Hsueh, Ming kai 22045720

Research paper

Characterization of anomaly detection in hyperspectral imagery

Chein-I Chang and Mingkai Hsueh

Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, Maryland, USA

Abstract

Purpose – The paper aims to characterize anomaly detection in hyperspectral imagery.

Design/methodology/approach – This paper develops an adaptive causal anomaly detector (ACAD) to investigate several issues encountered in hyperspectral image analysis which have not been addressed in the past. It also designs extensive synthetic image-based computer simulations and real image experiments to substantiate the work proposed in this paper.

Findings – This paper developed an ACAD and custom-designed computer simulations and real image experiments to successfully address several issues in characterizing anomalies for detection, which are – first, how large size for a target to be considered as an anomaly? Second, how an anomaly responds to its proximity? Third, how sensitive for an anomaly to noise? Finally, how different anomalies to be detected? Additionally, it also demonstrated that the proposed ACAD can be implemented in real time processing and implementation.

Originality/value – This paper is the first work on investigation of several issues related to anomaly detection in hyperspectral imagery via extensive synthetic image-based computer simulations and real image experiments. In addition, it also develops a new developed an ACAD to address these issues and substantiate its performance.

Keywords Differential geometry, Correlation analysis, Image processing

Paper type Research paper

1. Introduction

Owing to very high spectral resolution provided by hyperspectral sensors, many subtle targets that cannot be resolved by multispectral imaging sensors can now be uncovered, such as special species in agriculture, unusual migrations in ecology, rare minerals in geology, toxic wastes in environments, drug trafficker or smugglers in law enforcement, vehicles/tanks in battlefields, cancerous cells or tumors in medical diagnosis, etc. just name a few. Such targets generally appear as anomalies in a form of abnormalities that are distinct from their surroundings. Unfortunately, the issue of how to characterize anomalies in terms of existence, presence and populations has not been addressed. The reason of being anomalies is because they are not known a priori. On the other hand, anomalies usually occur with low probabilities. Therefore, their existence generally cannot be detected by any supervised means or visual inspection. As for presence, the spatial extent of anomalies is rather limited since they can be present as subpixel targets with their size smaller than pixel size or as pixels mixed with the background or other substances. Most importantly, once anomalies do appear, their population is relatively small due to the nature of being anomalies. Interestingly, several issues arising in anomalies have not been investigated or explored in the past. First of all, how large is a target to be considered as an anomaly in terms of size? Second, how does an anomaly respond to its surrounding neighborhood? Third, how sensitive is an anomaly to noise? Fourth, how can an anomaly distinguish itself from other anomalies? Finally, how can these anomalies be detected effectively by taking into account all the above-mentioned issues? Many algorithms have been developed for anomaly detection such as the well-known RX filter (RXF) developed by Reed and Yu (1990). However, the above-mentioned issues still remain unsolved. This paper develops an approach, called adaptive causal anomaly detector (ACAD) to resolve these issues. It can perform causal anomaly detection in real time as does the causal RXF (CRXF) (Chang and Chiang, 2002; Chang, 2003) while adaptively building up a library for anomalies it detected. Several advantages are derived from the proposed ACAD. It resolves the issue caused by RXF or CRXF that an early detected anomaly may impair detectability of follow-up anomalies. Second, it can use the built library to map out the detected anomalies in the original image. This map is generated simultaneously as the detection process moves on. Third, it provides an effective means to investigate various issues related to sensitivity of anomaly detection for performance evaluation and analysis.

2. Anomaly detection

In this section, we describe three types of anomaly detectors, a filter implemented by RXF (Reed and Yu, 1990), a CRXF

The current issue and full text archive of this journal is available at www.emeraldinsight.com/0260-2288.htm

Sensor Review
26/2 (2006) 137–146
 [ISSN 0260-2288]
[DOI 10.1108/02602280610652730]

(Chang and Chiang, 2002; Chang, 2003) and dual window-based eigen separation transform (DWEST) (Kwon *et al.*, 2003). It should be noted that since a hyperspectral image is an image cube, its image pixels are actually a column vector. However, "pixel" instead of "pixel vector" is still used in this paper for simplicity.

