

Improving the RX Anomaly Detection Algorithm for Hyperspectral Images using FFT

Mohsen Zare-Baghbidi, Saeid Homayouni, and Kamal Jamshidi

Abstract

Anomaly Detection (AD) has recently become an important application of target detection in hyperspectral images. The Reed-Xialoi (RX) is the most widely used AD algorithm that suffers from “small sample size” problem. The best solution for this problem is to use Dimensionality Reduction (DR) techniques as a pre-processing step for RX detector. Using this method not only improves the detection performance of algorithm, but also significantly reduces its runtime. This paper presents a novel DR technique that uses the Fast Fourier Transform (FFT) to perform the band reduction for RX detector. We compared the proposed method, named FFT-RX, with several well-known detectors such as RX, RX-UTD, Kernel-RX, PCA-RX and DWT-RX. These algorithms applied to two hyperspectral datasets acquired by both the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Hyperspectral Mapper (HyMap) sensors. The evaluation of algorithms was based on Receiver Operation Characteristic (ROC) curve, visual investigation, and runtime of algorithms as well. Experimental results show that the proposed method improves the detection performance and runtime of RX detector significantly and has the best runtime and detection performance among all methods .

Keywords—Hyperspectral Remote Sensing, Anomaly Detection, Dimensionality Reduction, Fast Fourier Transform.

I. INTRODUCTION

Hyperspectral image data are very essential and useful in order to detect, recognize, quantify, and map many targets for various applications such as search-and-rescue operations, mine detection, and military usages [1].

Hyperspectral sensors are a powerful tool for distinguishing between different materials on the basis of each object's unique spectral signatures; these sensors are able to do this because they collect information about surfaces and objects in hundreds of

narrow contiguous spectral bands in the visible and infrared regions of the electromagnetic spectrum [1].

Anomaly Detection (AD) is a special kind of target detection (TD) techniques with no a priori information about the targets. The main purpose of these algorithms is to find the objects in a given image that are anomalous with respect to their surrounding background [1]. In other words, the point of anomaly detectors is to find the pixels whose spectra significantly differ from the background spectra [2]. The main advantage of these methods is that they don't need a priori information about the target signature, nor do they need any form of atmospheric or radiometric corrections on data [3].

The Reed-Xialoi (RX) is the most widely used AD algorithm [4], it is known as a benchmark anomaly detector for multi/hyperspectral images. This algorithm, which is derived from the generalized likelihood ratio test (GLRT), assumes that the background pixels in a local neighborhood around the target can be modeled by the multivariate normal (Gaussian) distribution [5,6]. RX assumes that r is an image pixel vector, which has L elements. L is the number of image's spectral bands. The RX detector is defined by equation (1).

$$\delta_{rxd}(r) = (r - \mu)^T C_{L \times L}^{-1} (r - \mu) \quad (1)$$

In this equation, μ is the sample mean vector. C is the sample data covariance matrix. Finally $\delta_{rxd}(r)$ is the well-known *mahalanobis* distance that shows the abnormality amount of pixel under test (PUT). The result of AD process is a two dimensional detection matrix. To determine the exact location of targets (anomalies), a threshold should be applied on the detection matrix.

The most reported problem for the RX and many of its modified versions is the “small sample size”. This problem concerns the estimation of a local background covariance matrix from a small number of very high dimensional samples. This may result in a badly conditioned and unstable estimate of local background covariance matrix that strongly affects the detection performance of the AD algorithm [7]. The first solution to this problem is enlarging the sample size by expanding the local window size. This solution tries to resolve the non-homogeneity of the local

Manuscript received August 6, 2014; revised January 31, 2015; accepted February 17, 2015.

M. Zare-Baghbidi, Department of Computer Engineering, College of Engineering, University of Isfahan, Isfahan, Iran (e-mail: mohsen.zare@zoho.com).

S. Homayouni, Department of Geography, University of Ottawa, Ottawa, Canada (e-mail: saeid.homayouni@uottawa.ca).

K. Jamshidi, Department of Computer Engineering, College of Engineering, University of Isfahan, Isfahan, Iran (e-mail: jamshidi@eng.ui.ac.ir).

background, which undermines the effectiveness of the covariance matrix estimation. Another solution for this problem is using the Dimension Reduction (DR) [7, 8].

