are unknown. Let each input spectral signal be denoted by a vector $\mathbf{x}(n) = (x_1(n), x_2(n), \dots, x_J(n))^T$ consisting of J spectral bands. Define \mathbf{X}_b to be a $J \times M$ matrix of the M reference background clutter pixels. Each observation spectral pixel is represented as a column in the sample matrix \mathbf{X}_b

$$\mathbf{X}_b = [\mathbf{x}(1) \ \mathbf{x}(2) \ \dots \ \mathbf{x}(M)]$$

The two competing hypotheses that the RX-algorithm must distinguish are given by:

$$\mathbf{H}_0 : \mathbf{x} = \mathbf{n}, \quad \text{Target absent}$$

$$\mathbf{H}_1 : \mathbf{x} = \mathbf{a}s + \mathbf{n}, \quad \text{Target present}$$

where $a = 0$ under \mathbf{H}_0 and $a = 1$ under \mathbf{H}_1 , \mathbf{n} is a vector that represents the background clutter noise process, and \mathbf{s} is the spectral signature of the signal (target) given by $\mathbf{s} = [s_1, s_2, \dots, s_j]$. The target signature \mathbf{s} and background covariance C_b are assumed to be unknown. The model assumes that the data arises from two normal PDFs with the same covariance matrix but different means. Under \mathbf{H}_0 the data (background clutter) is modeled as $N(\mathbf{0}, C_b)$ and under \mathbf{H}_1 it is modeled as $N(\mathbf{s}, C_b)$. The background covariance C_b is estimated from the reference background clutter data. The estimated background covariance \hat{C}_b is given by:

$$\hat{C}_b = \frac{1}{M} \sum_{i=1}^M (\mathbf{x}(i) - \hat{\boldsymbol{\mu}}_b)(\mathbf{x}(i) - \hat{\boldsymbol{\mu}}_b)^T$$

where $\hat{\boldsymbol{\mu}}_b$ is the estimated background clutter sample mean given by:

$$\hat{\boldsymbol{\mu}}_b = \frac{1}{M} \sum_{i=1}^M \mathbf{x}(i)$$

Assuming a single pixel target \mathbf{r} as the observation test vector, the expression for the RX-algorithm is given by:

$$D_{RX}(\mathbf{r}) = (\mathbf{r} - \hat{\boldsymbol{\mu}}_b)^T \hat{C}_b^{-1} (\mathbf{r} - \hat{\boldsymbol{\mu}}_b)$$

which is the Mahalanobis distance of the test pixel spectrum from the mean of the background clutter pixels. The quantity D is compared to a certain threshold η to identify either this test pixel is under \mathbf{H}_1 or \mathbf{H}_0 (target or not).

In order to implement the RX for the dual windows, the RX in equation is modified as:

$$D_{RX}(\mathbf{r}) = (\boldsymbol{\mu}_{in} - \boldsymbol{\mu}_{out})^T \hat{C}_b^{-1} (\boldsymbol{\mu}_{in} - \boldsymbol{\mu}_{out})$$

So, it is clear that the RX algorithm is a locally adaptive constant false alarm rate (CFAR) detector, which assumes spatially uncorrelated Gaussian image with unknown spectral distributions and covariance matrices.

In the conventional RX it was suggested that the assumption of spatially uncorrelated image is valid. When the image contains significant correlation, this will has a strong effect on algorithm performance. So, the task is how to treat the problem of spatial correlation.

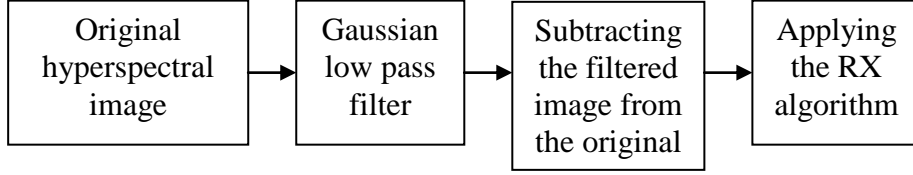
4. The Pre-Processing Techniques

4.1. Filtering the Image

As mentioned before, the RX algorithm is optimal when the hyperspectral image can be represented by a local normal model (target and background).

