A Background Self-Learning Framework for Unstructured Target Detectors

Ting Wang, Bo Du, and Liangpei Zhang, *Senior Member, IEEE*

***Abstract*—Unstructured background model based detectors have been successfully applied in various hyperspectral target detection applications. The background statistics of an image can be estimated in a global way or a local way. The global approach involves modeling the background directly from the whole image, which can prove to be inaccurate due to target contamination of the background information. The local approach usually involves estimating the background statistics using a spatially sliding local window. However, this approach can also fail to reflect reality, due to sensitive parameters, like the window size, and presents high computational costs. This letter proposes a self-learning method to adaptively determine the background statistics for unstructured detectors, with the consideration of exploiting both the spatial and spectral information, and accelerating the computation speed. The experimental results with two real hyperspectral images confirm the superior performance when compared to the other two approaches to modeling background statistics.**

***Index Terms*—Background self-learning, global detector, local detector, target detection.**

I. Introduction

HYPERSPECTRAL remote sensing has received considerable interest in recent years as hyperspectral imaging sensors have both broadband wavelength and very fine spectral resolution, which can provide critical information for material classification and object identification [1]. Target detection has been undertaken from both full pixels and mixed pixels in many different scenarios, including the detection of relatively small targets such as vehicles and aircrafts, and intrinsically mixed pixel targets such as minerals and crop species [2]. Target detection has become an increasingly ubiquitous and challenging area, especially when the targets of interest lie in mixed pixels [3]. The performance of the conventional algorithms is susceptible to inaccurate assumptions and an inability to discriminate the subtle signatures from noise [2]. This is mainly caused by inaccurate models describing the background information of the image. In many state-of-the-art unstructured background detection algorithms, a single multivariate Gaussian random variable with additive independent and identically distributed (i.i.d.) Gaussian noise is used to globally model the background. These include the Kelly's generalized likelihood ratio test (K'GLRT) [4], the adaptive cosine estimator (ACE) [5], and the adaptive matched filter (AMF) [6]. Although they are widely used, the accuracy of these background description models has become the main limiting factor for obtaining a high confidence and low false alarm detection from mixed pixel targets [7], [8]. In addition, a number of other methods have been developed for similar purposes, such as general tensor discriminant analysis and Gabor features for gait recognition [9], representative multiple kernel learning for classification in hyperspectral imagery [10], a kernel-based target-constrained interference-minimized filter for hyperspectral sub-pixel target detection [11], supervised tensor learning [13], unsupervised transfer learning for target detection from hyperspectral images, and the geometric mean for subspace selection [13].

In general, globally determining the background statistics is the easiest but the least accurate way of describing the background model. Removing the probable anomalies and targets requires an accurate grasp of the amount of anomalies and target candidates in the scene; therefore, the detection results of global-based anomaly-removal detectors are consequently very sensitive to the threshold set to these pixels, which makes the performance susceptible to the advanced estimation of anomalies and targets. In addition, the global approaches to estimating the background characteristics do not fully exploit the spatial relationship between the high-dimensional pixels, which is not the best way to use the hyperspectral image information.

With the aim of overcoming the problems stated above, this letter proposes a new background self-learning (BSL) approach, to automatically and rapidly suppress the background for a better detection result, with easy parameter selection.

II. Background Self-Learning Framework

In the traditional methods, including KGLRT, ACE, and AMF, the background statistics are directly calculated from the image data set but, as stated above, they are not, theoretically, robust enough. Global moment based detectors are subject to the biases resulting from the inclusion of possible targets and anomalies in the background statistics.

The local methods have been well researched over the years. These approaches take the spatial information into consideration and apply the detectors with locally constructed statistics, like the mean and covariance. One common way of applying a local-based detector is to process the spatial segmentations of the image, and then choose a region that is spatially proximal to the test pixel to estimate the background statistics [8].