Performance of many AD algorithms can be improved by using a pre-processing DR step. Because the hypercube is a relatively large empty space and the most important or interesting information could be represented in a few features [9, 10]. The DR step, used as a pre-processing step of the AD algorithm, can reduce the inter-band spectral redundancy and ever-present noise. Although the DR is lossy, it increases the separation between anomaly and background signatures. Thus, the detection performance of the anomaly detector is improved. DR techniques are divided into two categories: linear and nonlinear. Although linear techniques do not exploit the nonlinear properties in hyperspectral data, they can be fast enough for real time applications. A popular linear DR method, which is ideally used for small target detection is Principle Component Analysis (PCA) [11]. There are other linear DR methods, such as the Discrete Wavelet Transform (DWT), which is used to improve the detection performance and runtime of AD algorithms [12]. Other linear DR methods, such as Fourier Transform (FT) have not been investigated for AD methods.

A general framework of an AD scenario is shown in Fig. 1. In the first step, the spectral dimension of an image cube is reduced using a DR method (Fig. 1(b)). The AD algorithm is then used to analyze new image; the result is a two dimensional matrix named ‘‘AD matrix’’ (Fig. 1(c)).

To specify the locations of anomalies or targets in the image, a post-processing threshold step can be added to the algorithm. The final result is an image that shows the exact location of the targets (Fig. 1(d)).

In this study we introduce a new DR method for RX detector based on the Discrete Fourier Transform (DFT). We compared the proposed method, namely FFT-RX, with RX [4], RX-UTD [13], Kernel-RX [5],

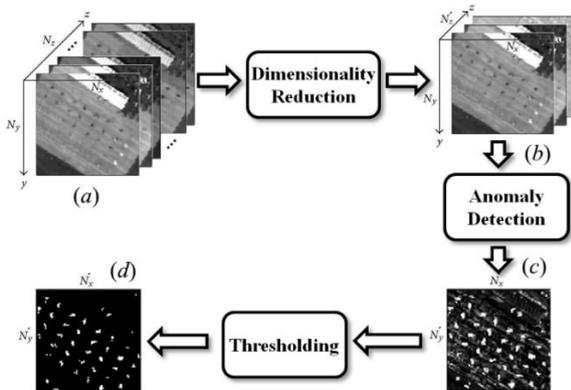


Fig. 1. Flowchart of hyperspectral AD using the pre-processing DR method

PCA-RX [8, 11] and DWT-RX [12]. The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Hyperspectral Mapper (HyMap) datasets are being used to both apply and evaluate the proposed methodology on real hyperspectral remotely sensed images.

II. PROPOSED DIMENSIONALITY REDUCTION METHOD

The Fourier Transform (FT) is an essential tool for the analysis of signals and systems. This transformation maps a time series onto the series of frequencies. The FT determines all signal frequencies; it does not have a time resolution [14]. However, because in dimensionality reduction methods, the time resolution is not important, using this transformation as a dimensionality reduction technique is not a negative attribute. In order to find the frequency spectrum of a digital signal, we used Discrete Fourier Transform (DFT), which is the discrete version of the FT. A practical method for computing the DFT, which requires much less computational effort, is the Fast Fourier Transform (FFT) [15, 16]. Cooley and Tukey introduced this method, which can lead to major changes in computational techniques [17].

In general, most signals mainly include low frequency components [18]. Therefore, when calculating the amplitude of the DFT of a signal, such as a spectrum pixel of a hyperspectral image, the early component of the amplitude, which is related to low frequencies in the main signal, are very high compared to other components. This problem is shown in Fig.2. In this figure, part (a) shows a pixel spectrum of a hyperspectral image that has 64 spectral bands; part (b) shows the DFT amplitude of part (a). According to this image, the amplitude of coefficients related to low frequencies is very high. This important characteristic can reduce the dimensions of hyperspectral images.