Applying a Gaussian low pass filter to the image, then subtracting the filtered image from the original image, results in a pure Gaussian spatially uncorrelated image, which is more suitable for the RX algorithm.

The RX algorithm can then be applied to the pre-processed image. We can summarize the processing sequence in the shown block diagram:



4.2. Whitening the Image

Another useful preprocessing strategy is to first whiten the image. This means that before the application of the RX algorithm (and after centering), we transform each pixel vector x linearly so that we obtain a new vector x_{whitened} which is white, i.e. its components are uncorrelated and their variances equal unity. In other words, the covariance matrix of global image equals the identity matrix.

The whitening transformation is always possible. One popular method for whitening is to use the eigen-value decomposition (EVD) of the covariance matrix of the global image ($C_{J \times J}$)[3].

$$C_{J \times J} = A \Lambda A^T$$

where A is the orthogonal matrix of eigenvectors of $C_{J \times J}$ and Λ is the diagonal matrix of its eigenvalues, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_j)$.

Note that $C_{J \times J}$ can be estimated in a standard way from the available whole image. Whitening can now be done by:

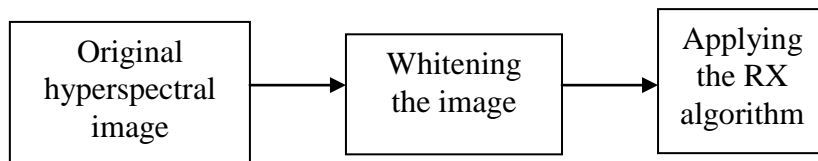
$$x_{\text{whitened}} = \Lambda^{-1/2} A^T x$$

where the matrix $\Lambda^{-1/2}$ is computed by a simple component-wise operation as $\Lambda^{-1/2} = \text{diag}(\lambda_1^{-1/2}, \lambda_2^{-1/2}, \dots, \lambda_j^{-1/2})$. It is easy to check that now $C_{J \times J} = I$.

Using the whitening transformation, the RX anomaly detector can be expressed as:

$$D_{\text{whitened RX}} = (\hat{\mu}_{\text{in}_{\text{whitened}}} - \hat{\mu}_{\text{out}_{\text{whitened}}})^T (\hat{\mu}_{\text{in}_{\text{whitened}}} - \hat{\mu}_{\text{out}_{\text{whitened}}})$$

which is the Euclidian distance of the region under test mean from the background mean in the whitened space. We can summarize the processing sequence in the shown block diagram:



5. Simulation Results

The conventional RX and the pre-processing techniques were applied to several hyperspectral images (about 5 images) and the detection performance of each method was evaluated. The tested hyperspectral images were taken from the Sherlock spectral camera and they are in the mid IR band (3 to 5 μm). Each hyperspectral image has 50 band images across the whole spectral range with step size of approximate 40 nm. These images were taken in different views. This number of spectral bands (50) is sufficient for spectral signature to distinguish between different materials in mid IR band.

An example of our images is a scene in grass has 9 man-made targets as shown in Fig. (1-a). We have used the dual window technique whose inner window size (target size) was set to enclose most of the targets in our data set whose approximate size was predetermined via prior knowledge of the range and field of view (FOV). Similarly, the size of the OWR (background region) is set to include sufficient number of spectral vectors from the neighboring background. The size of the inner window was 35×25 pixel and the outer window was 70×50 pixel.

Figures (1-b, c, d) show the results of the conventional RX algorithm, the filtered image RX algorithm and the whitened image RX algorithm respectively at the same threshold.