When sliding windows are applied to detectors, an exclusion window is set up to prevent the selection of possible targets

Manuscript received March 2, 2013; revised April 28, 2013; accepted April 29, 2013. Date of publication July 22, 2013; date of current version October 10, 2013. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2012CB719905 and 2011CB707105, the National Natural Science Foundation of China under Grants 40930532 and 61102128.

T. Wang and L. Zhang are with the Remote Sensing Group, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: wangtingwhu@126.com; zlp62@public.wh.hb.cn).

B. Du is with the School of Computer Science, Wuhan University, Wuhan 430072, China (e-mail: gunspace@163.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2013.2262133

in the local area, which is also called an inner window. The pixels in the inner window are neglected, after which the pixels left in the outer window are used to compute the statistics for the test pixel. What is worth mentioning here is that the window size of both the outer and inner window has a great impact on the detection performance. The inner window should exclude the possible targets; however, in real applications, the size of the targets is often unknown, and, even if known, the targets may either vary in size or be of irregular shape the local approaches are always computationally expensive as they need to calculate the inverse of the covariance matrix for as many times as the number of pixels, and the computational speed will be even slower with the enlarging of the outer window. A more appropriate model for the background is a varying multivariate normal distribution corresponding to the different kinds of materials. Therefore, organizing the background pixels into different classes according to their spectral information, with varying background statistics, is a better way to reflect the real features of certain kinds of materials. When the detection is conducted, the criterion of choosing the suitable background elements from the full candidate sub-background set should be determined both by the test pixel and the surrounding it is embedded in. This letter proposes an algorithm with high automaticity. It aims at enabling the traditional unstructured background with the ability to better suppress background, with adaptive and flexible statistics, instead of the background estimated in a fixed mode from the image, as well as decreasing the computational costs. By labeling each pixel with a sorted cluster number controlled by a spatial information based decision, the hypotheses to determine if a pixel is target-class or non-target-class will be more accurate.

Specifically, the framework can be stated as five steps, as follows.

A. Estimate the Endmember Numbers in the Image

The hyperspectral signal identification by the minimum error (HySime) technique is eigen decomposition based, unsupervised, and fully automatic (without tuning parameters) [15]. The BSL algorithm makes use of HySime to automatically determine the number of clusters or endmembers in the image. An initial number kf can be obtained from HySime. We then automatically divide the data into a proper number of clusters by conducting a k-means unsupervised classification.

B. Cluster Merging in the Background Self-Learning Framework

As the number kf is just a reference for k-means clustering to obtain a stable and reliable number of clusters. By evaluating the similarity between two vectors, the spectral angle criterion is utilized. In detail, the center of each cluster is calculated and compared to the others, and the clusters with spectral angles between the other centers smaller than the given threshold are merged. In our experiments, the spectral angles are calculated pairwise, then compared to the threshold Ts. The threshold is set to 0.2 according to an empirical value to differentiate one cluster from the others, which has been analyzed in detail in Section III. In addition, the clusters with a small number of group members are merged into the spectrally closest cluster, in the case of the covariance matrix calculated from the cluster being ill-conditioned. In our experiments, the smallest number of pixels required to construct a cluster is equal to the band number. After the adjustment, the data are re-clustered into a new number of clusters, whose statistics can then be calculated.

C. Target Contaminating Cluster Exclusion in the Background Self-Learning Framework

In some circumstances, the target class can contain a large amount of pixels, which could possibly become a cluster after the merging stage. To prevent this cluster from being collected into our candidate sub-background set, the cluster in the re-clustering results with the most similar spectral features to the priori target, which is calculated from the spectral angle criteria, is abandoned if the spectral angle is less than the given threshold Tt. If there is no cluster that is related to the priori target, this step will be skipped. As Tt is set equally to Ts, so far there are only two parameters in our framework: Ts and the smallest cluster pixel number, which are both fixed for given data.

