Kernel and Stochastic Expectation Maximization: A Fusion Algorithm for Target Detection in Hyperspectral Imagery

M. I. Elbakary[1], M. S. Alam[2]
[1]Department of Computer Engineering
Taif University
Taif, Saudi Arabia
[1,2]Department of Electrical and Computer Engineering
University of South Alabama
Mobile, AL, USA
m.elbakary@tu.edu.sa, malam@usouthal.edu

ABSTRACT: *In this paper, we present a new algorithm for target detection using hyperspectral imagery. The proposed algorithm is inspired by the outstanding performance of nonlinear Kernel method and the robustness of the stochastic expectation maximization (SEM) algorithm. The traditional technique of using SEM algorithm for target detection in hyperspectral imagery is associated with dimensionality reduction of the input data using binning or principal components analysis (PCA) algorithm. Although, the data reduction of the input data is enforced to reduce the computational burden on SEM algorithm, but it affects the results of target detection, especially the challenging one, due to not using the entire information of the potential targets. To facilitate detection of the target by using the entire targets information and simultaneously reducing the computational burden on SEM algorithm, we propose a new scheme for data reduction based on using Kernels. Kernel-based input data reduction is a nonlinear filtering technique in which the input data are mapped to the feature space where most of the background data is filtered using an easily selected threshold. Then, Gaussian mixture model is generated for the reduced input-data and SEM algorithm is employed to estimate the model parameters and to classify that input data. Finally, we allocated the target's class and isolated the target pixels. The proposed scheme for fusion the kernel with SEM algorithm has been tested using real life hyperspectral imagery and the results show superior performance compared to alternate algorithms.*

Keywords: Hyperspectral imagery, Pattern recognition, Target detection, Gaussian mixture model, SEM algorithm

Received: 1 September 2011, Revised 18 October 2011, Accepted 28 October 2011

1. Introduction

Hyperspectral imaging sensors are powerful tools used for detecting and recognizing objects in challenging environment.

Hyperspectral sensors can facilitate automatic target detection in cluttered background since man made objects often differ considerably from the natural background in absorbing and reflecting the radiation at various wavelengths i.e., the identification of the objects (targets) is based on spectral signature of the objects in the scene. Over the past few years, several algorithms for target detection in hyperspectral imagery were presented in the literature [1]-[12]. Statistical based methods suffer from the limited data to train the employed classifier which affects the performance of the classifier [1]. Another category of algorithms were developed based on knowing the covariance matrix of the background which may not be available in many applications [2]. Anomaly detection algorithms are introduced as unsupervised algorithms which suffer from automatic thresholding and hence

false alarms [3]. Also, linear and nonlinear mixing models are introduced in the literature for subpixel detection, which assume that a given pixel is a linear or nonlinear combination of a number of unique deterministic spectral signatures [4]. However, these algorithms are computationally intensive and sensitive to noise and other artifacts. Another class of algorithms based on the independent component analysis (ICA) and principal component analysis (PCA) appeared in the literature, which suffer from poor discrimination ability and high false alarm rates [5]. In general, an algorithm is characterized as supervised if the spectral signature of the target is known *a priori*, otherwise it is considered unsupervised. In this paper, we focus on the supervised methods. Recently, various supervised target detection methods have been developed [5]-[7],[21-22]. Although, an effective solution for object recognition in noisy hyperspectral data remains a challenging problem, this paper introduces a new algorithm for efficient supervised target detection in hyperspectral imagery.

SEM algorithm is employed for classification in remote sensing applications [8]-[12] and for anomaly detection in hyperspectral data [13]. The idea of detecting the targets using SEM algorithm is based on determining the class of the target since the pixels of the same target constitute a class. The implicit assumption in these algorithms and in the proposed algorithm in this paper is that the number of target pixels is larger than the number of bands in the input data [1]. To use the SEM algorithm for classification of high dimensional data, the authors used imagery of few bands, multispectral data [8], or dramatically reduced the dimensionality using PCA or binning [e.g., 16 bands out of 224 bands] [12],[13]. However, the effect of reducing the data dimensionality, the number of bands, on the results of target detection is not certain, [14]. Consequently, in this paper, we introduce a new algorithm for target detection in hyperspectral imagery based on global data classification using the SEM algorithm integrated with a new data reduction technique. In the proposed algorithm, instead of reducing the number of bands, we introduced a novel idea to reduce the background data. Reducing the background pixels and not the number of bands facilitates the detection of targets based on the entire input information of the targets and simultaneously reduce the computational burden on SEM algorithm. The reduction step is accomplished in the feature space where the input data is mapped using the Gaussian RBF kernel [19,20]. This fusion scheme of the SEM algorithm with kernel approach reduced the computational burden on the SEM algorithm and reduced the possibility of false alarms that may arise from noisy input data. Then, Mahalanobis distance and Bayesian decision law [10] are employed to isolate pixels corresponding to the targets of interest.

The rest of the paper is organized as follows: the Gaussian mixture model is described in Section II. Section III summarizes the SEM algorithm. Section IV presents the proposed algorithm. Simulation results are presented in Section V, and concluding remarks are included in Section VI.