To make statistical analysis of the algorithms, there is a way to measure the whitening of the image. Chen & Reed have introduced in [8] a heuristic criterion to evaluate the diagonality d_{cov} of the covariance matrix of the whole image. This criterion was given by:

$$d_{\text{cov}} = \frac{\sum_{m=1}^J \sum_{n=1}^J \text{cov}_{mn}^2 - \sum_{m=1}^J \text{cov}_{mm}^2}{\sum_{m=1}^J \text{cov}_{mm}^2}$$

where cov_{mn} are the elements of the covariance matrix. It is evident that the smaller d_{cov} is, the more likely the covariance matrix to be diagonal.

It was found that d_{cov} in the conventional RX equal 41.65 and in filtered image RX equal 28.8 and in whitened RX equal zero.

The ROC curve of the three algorithms is shown in Fig. (2). It is known that the ROC is the relation between the probability of detection (P_d) and the probability of false alarm (P_{fa}). It is known that the higher the curve is (the more concave), the better the detector.

6. Conclusion

In this paper we have presented two pre-processing techniques used before applying the conventional RX algorithm which is one of anomaly detectors in multispectral and hyperspectral imagery. The two techniques are filtering the image cube with a Gaussian low pass filter, and the other is whitening the image cube. We have justified the interest of these methods to ensure the requirements or assumptions of the conventional RX anomaly detector (Gaussian assumption and spatial uncorrelation of the image cube). Then we have shown the superiority of applying the presented pre-processing over the conventional RX algorithm alone. Statistical analysis was made to the 3 algorithms and the ROC curve was established to ensure the improvement of the RX algorithm performance.