D. Different Ways of Assigning the Test Pixel to a Cluster

The final detection process should be an assigning process before the unstructured background detector, which labels the test pixels into an appointed cluster and executes the detection with the statistic of that cluster. There are, however, two ways of choosing the background characteristics of the properly assigned clusters for the test pixels, as illustrated in the diagram shown in Fig. 1. As false alarms often occur due to the confusion of those pixels with similar spectra to the targets. The "target guided" (TG) method uses the cluster that is spectrally closest to the target signature to estimate the background statistics [7]. When deciding the proper cluster features for the test pixels, the dark red cluster center in Fig. 1 is chosen since it is spectrally closest to the priori target. In this way, the estimated covariance can best suppress the pixels that lead to false alarms, as they may not be so distinguishable in the target-similar environment. Another approach is called "pixel guided" (PG), which uses the spectrally closest cluster to the test pixels to calculate the mean and covariance matrix, shown as the purple cluster in Fig. 1. This approach is more practicable than the TG method, as TG only works on the premise that the possible targets and anomalies are excluded before clustering. As mentioned above, there are no exclusions for anomalies and possible targets in our work; the principle of the PG method is then chosen to assign the test pixels into clusters.

E. Strategic Decision for the Labels Combining the Spatial and Spectral Information

When the labels of the test pixels are being determined, the spatial information should also play a key role. To combine the spectral information with the advantage of locally estimated background information, a sliding local window of the test pixel is established to decide which cluster the test pixel is more likely to belong to. In our framework the label of the test pixel is assigned to the cluster spectrally closest to a locally determined "pixel." This more representative "pixel" is calculated from the average of its neighborhood pixels in that area, to better describe the test pixel when it is unknowingly mixed with other spatially adjacent materials. This is especially valid when

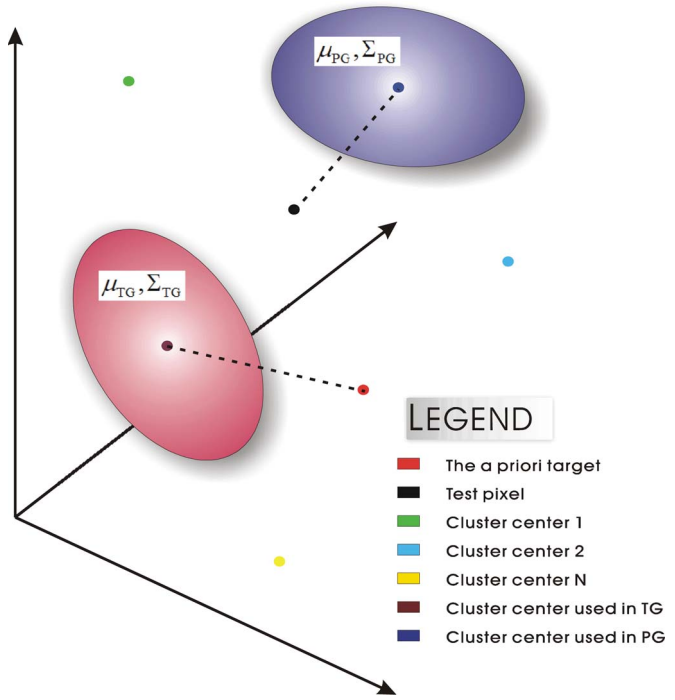

Fig. 1. Diagram illustrating how to properly assign clusters for the test pixels.

detecting the mixed pixels, as the spatial average vector can fully exploit the information of the composition of the mixed pixel, and thus assign the pixel to a more precise cluster.

As mentioned above, in the re-clustering stage, the cluster closest to the priori target is deleted from the image, and the pixels in that area are assigned to a status without a cluster label. These pixels are then re-labeled with the cluster label closest to the deleted cluster in the remaining clusters, to make sure every pixel in the image is properly assigned to a cluster.

The BSL method also takes the computational costs into consideration. In the BSL method, only a limited number of covariance matrices need to be calculated, corresponding to the cluster number, which greatly saves on detection time.

III. Experiments

A series of experiments were carried out to test the three ways of modeling the background. These different background estimation methods were conducted before the use of three unstructured background target detectors: ACE, AMF, and Kelly's GLRT (KGLRT). Two benchmark hyperspectral image data sets were chosen to validate the proposed BSL-based detectors: the RIT data set and the San Diego data set.