2. Gaussian Mixture Model

Mixture models are widely used in many areas including anomaly detection, image segmentation, and computer vision applications [8],[12]. The Gaussian mixture model corresponds to the sum of weighted Gaussian components, where it combines the flexibility of nonparametric methods while exploiting the quantitative advantages of parametric methods [13]. In the Gaussian mixture model, it is assumed that the hyperspectral imagery $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, ...\mathbf{X}_N\}$ in a d - dimensional feature space (R^d) arises from the linear combination of density functions resulting in a mixture *pdf*, expressed as [12]

$$f(\mathrm{X}) = \sum_{l}^{L} \alpha_l \, f_l(\mathrm{X} \,|\, \mu_l, \textstyle\sum_l) \qquad (1)$$

where L is the number of mixture components, and α_l represents the mixing proportions. The mixing proportions must satisfy the constraint $\sum_l^l \alpha_l = 1$, where $0 \le \alpha_l \le 1$, and $l = 1, 2, ..., L$. In Equation. (1), $f_l(\mathbf{X} \mid \mu_l, \sum_l)$ denotes the *pdf* of class l, where μ_l represents the mean vector and $\sum_l$ represents the covariance matrix of that class. The *pdf* for multivariate Gaussian distribution is defined as [10]

$$f_l(\mathbf{X} \mid \mu_l, \Sigma_l) = \frac{1}{(2\pi)^{d/2} \mid \Sigma_l \mid^{1/2}} \exp(-\Delta^2{}_l / 2) \qquad (2)$$

where

$$\Delta^2_l = (\mathbf{X} - \mu_l)^T \overset{-1}{\underset{l}{\Sigma}} (\mathbf{X} - \mu_l)$$

represents the Mahalanobis distance. To estimate the parameters of the Gaussian mixture model, we used SEM for characterizing the spectral classes. The SEM algorithm is based on the expectation maximization (EM) algorithm as discussed in Section III.

3. Sem Algorithm

To estimate the parameters of Equation (1), many iterative methods can be used [9],[15]. However, all of these methods except the SEM suffer from the following constraints [8]:

- the number of classes must be known, and
- the solution depends strongly on the initialization.

These two constraints impose a serious limitation on the performance of these algorithms [8]. For example, knowing the number of classes is a serious problem because in real life images it is highly unlikely to be known *a priori*.

To overcome the aforementioned limitations, the SEM algorithm has been introduced by incorporating a stochastic processing step in the EM algorithm [9]. The EM algorithm is a popular iterative algorithm used for parameter estimation using the maximum likelihood approach. The addition of the stochastic processing step in the EM algorithm results in the following improvements [8]:

- significant improvement in convergence rate,
- only the upper bound on the number of classes needs to be known, and
- unlike the EM algorithm, the solution does not depend on specific initialization.

The main processing steps of the SEM algorithm are summarized in Figure 1.

4. The proposed algorithm

In this section, we present a new algorithm for supervised automatic target detection (ATD) based on fusion of SEM classification algorithm and Gaussian RBF kernel method. The main motivation is using the robustness of SEM with respect to different image contents [8] and hence the robustness in the classification and detection of the targets. However, using SEM with high dimensional imagery is associated with dramatic reducing in the number of bands of the input imagery to few bands [8],[12],[13]. The reduction in the number of bands is enforced to reduce the computational burden on SEM.

In [14], it has been shown that the affect of data reduction, through the number of bands, in target detection algorithms are not certain and still under investigation. In the contrary, keeping all information in the target pixels helps in detecting all the targets especially the difficult one which its spectral signature is similar to the background.

To solve the dilemma of reducing the computational burden on SEM and keeping all the information in the target pixels, we propose a new algorithm for data reduction that reduces the background in the input data and kept all the pixels related to targets intact.

4.1 Data Reduction

To reduce the background pixels in the input data, we used the assumption in [3, 19] that the input data consists of two Gaussian distributions, background distribution and target distribution. In other words, the target pixels are drawn from multivariate Gaussian distribution. The mean vector s_l of that distribution is selected to be the reference target signature, and the covariance matrix Σ is approximated by using a diagonal matrix whose elements correspond to the variance of each band. Figure 2(a) shows one band from a hyperspectral data set and Figure. 2(b) shows the truth mask of the targets of interest in the data. Although, there are 11 targets in the input data but all the targets are the same kind of object [e.g., tank] and the same brand [e.g., M16]. Figures 2(c) shows the spectral signatures of one target in that data set. We expect that 11 targets, the spatial size of each target is approximately 10 x 10 pixels, will constitute together a class of 1100 pixels that represents the targets of the interest.

Mahalanobis distance is effectively used in classification [10] in the input data space, however, classification in the feature space shows outstanding results [19,20]. Therefore, we formulated the proposed data reduction technique in the feature space as following.

1. Obtain an upper bound on the number of classes , L.
2. *Step I (Initialization)*

   Assumes that the probablity of class l has uniform distribution, i.e., $f_l^{o}(\mathbf{X} \mid \mu_l, \Sigma_l)$, where l = 1, 2,..., L and the superscript denotes the iteration number.
3. *Step S (Stochastic):*

   Define a partition Q_1^{n} , Q_2^{n},.....,Q_n^{L} for the data and select the between $P_1^{n}(\mathrm{X})$, $P_2^{n}(\mathrm{X})$,....., $P_L^{n}(\mathrm{X})$. where $P_1^{n}(\mathrm{X})$ is the posterior probablity of class l given the vector **X**. In other words, $\mathbf{X} \in Q_1^{n}$ if $P_1^{n}(\mathbf{X}) = \max_{l=1,2,...L} P_l(\mathbf{X})$
4. *Step M(Maximization)*

   Estimate the empirical mean μ_l^{k+1} *and the covariance matrix* Σ_l^{n+1} of each class l, where l = 1, 2,..., L, using the following equations:

$$\mu_l^{n+1} = \frac{1}{C_1^n} \sum_{i=1}^{C_1^n} X_{i,l}^n \tag{3}$$

$$\sum_l^{n+1} = \frac{1}{C_1^n} \sum_{i=1}^{C_1^n} (X_{i,l}^n - \mu^{n+1})(X_{i,l}^n - \mu^{n+1})^T \tag{4}$$

   where , C_1^{n} is the number of pixels in class l.