2.1 RX filter

A widely used anomaly detector was developed by Reed and Yu (1990), referred to as RXF, which is a constant false alarm rate (CFAR) detector and derived from the generalized likelihood ratio test. It is specified by:

$$\delta^{\mathrm{RXF}}(\mathbf{r}) = (\mathbf{r}-\mu)^{\mathrm{T}}\mathbf{K}^{-1}(\mathbf{r}-\mu) \tag{1}$$

where μ is the sample mean and $\mathbf{K}$ is the sample data covariance matrix. It should be noted that the form of $\delta^{\mathrm{RXF}}(\mathbf{r})$ in equation (1) actually performs the well-known Mahalanobis distance.

2.2 Causal RX filter

Since RXF involves the computation of the mean and covariance matrix in equation (1), RXF cannot be implemented in real time. In order to mitigate this problem, we replace $\mathbf{K}$ in equation (1) with the sample data correlation matrix to construct a real-time anomaly detector, called CRXF, $\delta^{\mathrm{CRXF}}(\mathbf{r})$ by:

$$\delta^{\mathrm{CRXF}}(\mathbf{r}_k) = \mathbf{r}_k^{\mathrm{T}}(\mathbf{R}^{-1}(\mathbf{r}_k))\mathbf{r}_k \tag{2}$$

where:

$$\mathbf{R}(\mathbf{r}_k) = (1/k)\sum_{i-1}^{k}\mathbf{r}_i\mathbf{r}_i^{\mathrm{T}}$$

is called causal sample correlation matrix at $\mathbf{r}_k$ and $\mathbf{R}(\mathbf{r}_N) = \mathbf{R}$. It should be noted that the sample data correlation matrix $\mathbf{R}(\mathbf{r}_k)$ used in CRXF is the sample correlation matrix formed by the sample vectors $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k\}$ up to the pixel to be processed, $\mathbf{r}_k$.

It is worth noting that we can make the causal sample correlation matrix independent of the pixel to be processed by replacing $\mathbf{R}(\mathbf{r}_k)$ with the sample auto-correlation matrix $\mathbf{R}$. The resulting anomaly detector is called correlation-based RXF (R-RXF), $\delta^{\mathrm{R-RXF}}(\mathbf{r}_k)$ given by:

$$\delta^{\mathrm{R-RXF}}(\mathbf{r}_k) = \mathbf{r}_k^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}_k \tag{3}$$

Comparing equations (3) to (2), the only circumstance that the $\delta^{\mathrm{R-RXF}}(\mathbf{r}_k)$ and $\delta^{\mathrm{CRXF}}(\mathbf{r}_k)$ perform identically is the case when $\delta^{\mathrm{CRXF}}(\mathbf{r}_k)$ reaches the last pixel $\mathbf{r}_N$ in the image, i.e. $k = N$.

2.3 DWEST

A recently developed adaptive anomaly detector, referred to as DWEST in Kwon *et al.* (2003) implements two windows, called inner and outer windows which are designed to maximize the separation between two-class data, target class and background class. Let $\mu_{\mathrm{outer}}(\mathbf{r})$ and $\mu_{\mathrm{inner}}(\mathbf{r})$ be the means of the outer and inner windows, respectively , and $\mathbf{K}_{\mathrm{outer}}$ and $\mathbf{K}_{\mathrm{inner}}$ be their respective covariance matrices. Define $\mathbf{K}_{\mathrm{diff}} = \mathbf{K}_{\mathrm{inner}} - \mathbf{K}_{\mathrm{outer}}$ as the difference covariance matrix between $\mathbf{K}_{\mathrm{outer}}$ and $\mathbf{K}_{\mathrm{inner}}$. As a result, the eigenvalues of $\mathbf{K}_{\mathrm{diff}}$ can be divided into two groups, negative values and positive values. Kwon *et al.* argued that the eigenvectors associated with a small number of the large positive eigenvalues of $\mathbf{K}_{\mathrm{diff}}$ could successfully extract the spectrally distinctive materials that are present in the inner window. If the eigenvectors represented by the positive eigenvalues in this small set are denoted by $\{\mathbf{v}_i\}$, the anomaly detector derived by DWEST, $\delta^{\mathrm{DWEST}}(\mathbf{r})$, projects the differential means of two windows, $\mu_{\mathrm{diff}}(\mathbf{r}) = \mu_{\mathrm{outer}}(\mathbf{r}) - \mu_{\mathrm{inner}}(\mathbf{r})$ onto $\{\mathbf{v}_i\}$ specified by:

$$\delta^{\mathrm{DWEST}}(\mathbf{r}) = \left|\sum_{\mathbf{v}_i}\mathbf{v}_i^{\mathrm{T}}\mu_{\mathrm{diff}}(\mathbf{r})\right| \tag{4}$$

Two remarks are noteworthy:

1. The images generated by RXF and CRXF are gray scale. Image analysts must interpret themselves to determine if a pixel is anomalous pixel according to its intensity. However, such human interpretation is generally very subjective. In order for interpretation to be objective, a threshold must be implemented to threshold a gray scale image to a binary image which segments anomalies from the image background. In this case, how to choose the threshold is crucial. This issue will be discussed in our proposed ACAD.
2. One key parameter to determine the anomaly detection performance is the ratio of image size to anomaly size. This fact will be demonstrated by the experiments later.

3. Adaptive causal anomaly detection

One major issue encountered in RXF, R-RXF or CRXF is that if an earlier detected anomaly has a strong signature, it may have significant impact on the detection of later detected anomalies. This phenomenon is mainly caused by an inappropriate use of the sample correlation matrix. According to Chang (2005), an appropriate sample correlation matrix should be one that removes all the earlier detected target pixels being included in the sample correlation matrix. Therefore, we should replace the $\mathbf{R}(\mathbf{r}_k)$ in equation (2) with a sample spectral correlation matrix that removes all detected anomalies, referred to as causal anomaly-removed sample spectral correlation matrix, $\tilde{\mathbf{R}}(\mathbf{r}_k)$ defined by:

$$\tilde{\mathbf{R}}(\mathbf{r}_k) = \mathbf{R}(\mathbf{r}_k) - \sum_{\mathbf{t}_j\in\Delta(k)}\mathbf{t}_j\mathbf{t}_j^{\mathrm{T}} \tag{5}$$

where $\Delta(k)$ is the set of all earlier detected anomalous target pixels $\mathbf{t}_j$ prior to the currently being processed image pixel $\mathbf{r}_k$.

Another major issue arising in RXF, R-RXF and CRXF is that the size of anomalies to be detected cannot be too large. This actually fits the nature of anomaly. However, the dilemma is how large the size can be for a target to be considered as an anomaly? Furthermore, a third issue is how close is too close for two anomalies to be detected as two separate anomalies? Finally, a fourth issue is how to distinguish two detected anomalies one from another?

In order to address these issues, we develop a new adaptive anomaly detector, called ACAD, $\delta^{\mathrm{ACAD}}(\mathbf{r}_k)$ via the causal anomaly-removed sample spectral correlation matrix, $\tilde{\mathbf{R}}(\mathbf{r}_k)$ defined in equation (5). It is also a matched filter but implements the following causal filter form which replace $\mathbf{R}(\mathbf{r}_k)$ in CRXF in equation (2) with $\tilde{\mathbf{R}}(\mathbf{r}_k)$:

$$\delta^{\mathrm{ACAD}}(\mathbf{r}_k) = \mathbf{r}_k^{\mathrm{T}}\tilde{\mathbf{R}}^{-1}(\mathbf{r}_k)\mathbf{r}_k \tag{6}$$

Despite the fact that ACAD is a result of a simple replacement of $\mathbf{R}(\mathbf{r}_k)$ with $\tilde{\mathbf{R}}(\mathbf{r}_k)$ in CRXF, ACAD has