The RIT image was collected by the HyMap (Hyperspectral Mapper) sensor, covering the small town of Cooke City, MT, USA. HyMap is an airborne sensor, which was flown at approximately 1.4 km above the terrain, yielding a 3 m ground resolution. A total of 126 bands remained after removing the water absorption bands and low-SNR bands, covering the wavelength range of 0.45–2.5 μm. A sub-image containing only F1 targets, surrounded by a relatively monotonous background, was chosen, and the false color image is shown in Fig. 2(a). The target ground truth is presented in Fig. 2(b), according to the research project.

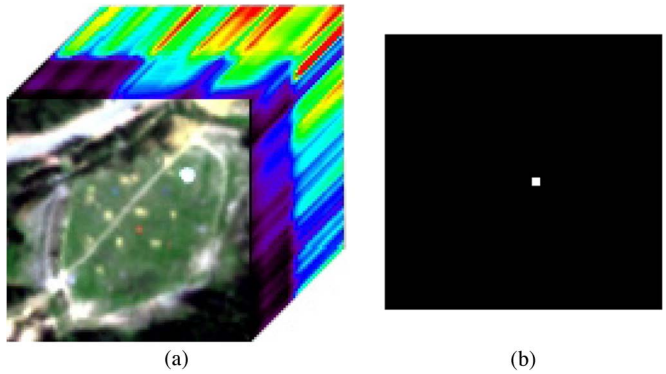

Fig. 2. HyMap data set and the distribution of the targets. (a) Hyperspectral cube of the RIT image; (b) true distribution of the targets.

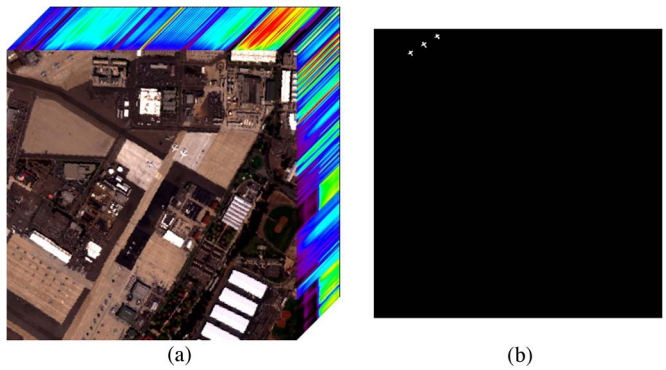

Fig. 3. AVIRIS data set and the distribution of the targets. (a) Hyperspectral cube of the San Diego image; (b) The true locations of the targets.

The San Diego image was collected by AVIRIS (the Airborne Visible/Infrared Imaging Spectrometer), which is a spaceborne sensor, covering a naval air station in San Diego, CA, with 400×400 pixels, and the color image is shown in Fig. 3(a). A total of 189 bands remained after removing the water absorption bands and low-SNR bands, covering the wavelength range of 0.4–1.8 μm. The image is mainly composed of buildings with different roofs, parking aprons with different materials, an airport runway, and a small quantity of vegetation. Unlike the RIT data set, the detection difficulty of this image is the complex composition of the background materials. We chose the airplanes lying in the upper-left corner as the targets to detect. The location of the targets is shown in Fig. 3(b).

A. HyMap Dataset

GLBACE, LCACE, and BSLACE were conducted on the RIT image to detect the F1 target panel. The priori spectra of the F1 target was provided by the RIT website. In LCACE, the size of the inner and outer window was set as 5×5 and 15×15, respectively; in BSLACE, the window size was set as 5×5. The window size of the LC-based algorithm is chosen on the principle that the inner window size is bigger than the target size, and the number of pixels between the inner window and the outer window is larger than the number of bands. The window size of the BSL-based algorithm is set as equal to the inner window size of the corresponding LC-based algorithm. All the three ways of estimation of the mean and covariance matrix worked on this HyMap data set, which is composed