   • The prior is defined as $\alpha_l^{n+1} = \frac{C_1^{n}}{N}$, where N is the total number of pixels in the image. For α class l, if the prior ais less than the prescribed threshold , then eliminate that class and go to the *step I*.

   • Compare μ_l^{n} with μ_l^{n+1} and if the difference falls below the prescribed threshold, terminate the algorithm.
5. *Step E (Estimation)*

   • Define $P_1^{n}(\mathrm{X})$, $P_2^{n}(\mathrm{X})$,....., $P_L^{n}(\mathrm{X})$ fro each **X** on the set of classes using maximum a posteriori (MAP) distribution based on the current values of the mean, the covariance matrix, and the prior(i.e.,) μ_l^{n+1}, Σ_l^{n+1}, α_l^{n+1} , where l = 1, 2,..., L as follows.

$$P_l^{n+1}(X) = \frac{\alpha_l^{n+1} f_l^{n+1}(X)}{\sum_{i=1}^{L} \alpha_{i,l}^{n+1} f_{i,l}^{n+1}(X)} \tag{5}$$

   where f_l^{n+1} designates the multivate Gaussian distribution corresponding to μ_l^{n+1} and Σ_l^{n+1}, respectively.
6. Return to *Step S.*

Figure 1. The SEM algorithm

Define the Mahalanobis distance between any pixel vector X in the data and the target's distribution by the following equation

$$\Delta^2(X) = (X - \mu_s)^T \sum^{-1} (X - \mu_s) \quad (6)$$

Then, we mapped $\Delta^2(X)$ to the feature space using the Gaussian RBF kernel as following

$$V(X) = exp\,(\Delta^2(X)) \quad (7)$$

Equation (7) is a nonlinear mapping of the Mahalanobis distance into the feature space. To reduce the background in the input data, we removed those pixels from the input data whose $V(X)$ values (Equation (7)) are less than an experimentally determined threshold value. The selected value of the threshold is depending on the amount of reduction required in the data cube. Let t represents the selected threshold, X represents a pixel vector in the input data, and C represents the set of data of interest which includes the target pixels,