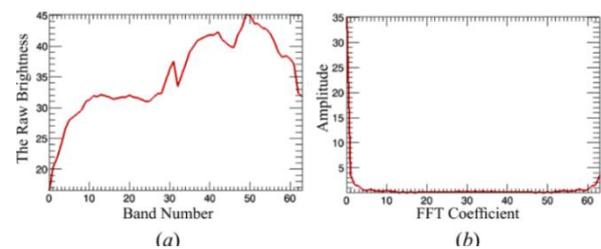


Fig. 2. (a) A spectrum pixel of a hyperspectral image, (b) FFT amplitude of the main signal

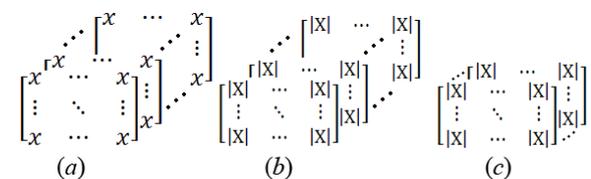


Fig. 3. (a) Hyperspectral image matrix, (b) FFT Amplitude of hyperspectral matrix, (c) Abstract of FFT Amplitude matrix.

We used DFT in a three-step anomaly detection framework (see Fig. 3). In the first step, the DFT of every image pixel is calculated using the FFT. The amplitude of the DFT values is, then, calculated. The results are placed in a matrix named “amplitude” (Fig. 3(b)). The last step uses a few bands of the amplitude matrix which are related to low frequencies of the main image (and which, in addition, have high values). In this way, a new matrix is formed which is the abstract of the FFT amplitude matrix (Fig. 3(c)). This matrix is a good approximation of the main image and any anomaly detectors can employ it. The number of abstract matrix’s bands which shows the amount of band reduction, can be selected during the experiment.

III. EXPERIMENTAL RESULTS

A. Hyperspectral images

1) HyMap data with implanted targets (Img-I)

This hyperspectral image, which released for Target Detection Blind Test project, has 126 spectral bands [19, 20]. During the image acquisition campaign, 12 real targets were located in an open grass region. Targets of this image are divided into two parts: self-test and blind-test. Because only the real location of the self-test targets is available, we cannot use this part of the image to evaluate the performance of the AD algorithms. Due to this limitation, some self-test targets (red cotton, blue cotton, yellow nylon, and red nylon) were selected and implanted in another part of the image. To implant the targets in this sub-image (named “Img-I”) a target implanted method [21, 22] has been used. For this method, asynthetic sub-pixel anomaly, z , is a combination of both the target and background, as shown in equation (2). In this equation, t and b shows (i.e., denotes) the target and background vectors, respectively. Therefore, sub-pixel (z) consists of the target’s spectrum with fraction f , and the background’s spectrum with fraction $(1-f)$ [22].

$$z = f \cdot t + (1 - f) \cdot b \quad (2)$$

This implantation method does not include the adjacency effects of the target spectrum on the local background pixels. To have a more realistic condition, the background pixels, which are neighbors of the targets, can be affected by a target pixel. This effect can be achieved by using a Gaussian function with a width of w , as shown in equation (3), where p_i is the spatial distance between background pixel (z_i) and the target pixel (t) [6].

$$z_i = \exp\left(-\frac{\rho_i^2}{w^2}\right) \cdot f \cdot t + \left(1 - \exp\left(-\frac{\rho_i^2}{w^2}\right) \cdot f\right) \cdot b_i \quad (3)$$

In order to construct the desired image, according to Fig. 4, a part of the main image is selected; the targets are then implanted in the selected sub-image (Fig. 4(a)). To apply the effect of background on targets and, in addition, make sub-pixels, outlines of targets have been selected and combined with their adjacent background according to equation (2) with the coefficient $f=0.6$. This means that every target pixel is composed of 60% target and 40% background.

Then, we used equation (3) in order apply the effect of anomalies on the background pixels. The final image with implanted targets includes sub-pixel and full-pixel (or multi-pixel) targets. As a result, this image seems to be a perfect data for testing AD and TD algorithms. Fig. 4(b) shows the truth location of the targets that are either sub-pixel or full-pixel.

2) AVIRIS data with real targets (Img-II and Img-III)

Two other sub-images have been extracted from a hyperspectral image of a naval air station in San Diego, California, collected by the AVIRIS sensor [23]. This data cube has 189 useful spectral bands with a ground resolution of 3.5 meters (see Fig. 5). The first sub-image, named Img-II, is an 80×80 pixel data cube that contains some military targets as anomalies and is used to evaluate the exact detection performance of algorithms using Receiver Operation Characteristic (ROC) curve (Fig. 5 (a)). The truth location of targets in this sub-image is shown in Fig. 5 (b). The second sub-image, named Img-III, is an image window with 100×100 pixels. This sub-image contains 38 anomalous targets which may be either helicopters or helipads, as shown in Fig. 5(c). This sub-image is used in some TD works [8, 24]; it is also used to evaluate the runtime of anomaly detectors.