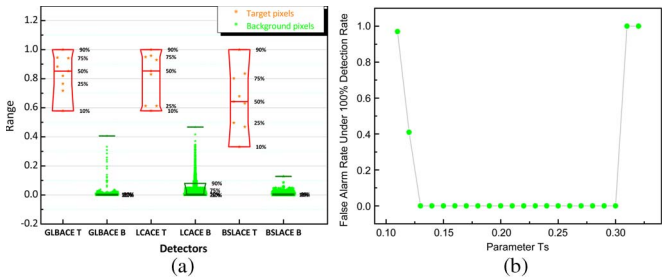

Fig. 4. (a) Target-background separation diagram for the RIT image. (b) Sensitivity analysis of parameter Ts.

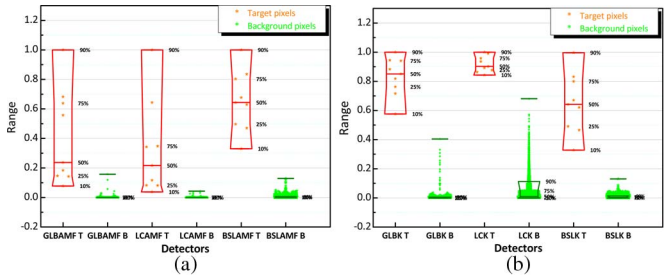

Fig. 5. Target-background separation diagram for the RIT image. Target-background separation diagram for the RIT image.

of monotonous background materials, which makes the global covariance matrix close to the local one and the proposed one. As the detection result maps were not visually significantly different, we calculated the distribution of the values of the targets and backgrounds in the detection results, respectively, and display them in the target-background separation diagram in Fig. 4(a).

In Fig. 4(a), the orange and green points are drawn according to the range of the target and background values in the detection results, respectively. The different columns suggest the value ranges of the targets and backgrounds, respectively, where T refers to the target, and B refers to the background on the horizontal axis. The lines at the top and bottom of each column are the extreme values. After statistical calculation of the data points, boxes are drawn to enclose the main parts of the points, excluding the largest 10% and the smallest 10%. The red boxes illustrate the distribution of the target points' values, and the line in the middle of the box is the mean of the points. In a similar way, the olive green boxes enclose the middle 80% of the main points of the point cloud in each background column. In other words, the position of the boxes reflects the tendency and compactness of the distribution of the point cloud. It is clear from Fig. 4(a) that the BSLACE detector can best suppress the background pixels to a small range, as the olive green box is the shortest. With this image, the BSL framework can achieve some improvement and can suppress the background pixels to a narrower range.

As the parameter Ts affects the detection performance, a series of detection experiments with different Ts were compared to decide a reliable candidate value for Ts. Fig. 4(b). was drawn by the false alarm rate under full detection rate with different Ts values. It is easily observed that the detection rate was controlled to 0 in a wide region, in which the parameter value was reliable. This is reasonable because a too small Ts will prevent the clustering process while a too large Ts will cluster the image into too few classes. The middle value of the region is believed to a reliable choice of the parameter Ts, so it is fixed at 0.2 throughout our experiments.

The same experimental set was conducted on the RIT image, with different ways of valuating the background statistics for the AMF and Kelly's GLRT detector. In LCAMF and LCK-GLRT, the size of the inner window and the outer window was set to 5 $\times$5 and 15 $\times$ 15, respectively. In BSLAMF and LCKGLRT, the window size was set as 5 $\times$ 5. Similarly, the target-background separation map is used to quantitatively assess the proposed algorithm shown in Figs. 5(a) and Fig. 5(b). From Fig. 5(a) it can be observed that only in BSLAMF can the target and background columns be separated. Both in GLBAMF and LCAMF, the value ranges of the target pixels and the background pixels have some obvious overlapping areas, which would make it difficult to determine a threshold. If all the target pixels are separated, many false alarms will present in GLBAMF and LCAMF, while in BSLAMF it is much easier to determine the threshold to separate the targets without any false alarms. In Fig. 5(b), the distribution of targets and backgrounds shows that with the GLBKGLRT, LCKGLRT, and BSLKGLRT detectors, targets can be detected by a threshold, without a false alarm. However, in BSLKGLRT, the range of the background pixels is compressed to the smallest range, which confirms the superiority of BSL in obtaining better statistics for suppressing the non-targets.