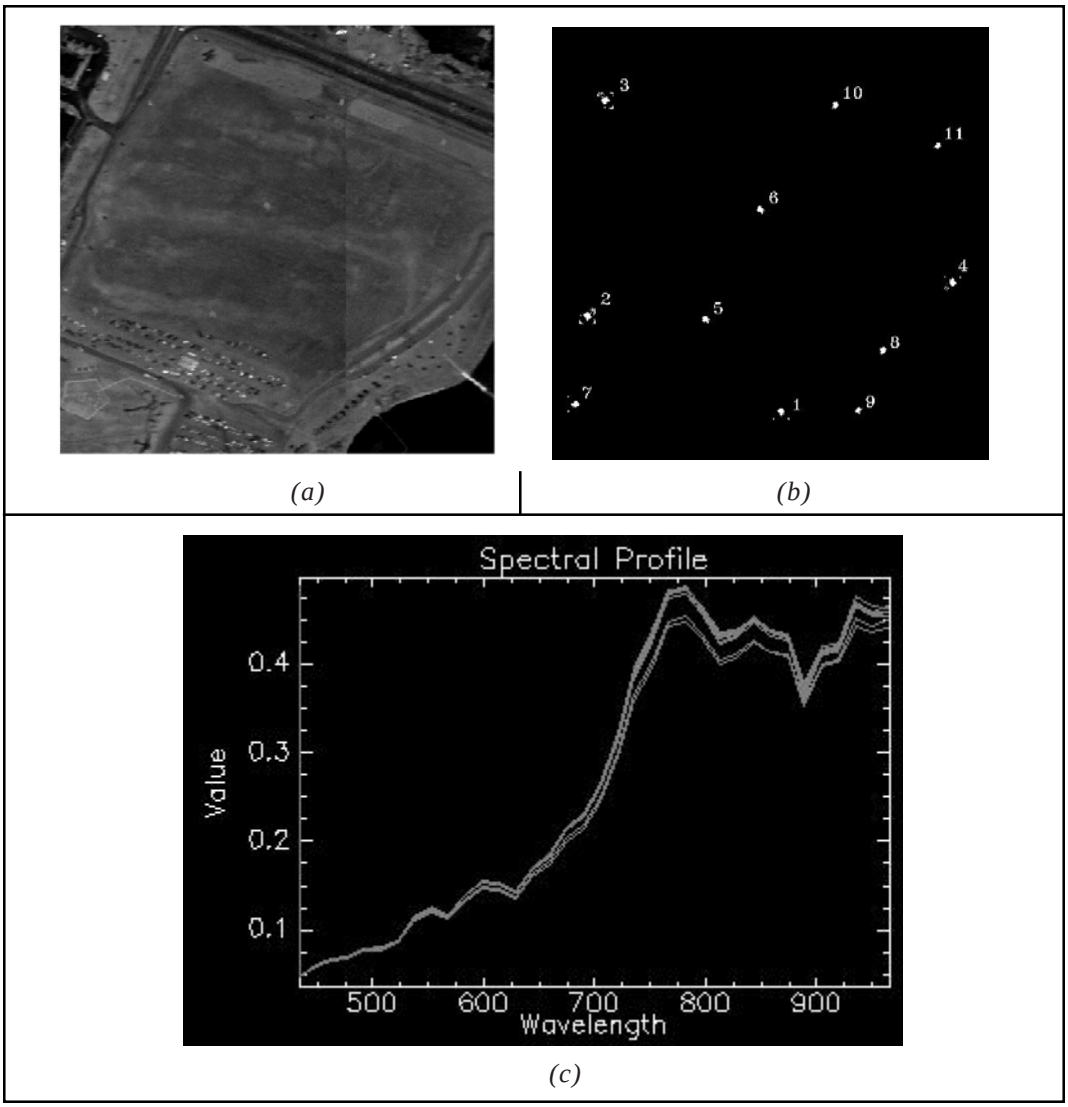


Figure 2. *(a)* One band from hyperspectral data set, *(b)* the ground truth of targets in that data set , (*c*) the spectral signatures of one target in that hyperspectral data

then

$$X \in C \text{ if } V(X) \geq t$$

$$X \notin C \text{ if } V(X) < t \tag{8}$$

This nonlinear transformation shown in Equation (8) successfully filters out most of the background pixels and significantly reduces the amount of data to be processed in the subsequent SEM algorithm. This data reduction technique is integrated with the traditional SEM algorithm and the combination is called enhanced SEM, ESEM, algorithm as explained in the following section.

4.2 Classification Using Enhanced SEM (ESEM)

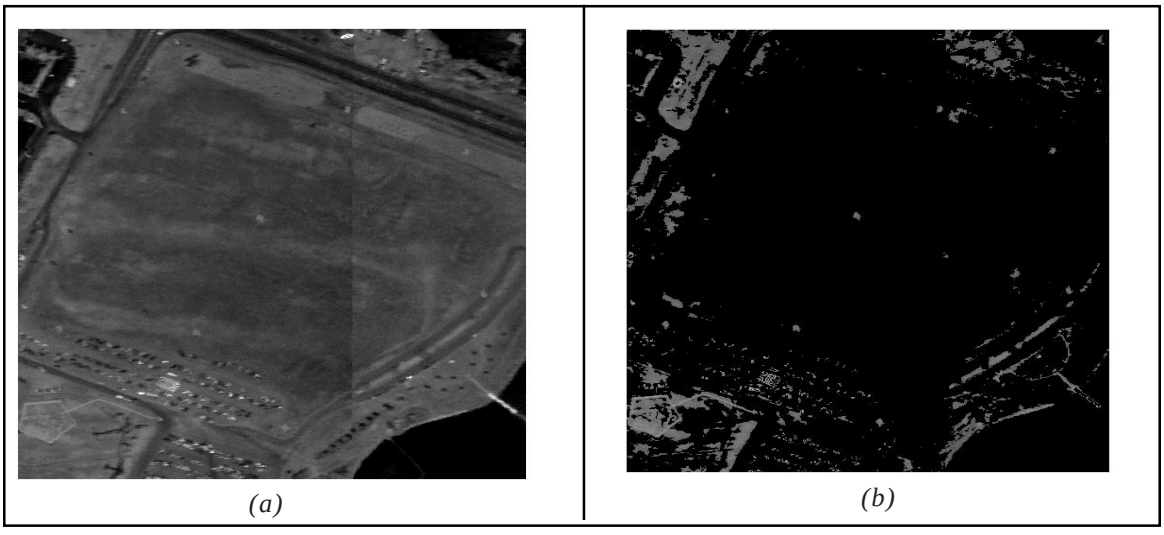
SEM is presented as viable classification algorithm in the remote sensing and many numerous simulations show correct behavior of that algorithm [8]. Integrating the reduction step with the SEM algorithm reduced the computational burden on SEM and facilitated the detection of the pixels related to the target of interest using all the spectral information in the target pixels. Figure 3(a) shows one band from a hyperspectral imagery before the application of the proposed data reduction technique, while Figure 3(b) shows the same band after the application of the data reduction technique. From Figure 3(b), we observe dramatic reduction in the amount of background pixels which reduces significantly the computational burden in the classification.

4.3 Target Isolation

Once the input data is classified into different unique classes using the ESEM algorithm, we identify automatically the class of interest [class of the targets] by computing the traditional Mahalanobis distance between the reference spectral signature and all the classes. Then, we choose the class with smallest Mahalanobis distance as the target class. Finally, the Bayesian decision law [10] is employed to isolate the target pixels in the hyperspectral imagery. Figure 4 depicts the basic building blocks of the proposed algorithm.

The salient features of the proposed algorithm can be summarized as follows:

- alleviate the effect of noise in a preprocessing step using the average filter along each pixel vector,
- reduce the background pixels in the input data using the feature space,
- generate Gaussian mixture model for the reduced data,
- reduce the computational burden on SEM algorithm by integrating it with the data reduction technique, fusion step to build ESEM.
- employ the ESEM algorithm to estimate the parameters of the mixture model
- identify the target class using the traditional Mahalanobis distance.
- employ MAP algorithm or any suitable method to isolate the pixels of the target class.