A. Implementation

One of the most important decisions for AD algorithms is about the size of detection window [6]. Although, there is no specific method for choosing these windows [6], the size of the inner window should be almost as large as the biggest target in the scene. In addition, the size of the outer window should be large enough to provide a sufficient number of background samples for simulating the local background [25]. According to the both above-mentioned rules and the results of the experiment, the inner and outer window size for Img-I are selected 3×3 and 11×11 pixels, respectively. The inner and outer windows for both Img-II and Img-III are selected 5×5 and 13×13 pixels, respectively.

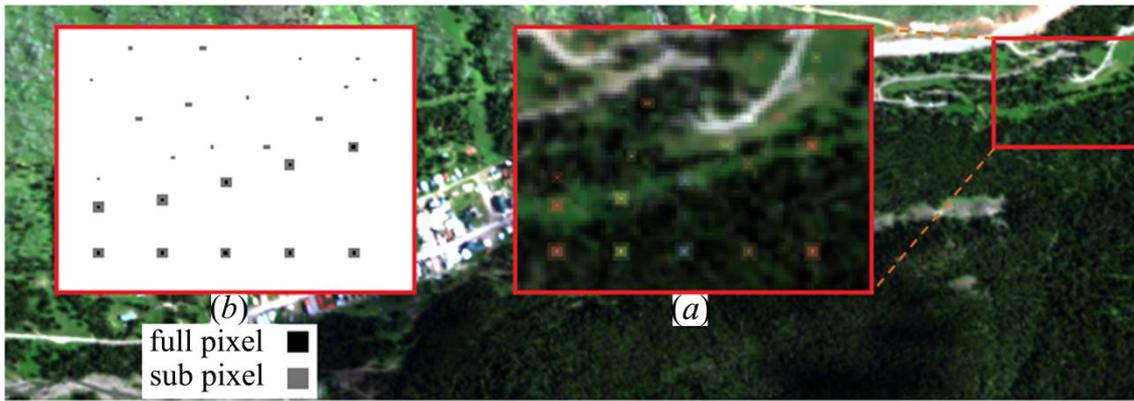


Fig. 4. A natural color composite of the HyMap data cube, (a) selected sub-image with implanted targets (Img-I), (b) truth location of targets.

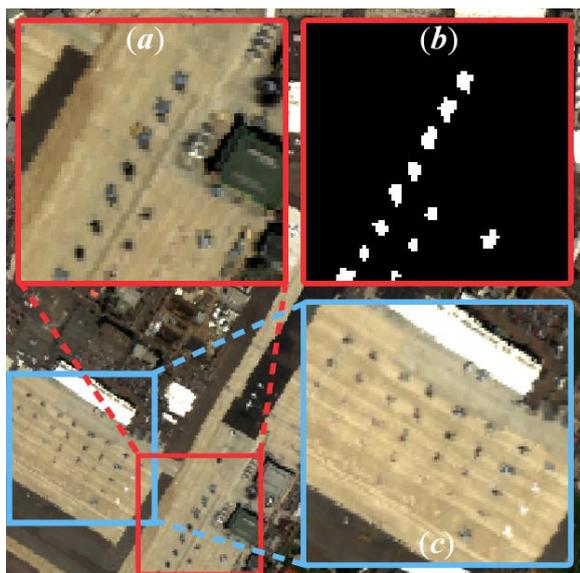


Fig. 5. A natural color composite of the AVIRIS data cube, (a) sub-image with real targets (Img-II), (b) truth locations of targets in Img-II and (c) sub-image with real targets (Img-III).

The proposed DR method (FFT-RX) acts as a pre-processing step for RX method and makes method. In addition, PCA-RX and DWT-RX methods are RX detectors that use PCA and DWT detectors as a pre-processing step. An important decision for the FFT, DWT and PCA DR methods is the amount of reduction rate that determines the number of image/feature bands after the DR step. This parameter should be selected according to two metrics: detection performance and runtime. In this study, according to the experiments conducted in different cases on Img-I and Img-II, the band number of output images is assumed to be 8 for all DR methods.