TABLE I
TIME COSTS (S) FOR THE DETECTORS IN DIFFERENT BACKGROUND ESTIMATION FRAMEWORKS

Algorithms	ACE	AMF	K'GLRT
GLB-	3.064808	3.389578	3.580115
LC-	22.753708	20.526966	21.035713
BSL-	14.664481	14.471008	15.365081

The computational cost factor is also considered, as shown in Table I. Although it takes a longer time, due to the clustering process, BSL still outperforms the local approaches in terms of computational cost.

B. AVIRIS Dataset

Detection with the San Diego data set is much more difficult as it consists of various complex background materials. GLBACE, LCACE, and BSLACE were applied to this image. The priori target was selected directly from the center part of the airplanes in the image. The ground truths of the airplane target points were carefully selected, as shown in Fig. 3(b), which are spatially continuous and spectrally similar to each other. Considering the size of the airplanes, the size of the inner window in LCACE was set to 15 $\times$ 15, the outer window to 23 $\times$ 23, and the window size in BSLACE was set to 15 $\times$ 15.

Fig. 6(a) shows the ROC curves generated by applying GLBACE, LCACE, and BSLACE to the AVIRIS image, based on the target reference. BSLACE performs best among the three methods, since it reaches a 100% detection rate at the lowest false alarm rate of below 1%. Therefore, the algorithm is shown to be effective in detecting the pixels with target signals in this data set. To further evaluate the ROC curves, the AUC

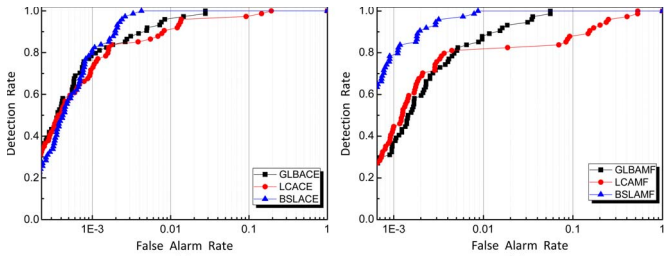

Fig. 6. (a) ROC curves of GLBACE, LCACE, and BSLACE for the San Diego image. (b) ROC curves of GLBAMF, LCAMF, and BSLAMF for the San Diego image.

TABLE II
AUC VALUES AND TIME COSTS (IN SECONDS) OF THE DETECTORS WITH DIFFERENT BACKGROUND ESTIMATION FRAMEWORKS

Algorithms	ACE	AUC	AMF	AUC	K'GLRT	AUC
GLB-	15.357489	0.9983	13.247820	0.9949	20.718537	0.9949
LC-	1409.457085	0.9939	1521.573171	0.9624	1617.815407	0.8002
BSL-	765.540878	0.9993	712.237980	0.9991	797.663796	0.9983

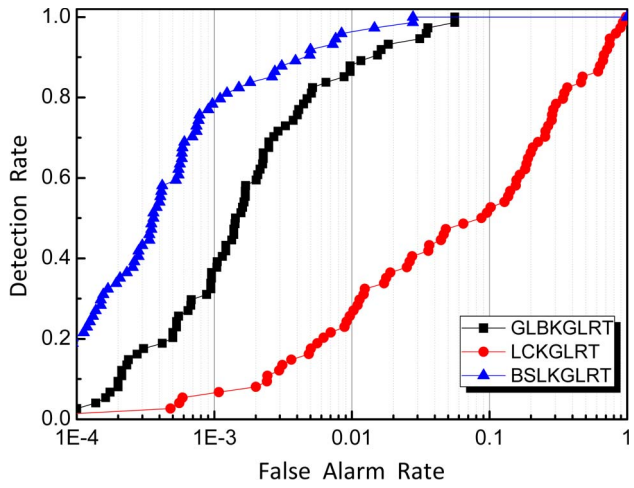

Fig. 7. The ROC curves of GLBKGLRT, LCKGLRT, and BSLKGLRT for the San Diego image.

was calculated from the integral of the area surrounded by the given curves and false alarm rate axis, as shown in Table II. A bigger AUC suggests that the curve lies to the upper-left in the figure, and consequently corresponds to a better detection performance.