(a) *(b)*

Figure 3. *(a)* One band from Hyperspectral imagery data cube and *(b)* the corresponding band after data reduction

5. Experimental Results

In this section, experimental results are presented in order to demonstrate the efficacy and robustness of the proposed algorithm compared to alternate techniques. Real life hyperspectral data are employed for performance evaluation of the proposed technique.

5.1 Data sets

To verify the accuracy and robustness of the proposed algorithm, we conducted many experiments using hyperspectral data sets with different levels of noise. The hyperspectral data sets used in this work were acquired by Compact Airborne Spectrographic Imager (CASI) sensor. The data sets contains 36 bands, each band consists of 512x512 pixels and covers the spectral wavelength range of 433 nm to 965 nm with spectral resolution of 15 nm. The reference spectral signature for each data set is computed from training samples using the truth mask of that data. It is worth to say that the data is provided in the form of ready to use by a well known company, please see the acknowledgement.

5.2 Results and Discussions

In the first experiment, we run the algorithm on Dataset1 which contains 11 targets from the same kind and the same brand as explained in subsection A in Section IV. The result obtained after the application of the proposed algorithm for that data set is shown in Figure 5. The truth mask is shown in Figure 5(a) and the result of the algorithm in Figure 5(b). From Figure 5(b), we observe that the proposed algorithm detected all targets in the data cube with zero false alarms. Another way to show the results is to use the confusion matrix. To generate the ground truth of the false alarms, we divided the image plane into blocks with a block size of 10x10 pixels. We chose 10x10 pixels based on the size of targets in the ground truth of the targets. Excluding targets' blocks, the rest are the potential false alarms. Therefore, for this data set, the probability of detection is $P_d = 1$, and the probability of false alarms is $P_{FA} = 0$ as shown in the following confusion matrix.

Confusion matrix	Targets	False alarms
Targets	11	0
False alarms	0	2590

To show the advantage of the proposed algorithm over alternate algorithms, we considered Spectral Angle Mapper (SAM) algorithm [5]. Constrained Energy Minimization (CEM) algorithm [20] is considered for the same purpose but its results show that it is sensitive for the noise and is not suitable for target detection in the current data sets. It is worth to pay attention to the thresholds selected in SAM for all the conducted experiments in this paper are selected to detect all the targets in the input data. Also, we considered that the target is detected if at least only one pixel of that target is detected.

The results of SAM demonstrates that SAM detected all targets in the scene with $P_d = 1$, but with 34 false alarms i.e., $P_{FA} = 0.013$ as shown in the following confusion matrix and Figure 5(c).

Confusion matrix	Targets	False alarms
Targets	11	0
False alarms	0	2590 - 34 = 2556

To investigate the performance of the proposed algorithm against scene details and noise, we used another challenging hyperspectral data set (Dataset2) which is acquired using the same sensor. The input data contains 10 targets from the same kind and brand. Again, to generate the ground truth of the false alarms, we divided the image plane into blocks with a block size of 5 x 5 pixels. This data set is challenging since the spectral signatures and many details in the scene mimics the target signature. Additionally, two noisy versions of that data set were generated by corrupting the input pixel-vectors via 5% and 10% additive random noise [i.e., the spectral are degraded by random noise, but the spatial information maintained]. This process of noise addition produced two versions from the data set corresponding to the noise levels. One band from Dataset2 is shown in Figure. 6(a) and the corresponding truth mask for the targets (10 targets), is shown in Figure 6(b). The results of using the proposed algorithm for the two data sets are shown in Figure 6(c) and 6(d), respectively. The confusion matrix for 5% noisy data is shown below which yields $P_d = 1$ and $P_{FA} = 0.00076$, respectively.