B. Evaluation of Detection Performance

The ROC curve is the best way to evaluate the detection performance of AD algorithms [1]. To evaluate the detection performance of algorithms more accurately, the area under the ROC curve (AUC) is

used. This value is an exact criterion; it is widely used to evaluate the detection performance of target detection algorithms [8]. Another way to evaluate the performance of algorithms is the visual investigation. This evaluation can be a good criterion using the post-processing threshold step. In this study the evaluation of algorithms for Img-I and Img-II datasets is done using the ROC curve and the AUC value; in addition, the Img-III data are used to evaluate algorithms visually.

1) Results of Img-I

Fig.6 shows the ROC curves of the AD methods for Img-I. Table 1 presents the AUC values of all algorithms. According to these criteria, the detection performance of the RX and RX-UTD methods is, in general, very weak. This is because of “small sample size problem” that concerns the estimation of local background covariance matrix [7]. However, the pre-processing DR methods not only reduce the inter-band spectral redundancy and ever-present noise but also increase the separability between anomaly and background signatures. Therefore they can significantly improve the performance RX detector.

According to the ROC curves and AUC values, the FFT-RX and DWT-RX ones perform best; the RX and RX-UTD methods perform worst and the performance of other methods are almost acceptable.

2) Results of Img-II

Fig.7 shows the ROC curves of the anomaly detectors for Img-II. The AUC values of these methods are shown in Table 2. According to these results, FFT-RX and Kernel-RX ones performed best, the performance of RX and RX-UTD is the worst and the performance of other methods is good. These results show the performance of RX detector and its family is, in general, very weak, but we can improve their performance using DR methods. In addition kernelizing RX detector can improve its performance significantly.

3) Results of Img-III

Img-III is used to evaluate the performance of anomaly detectors visually in a real scene. Because the truth location of the targets in this image is not available, the detection performance of AD algorithms is investigated visually. To achieve a better visual investigation, a threshold step is added at the end of the AD procedure. To execute this post-processing step, a cut-off threshold is needed; this value can be calculated adaptively using equation (4) [26]:

$$\tau_\alpha = \mu_d + Z_\alpha \times \sigma_d \quad (4)$$

TABLE I
AUC of AD methods applied to Img-I

AD Algorithm	AUC
RX	0.50
RX-UTD	0.49
Kernel-RX	0.94
PCA-RX	0.92
DWT-RX	0.95
FFT-RX	0.95

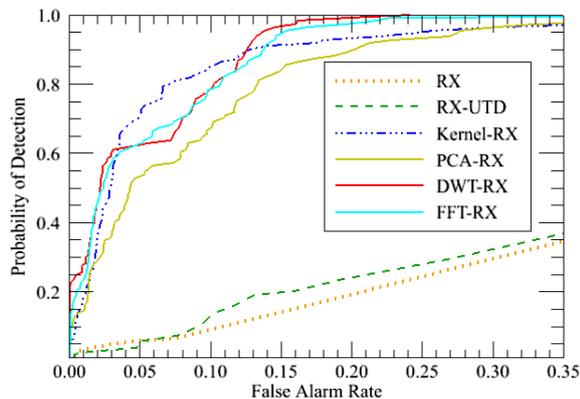


Fig. 6. ROC curves of Anomaly Detection algorithms for Img-I,

where τ_α is the cut-off threshold that declares whether a pixel is a target or not, μ_d and σ_d are the mean and standard deviation of the output of the AD algorithm, respectively, and Z_α is the z statistic at the significant level of α , which controls the number of pixels declared to be anomalies. Fig. 8 shows the output of the threshold step using the adaptive cut-off threshold of equation (4).

According to these results, the performance of RX and RX-UTD is very weak but DR step increases RX performance significantly. The Kernel-RX detector suffers from False Alarm Rate (FAR) that reduces its performance. Performance of FFT-RX and PCA-RX is almost the same but FFT-RX has the lowest FAR. In addition, performance of DWT-RX is good.