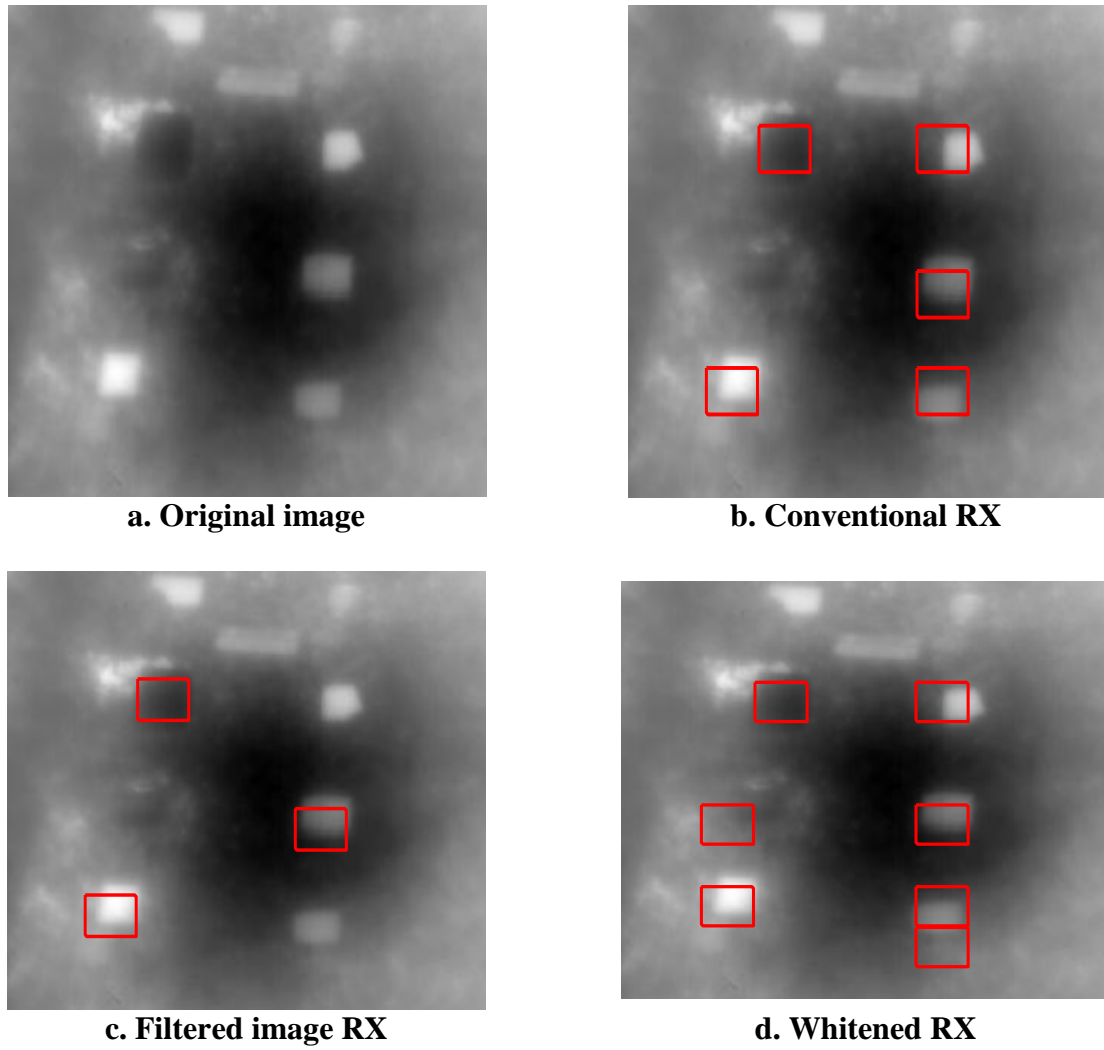


Fig. 1 The hyperspectral image used in this paper and the results of the 3 algorithms

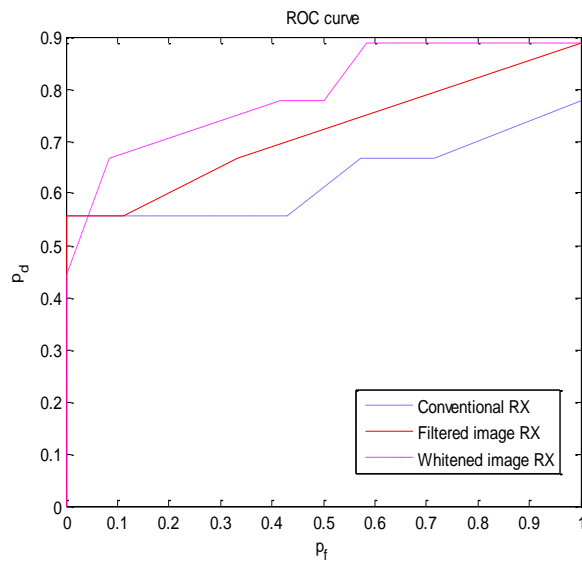


Fig. 2 The ROC curve of the 3 algorithms

7. References

- [1] C.I. Chang, "Hyperspectral Imaging," in *Theoretical Foundations of Computer Vision*, Kluwer Academic, Ed. Plenim, 2003.
- [2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Proc. Mag.*, pp. 29–43, Jan. 2002.
- [3] J.-M. Gaucel, M. Guillaume, and S. Bourennane, "Whitening Spatial Correlation Filtering For Hyperspectral Anomaly Detection," *IEEE Signal Proc. Mag.*, pp. 333–336, 2005.
- [4] G.D. Nash and G.W. Johnson, "Anomaly detection in dixie valley, nevada using hyperspectral data," in *Proc. 37th Workshop on geothermal Reservoir Engineering*, Stanford, California,, Jan. 2002, pp. 28–30.
- [5] I. S. Reed and X. Yu, "Adaptative multiple-band cfar detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
- [6] X. Yu, L. E. Hoff, I. S. Reed, and L. B. Stotts, "Automatic target detection and recognition in multiband imagery: A unified ml detection and estimation approach," *IEEE Trans. Image Processing*, vol. 6, pp.143–156, Jan. 1997.
- [7] S.M. Schweizer and J.M.F. Moura, "Hyperspectral imagery: Clutter adaptation in anomaly detection," *IEEE Trans. Inform. Theory*, vol. 46, no. 5, pp. 1855–1871, Aug. 2000.
- [8] J. Y. Chen, & I. S. Reed, A Detection Algorithm for Optical Targets in Clutter, *IEEE Trans. AES*, vol. 23, NO. 1, Jan. 1987, 394-405.