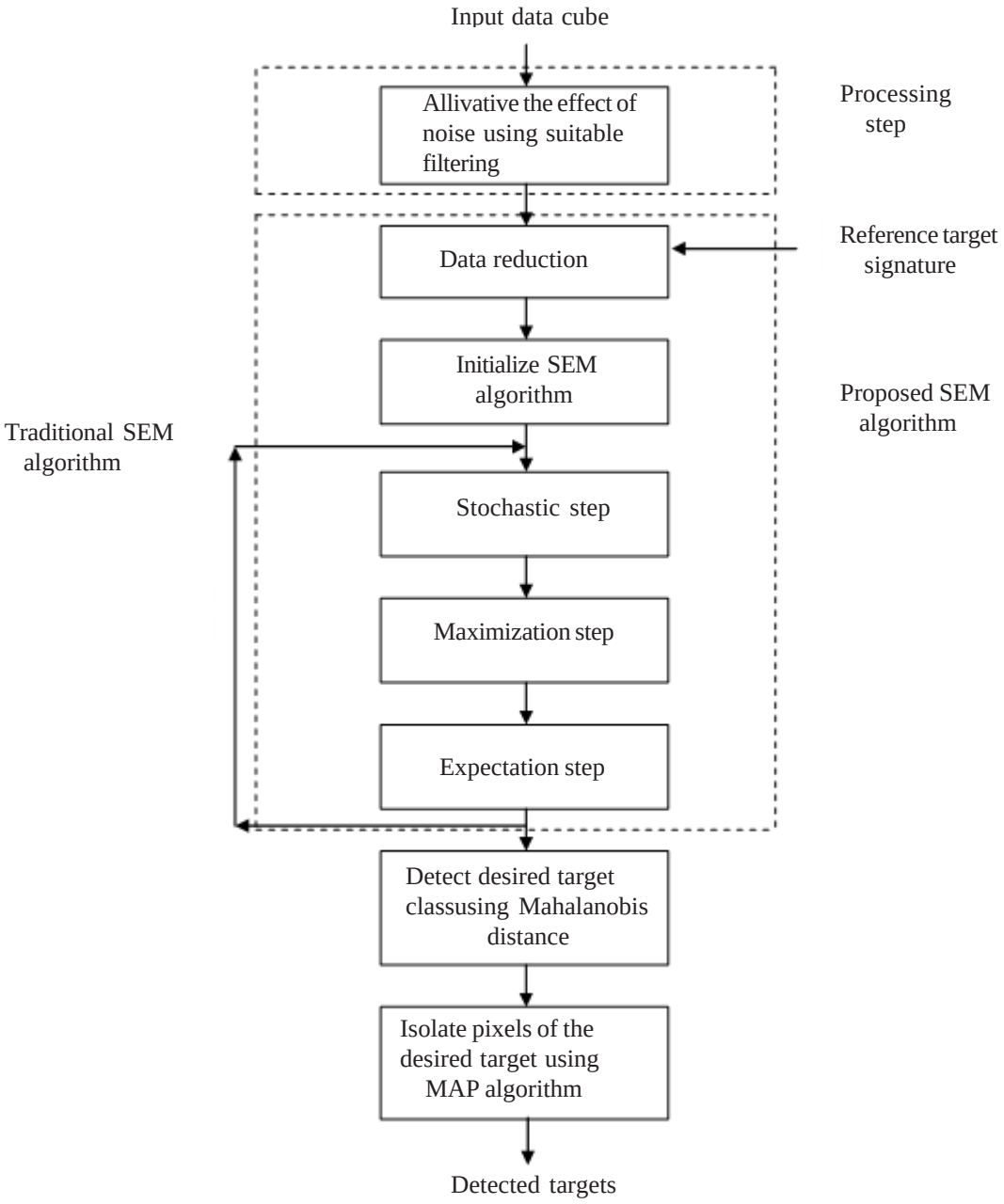


Figure 4. The flowchart of the proposed algorithm

Confusion matrix	Targets	False alarms
Targets	10	8
False alarms	0	10476 – 8=10468

The confusion matrix for 10% noisy data is depicted below which yields $P_d = 1$, and $P_{FA} = 0.0041$, respectively

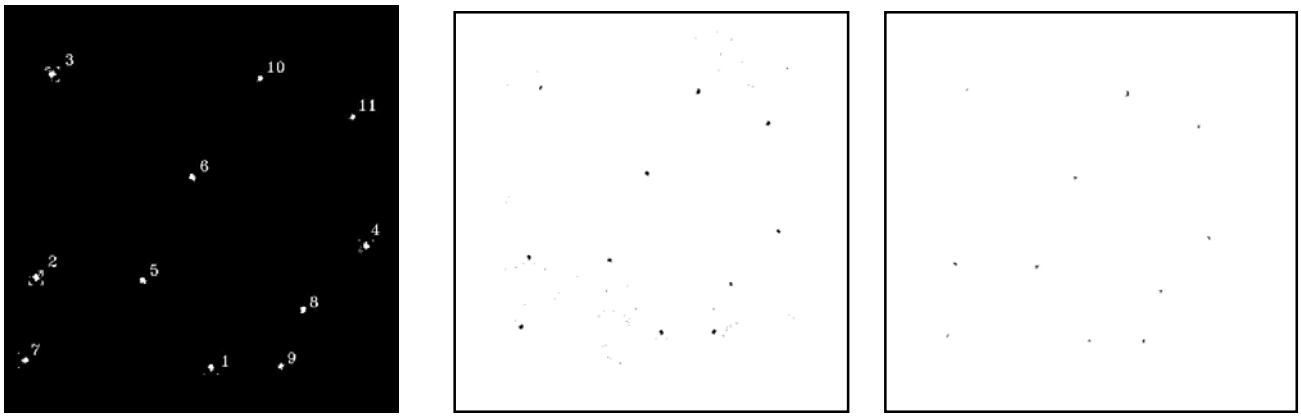
Confusion matrix	Targets	False alarms
Targets	10	8
False alarms	0	10476 – 43=10433

The results obtained by the proposed technique are much better than the corresponding results generated by using the SAM algorithm for the same data sets.. The SAM algorithm detects all targets but with numerous false alarms. The confusion matrix for 5% noisy data is shown below which shows that $P_d = 1$, and $P_{FA} = 0.1414$, respectively.

Confusion matrix	Targets	False alarms
Targets	10	1481
False alarms	0	10476 – 1481=8995

The confusion matrix for 10% noisy data is presented below and yields $P_d = 1$, and $P_{FA} = 0.3926$, respectively

Confusion matrix	Targets	False alarms
Targets	10	4113
False alarms	0	10476 – 4113=6363



(a) *(b)* *(c)*

Figure 5. Detection results for Dataset1. (a) Truth mask of the targets in the data cube, (b) results obtained using the proposed algorithm, and (c) results obtained using the SAM algorithm.