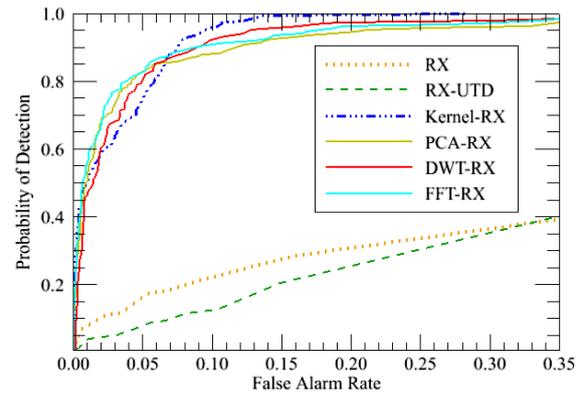


Fig. 7. ROC curves of Anomaly Detection algorithms for Img-II

TABLE II
AUC of AD methods applied to Img-II

AD Algorithm	AUC
RX	0.49
RX-UTD	0.50
Kernel-RX	0.97
PCA-RX	0.95
DWT-RX	0.96
FFT-RX	0.97

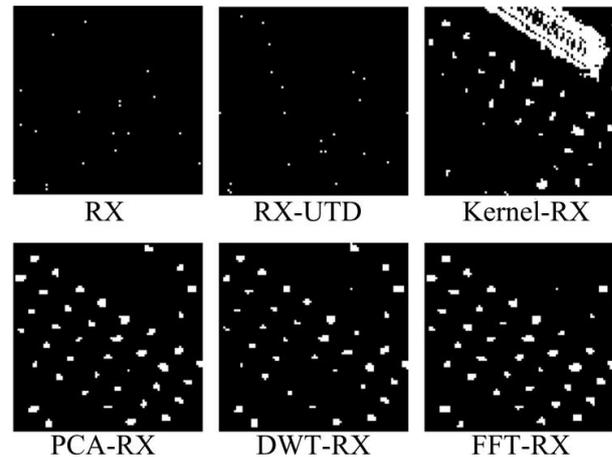


Fig. 8. Detection results of algorithms applied to Img-III

C. Runtime Evaluation

To evaluate the speed of the AD methods, a computer system with an “Intel Core i5-2410M, 2.3GHz” processor and four GB of Random Access Memory (RAM) is used to measure the runtime of algorithms on Img-III, in equal conditions. Runtime of the DR methods is shown in Table 3; the runtime of the AD methods, which includes the runtime of related DR pre-processing methods, is shown in Table 4.

TABLE III
Runtime of DR methods applied to Img-III

DR Method	PCA	FFT	DWT
Runtime(s)	0.98	0.26	9.78

TABLE IV
Runtime of AD methods applied to Img-III

AD Algorithm	AUC
RX	348.84
RX-UTD	348.70
Kernel-RX	1395.94
PCA-RX	7.83
DWT-RX	16.65
FFT-RX	7.16

According to these results, among the DR methods, FFT has the best runtime which is 3.7 and 37.6 times faster than PCA and DWT, respectively. Among the AD methods, FFT-RX has the best runtime which is almost equal to PCA-RX and 2.3, 48.7, 48.7 and 194.8 times faster than DWT-RX, RX, RX-UTD and Kernel-RX, respectively. These results show that FFT is fast enough to be used in real-time applications by using parallel processing [27].

IV. CONCLUSION

DR as a pre-processing step for RX detector reduces the inter-band spectral redundancy and increases the separation between anomaly and background. Consequently, it can resolve the "small sample size" problem of RX algorithm and improves the detection performance and runtime of this method as well. A new linear dimension reduction method, presented in this paper, uses FFT to implement the DR as a pre-processing step for RX method. The proposed method, named FFT-RX, is compared with some popular detectors: RX, RX-UTD, Kernel-RX, PCA-RX and DWT-RX. To evaluate the performance and runtime of methods two hyperspectral datasets acquired by both the AVIRIS and HyMap sensors, are used. According to the experimental results the proposed DR method improves the detection performance and runtime of RX detector significantly and has a better runtime than the PCA and DWT DR methods. Among all methods, the FFT-RX has the best runtime and detection performance that make it suitable for real-time applications of AD in hyperspectral remotely sensed data. An important problem in all DR methods is determining the amount of dimensionality reduction. According to the experiments, DR rate is directly related to runtime of algorithms, but its relation to the performance is not linear. Future research directions include evaluating reduction rate of DR methods on the performance of AD methods.

ACKNOWLEDGMENTS

The authors would like to thank the Digital Imaging and Remote Sensing group, Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, for providing the HyMap dataset, and the EXELIS VIS Company for making the AVIRIS dataset accessible.