GLBAMF, LCAMF, BSLAMF, GLBKGLRT, LCKGLRT, and BSLKGLRT were then applied to the San Diego image. As with the ACE series, the size of the inner window in LCAMF and GLBKGLRT was set to 15×15, the outer window to 23×23, and the window size in BSLACE and BSLKGLRT was set to 15×15. The ROC curves are drawn from the target reference distribution shown in Fig. 6(b) and Fig. 7. The quantitative detection results confirm the effectiveness of BSLAMF and BSLKGLRT. The AUCs of the three curves were calculated and are shown in Table II, where BSLAMF and BSLKGLRT again obtains the maximum value.

The computational cost factor is also well considered, as shown in Table II. As it is more time-consuming for the k-means clustering algorithm to converge on a bigger image, BSL takes a longer time than with the RIT image as it experiences the clustering process, but it still outperforms the local approaches in computational cost.

IV. CONCLUSION

This letter proposes a background self-learning (BSL) framework for unstructured background detectors, which is aimed at better modeling the test pixel's background and simultaneously taking computational costs into consideration. By applying global-based, local-based, and BSL-based unstructured detectors to real HSI images, extensive experiments were undertaken. From the detection results it can be concluded that: 1) the BSL-based detectors outperform the global-based detectors, with better ROC curves; and 2) the BSL-based detectors show obvious superiority over the local-based detectors, both in detection performance and efficiency.

REFERENCES

[1] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.
[2] M. S. Halper, "Global, local, stochastic background modeling for target detection in mixed pixels," in *Proc. Algorithms Technol. Multispec., Hyperspectral, Ultraspectral Imag. XVI*, 2010, pp. 769527-1–769527-13.
[3] B. Du and L. Zhang, "Random-selection-based anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1578–1589, May 2011.
[4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 2, pp. 115–127, Mar. 1986.
[5] L. L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Conf. Rec. 13th Asilomar Conf. Signals, Syst. Comput.*, 1996, vol. 1112, pp. 1114–1117.
[6] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.
[7] J. E. West, D. W. Messinger, and J. R. Schott, "Comparative evaluation of background characterization techniques for hyperspectral unstructured matched filter target detection," *J. Appl. Remote Sens.*, vol. 1, no. 1, p. 013520, Jan. 2007.
[8] S. Matteoli, N. Acito, M. Diani, and G. Corsini, "An automatic approach to adaptive local background estimation and suppression in hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 790–800, Feb. 2011.
[9] D. Tao, X. Li, X. Wu, and S. J. Maybank, "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, Oct. 2007.
[10] Y. Gu, C. Wang, D. You, Y. Zhang, S. Wang, and Y. Zhang, "Representative multiple kernel learning for classification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2852–2865, Jul. 2012.
[11] T. Wang, B. Du, and Z. Liangpei, "A kernel-based target-constrained interference-minimized filter for hyperspectral sub-pixel target detection," *IEEE J. Sel. Topics Appl. earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 626–637, Apr. 2013.
[12] D. Tao, X. Li, X. Wu, W. Hu, and S. J. Maybank, "Supervised tensor learning," *Knowl. Inf. Syst.*, vol. 13, no. 1, pp. 1–42, Sep. 2007.
[13] D. Tao, X. Li, X. Wu, and S. J. Maybank, "Geometric mean for subspace selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 260–274, Feb. 2009.
[14] J. Theiler and B. R. Foy, "Effect of signal contamination in matched-filter detection of the signal on a cluttered background," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 98–102, Jan. 2006.
[15] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
[16] J. P. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.