Comparing the above mentioned results, it is obvious that the performance of the proposed algorithm is robust with respect to noise and minute details in the input scene.

Finally, to ensure the robustness of the algorithm, we used an entirely different dataset (Dataset3). Again the data contains targets from the same kind and brand. Two versions of the noisy data sets (5% and 10%) were generated by following the procedure described earlier. Dramatic reduction in the background is obviously noticed and the confusion matrix of results of the proposed algorithm for 5% noisy data is shown in the following table which yields $P_d = 1$, and $P_{FA} = 0$, respectively.

Confusion matrix	Targets	False alarms
Targets	10	0
False alarms	0	10476 – 0 = 10476

The confusion matrix for 10% noisy data is shown below and shows $P_d = 1$, and $P_{FA} = 0.000095$, respectively

Confusion matrix	Targets	False alarms
Targets	10	1
False alarms	0	10476 -1 = 10475

The corresponding results obtained using the SAM algorithm are tabulated below. The confusion matrix for the results from 5% noisy data is depicted below which shows that $P_d = 1$, and $P_{FA} = 0.0266$, respectively.

Confusion matrix	Targets	False alarms
Targets	10	279
False alarms	0	10476 -279=10197

The confusion matrix for the results from 10% noisy data is shown below and yields $P_d = 1$, and $P_{FA} = 0.1199$, respectively.

Confusion matrix	Targets	False alarms
Targets	10	1256
False alarms	0	10476 -1256 = 9220

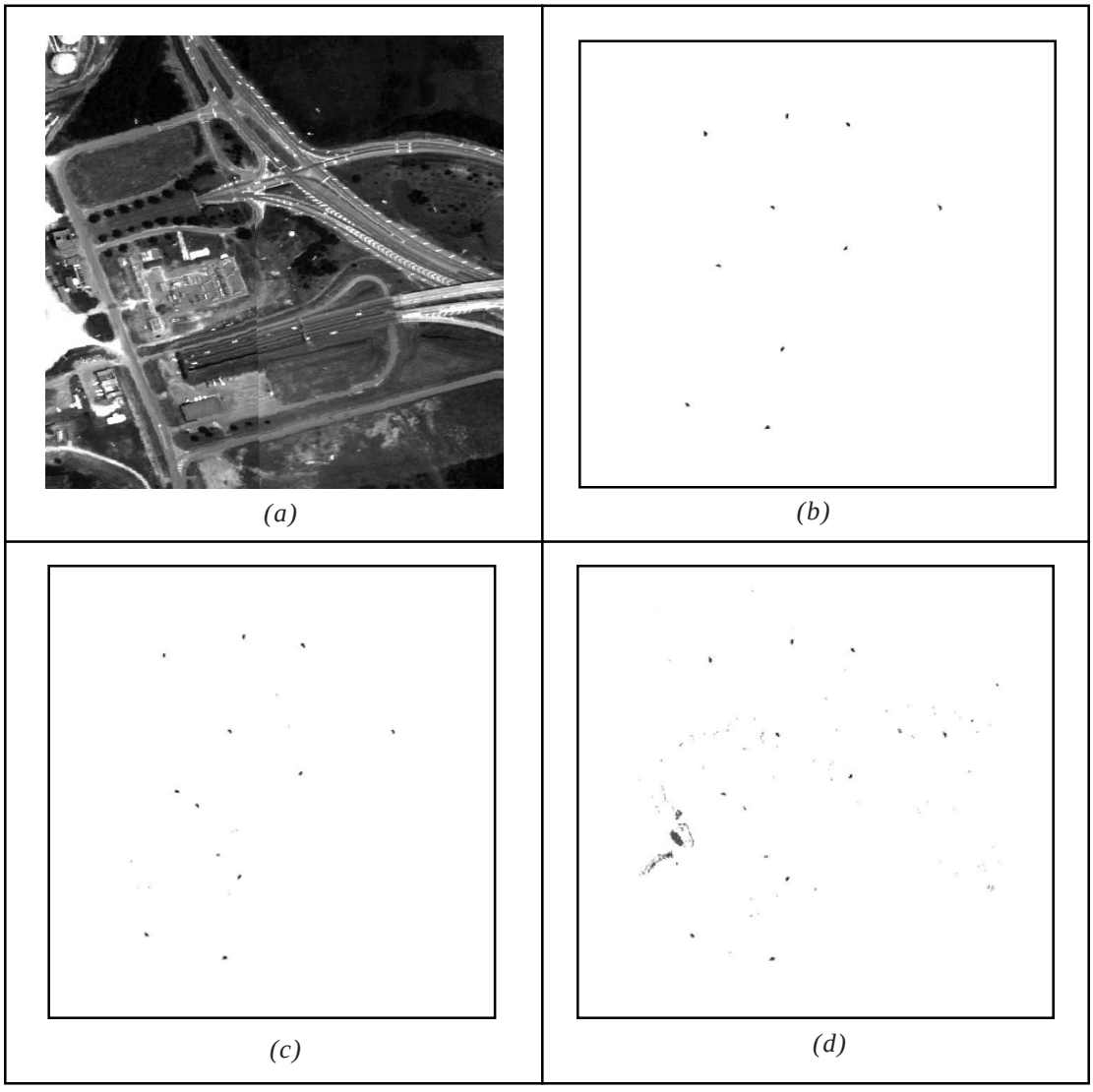


Figure 6. (a). One band from the data cube, (b) the truth mask of the targets in the scene, (c) the result of the algorithm for 5% noisy data set, (d) the result of the algorithm for 10% noisy data set

From the above listed results, it is obvious that the proposed algorithm yields superior performance compared to alternate algorithms.