REFERENCES

- [1] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 5-28, 2010.
- [2] D. Borghys, E. Truyen, M. Shimoni, and C. Perneel, "Anomaly detection in hyperspectral images of complex scenes," in *Proceedings of 29th Earsel Symposium*, MAI, Chania, 2009.
- [3] N. Acito, M. Corsini, and A. Cini, "Experimental performance analysis of hyperspectral anomaly detectors," in *Image and Signal Processing for Remote Sensing X*, Bellingham, WA, pp. 41-51, 2004.
- [4] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 10, pp. 1760-1770, 1990.
- [5] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 388-397, 2005.
- [6] S. Khazai, S. Homayouni, A. Safari, and B. Mojaradi, "Anomaly Detection in Hyperspectral Images Based on an Adaptive Support Vector Method," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 646-650, 2011.
- [7] S. Matteoli, M. Diani, and G. Corsini, "Different approaches for improved covariance matrix estimation in hyperspectral anomaly detection," *Proceedings of RiunioneAnnuale GTTI*, pp. 1-8, 2009.
- [8] S. Khazai, A. Safari, S. Homayouni, and B. Mojaradi, "A Performance Comparison Between Recent Anomaly Detectors on Hyperspectral Images" in *Int. Conf. on Sensors and Models in Photogrammetry and Remote Sensing*, Tehran, Iran, 2011.
- [9] L.M. Bruce, C.H. Koger, and J. Li, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2331-2338, Jan. 2002.
- [10] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, vol. 19, no.1, pp. 17-28, 2002.
- [11] L. Ma, M. M. Crawford, and J. Tian, "Anomaly Detection for Hyperspectral Images Based on Robust Locally Linear Embedding," *Journal of Infrared Millimeter and Terahertz Waves*, Springer, pp. 753-762, 2010.
- [12] M. Zare-Baghbidi, K. Jamshidi, A.-R. Naghsh-Nilchi, S. Homayouni, "Improvement of anomaly detection algorithms in hypespectral images using discrete wavelet transform," *Signal and Image Processing: An Int. Journal*, vol. 2, no. 4, pp. 13-25, 2011.

- [13] C.-I. Chang and S.S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no.6, pp. 1314–1325, 2002.
- [14] A. M. Grigoryan, "Fourier transform representation by frequency-time wavelets," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2489–2497, 2005.
- [15] W. T. Cochran et al., "What is the fast Fourier transform?," *Proceedings of the IEEE*, vol. 55, no. 10, pp. 1664-1674, 1967.
- [16] E. Kreyszig, "Fourier Series, Integrals, and Transforms," in *Advanced Engineering Mathematics*, 9th Ed., Wiley, pp. 477-534, 2005.
- [17] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.
- M. Eastaway, 2008, "The Discrete Wavelet Transform," [Online]. Available: <http://cnx.org/content/m18985/latest/> [Accessed 12 November 2012].
- J. Kerekes, D. Snyder, "Target Detection Blind Test," 2009. [Online]. Available: <http://dirsapps.cis.rit.edu/blindtest> [Accessed: 5-May-2012].
- [18] D. Snyder, J. Kerekes, and I. Fairweather, "Development of a web-based application to evaluate target finding algorithms," *IGARSS 2008*, pp. 915–918.
- [19] C. E. Cafer, M. S. Stefanou, E. D. Nielsen, et al., "Analysis of false alarm distributions in the development and evaluation of hyperspectral point target detection algorithms," *Optical Engineering*, vol. 46, no.7, 2007.
- [20] M. Stefanou and J. P. Kerekes, "A method for assessing spectral image utility," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 6, pp. 1698-1706, 2009.
- [21] R.O. Green, M. L. Eastwood, C. M. Sarture et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment*, vol. 65, pp. 227-248, 1998.
- [22] Y. He, D. Liu and S. Yi, "Recursive spectral similarity measure-based band selection for anomaly detection in hyperspectral imagery," *Journal of Optics*, vol. 13, no.1, pp. 1–7, 2011.
- [23] H. R. Goldberg, "A performance characterization of kernel-based algorithms for anomaly detection in hyperspectral imagery," M.S. thesis, University of Maryland, College Park, 2007.
- [24] D. S. Rosario, "Algorithm Development for Hyperspectral Anomaly Detection," Ph.D. dissertation, University of Maryland, College Park, 2008.
- [25] J. M. Molero, A. Paz, E. M. Garzón, et al., "Fast Anomaly Detection in Hyperspectral Images with RX Method on Heterogeneous Clusters," *Journal of Supercomputing*, vol. 58, no. 3, 2011.