6. Conclusion

In this paper, we presented a new algorithm for target detection in hyperspectral imagery based on the fusion of SEM algorithm and kernel method. The proposed scheme exploits the robustness of SEM algorithm against image contents in classification and the outstanding performance of kernel method. To facilitate detection of the target in a huge data set by using the entire information of the target without data compression to reduce the computational burden on SEM, a novel data reduction technique is introduced. The data reduction technique is based on reducing the background pixels in feature space using kernel. The proposed algorithm shows superior performance compared to alternate algorithms, such as the SAM and CEM. algorithms, for challenging scenarios. The proposed algorithm might be used for detection in applications such as military targets, materials in mining industries, and objects for rescue operations.

7. Acknowledgement

The authors would like to thank the anonymous reviewers for their constructive suggestions. The authors would also like to thank Radiance Tech for providing the hyperspectral data sets used in this work.

References

[1] Hoffbeck , J. P., Landgrebe, D. A. (1996).Covariance matrix estimation and classification with limited training data,. *IEEE Tran. on Pattern Analysis and Machine Intelligence,* 18 (7) 763-767.

[2] Muhammed, H. H. (2005).Hyperspectral crop reflectance data for characterizing and estimating fungal disease severity in wheat. *Biosyst Eng*., 91 (1) 9–20.

[3] Reed, I. S. and Yu, X. (1990).Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing,* 38 (10) 1760.

[4] Wang, J. and Chang, C. I. (2006). Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery. *IEEE Trans. Geoscience and Remote Sensing,* 44 (9) 2601–2616.

[5] Manolakis, D., Marden, D., Shaw, G. (2003). Hyperspectral image processing for automatic target detection applications. *Lincoln Laboratory journal,* 14 (1) 79-115.

[6] Alam, M. S., Ochilov, S. (2006). Pattern detection in hyperspectral imagery using one dimensional fringe-adjusted joint transform correlation, *In:* Proc. of SPIE, Orlando, Florida, April 15-19.

[7] Manolakis, D., Shaw, G., Keshava, N. (2000). Comparative analysis of hyperspectral adaptive matched filters detectors. Proc. SPIE, 2-17.

[8] Masson, P., Pieczynski, W. (1993). SEM algorithm and unsupervised statistic segmentation of satellite images. *IEEE Trans. On Geosciences and remote sensing,* 31. 618- 633.

[9] Chauveau, D. (1995) A Stochastic EM algorithm for mixtures with censored data. *Journal of Statistical Planning and Inference,* 46, 1-25.

[10] Duda, R. O., Hart,P. E., Stork, D. G. (2001). Pattern classification. Second edition, John Wiley & Sons Publication.

[11] ELbakary, M., Alam, M. S. (2006) Pattern recognition in multiband imagery using stochastic expectation maximization, *In:* Proc. of SPIE, 6311, p. 63110 W(1-10).

[12] Beaven, S. G., Hoff, L. E., Winter, E. M. (1999).Comparison of SEM and linear unmixing approaches for classification of spectral data. Proc. SPIE, V. 3753, p. 300-307.

[13] Willis, C. J. (2004) Mixture models for anomaly detection in hyperspectral imagery, *In*: Proc. SPIE, V. 5613, p. 119-128.

[14] Farrell, M., Mersereau, R. (2005). On the impact of PCA dimension reduction for hyperspectral detection of difficult targets. *IEEE Trans. On Geosciences and Rremote Ssensing Letters*, 2 (2) 192-195.

[15] Alam, M. S., Bognar, J. G., Cain, S., Yasuda, B. J. (1998). Fast registration and reconstruction of aliased low resolution frames using a modified maximum likelihood approach. *Journal of Applied Optics,* Vol. 37, p. 1319-1328.

[16] Landgrebe, D. (1999). Some fundamentals and methods for hyperspectral image data analysis. SPIE Photonics West, San Jose, CA, Jan. 23-29.

[17] Lennon, M., Mercier, G. (2003). Noise-adjusted non orthogonal linear projections for hyperspectral data analysis. IEEE IGARSS'03, p. 3760-3762.

[18] Stein , D., Beaven , S., Hoff , L., Winter , E., Schaum , A., Stocker, A. (2002) Anomaly detection from hyperspectral imagery. *IEEE Signal Processing Magazine,* p. 58-69.

[19] Kwon, H., Nasrabadi, M. N. (2005) Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery. IEEE *Trans. on Geoscience and Remote Sensing,* 43 (2) 388-397.

[20] Ruiz , A., Lopez-de-Teruel, P.E (2003). Nonlinear kernel-based statistical pattern analysis. *IEEE Trans. on Neural Networks,* 12 (1) 16-32.

[21] Shi, Z., Yang, S., Jiang, Z. (2011). Hyperspectral target detection using regularized high-order matched filter. *Opt. Eng*. 50 (05) 057201.

[22] Sakla, A., Sakla, W., Alam, S. M. (2011). Hyperspectral target detection via discrete wavelet-based spectral fringe-adjusted joint transform correlation. *Applied Optics*, 50 (28) 5545-5554.