Sensor Review
Volume 26 · Number 2 · 2006 · 137–146

several advantages over CRXF and RXF. One is that while performing anomaly detection, ACAD also simultaneously builds and updates an anomaly library as well as generates an anomaly map to provide spatial coordinates of all its detected anomalies in the original image. This anomaly map can be also used to classify all the detected anomalies and is similar to the one, target discrimination map proposed in Chang and Chiang (2002) and Chang (2003). In analogy with CRXF, ACAD can be also implemented in real time except that ACAD continues to remove an anomaly from its subsequent process, once a target pixel was detected as an anomaly. Here the term of "real time" implies that the process is carried out by ACAD in a causal manner so that only the pixels which were already processed and the currently being processed pixel are used for anomaly detection. As a result, ACAD can be still considered as a near real-time processor since the process of a data sample vector is completed nearly at the same time it comes in with probably a short time delay due to data processing. A tricky issue of implementing ACAD is how to determine if a pixel is an anomaly.

Since ACAD is a causal filter, the pixels to be considered as surrounding pixels must be those neighboring pixels which have been already processed. Let n_{ACAD} be the predetermined number of pixels preceding the currently being processed pixel. Then, n_{ACAD} indicates how far back the pixels in the past must be taken into account to find an appropriate threshold value to determine if the currently being processed pixel $\mathbf{r}_k$ is an anomalous pixel. In doing so, we first calculate the average of the abundance fractions of the past pixels, $\mathbf{r}_{k-1}, \mathbf{r}_{k-2}, \cdots, \mathbf{r}_{k-n_{\mathrm{ACAD}}}$ generated by $\delta^{\mathrm{ACAD}}(\mathbf{r}_k)$, denoted by:

$$\mu_k = \frac{1}{n_{\mathrm{ACAD}}} \sum_{i-1}^{n_{\mathrm{ACAD}}} \delta^{\mathrm{ACAD}}(\mathbf{r}_{k-i})$$

as well as the difference between μ_k and $\delta^{\mathrm{ACAD}}(\mathbf{r}_k)$, denoted by $t_k = \delta^{\mathrm{ACAD}}(\mathbf{r}_k) - \mu_k$. If the t_k is smaller than a predetermined threshold value τ, then the $\mathbf{r}_k$ will not be considered as an anomaly, in which case it will be retained in subsequent data processing. Otherwise, the $\mathbf{r}_k$ will be declared as an anomaly. In this case, this pixel will be removed from future data processing and added to the built anomaly library. In the mean time, the spatial coordinate of this pixel is stored in the anomaly map.

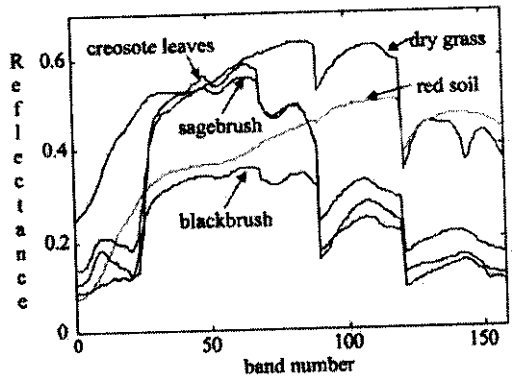
4. Issues arising anomaly detection

In this section, we conduct a comprehensive study of computer simulations to investigate four issues related to anomaly detection, which are:

1 How large for a target to be considered as an anomaly?
2 How an anomaly responding to its neighboring pixels?
3 How sensitive for an anomaly to noise?
4 How different anomalies to be detected?

The data set used for our computer simulations to investigate these four issues is the airborne visible/infrared imaging spectrometer (AVIRIS) reflectance data (Harsanyi and Chang, 1994) which contains five reflectance spectra, blackbrush, creosote leaves, dry grass, red soil and sagebrush shown in Figure 1 with spectral coverage from 0.4 to 2.5 μm. There are 158 bands after water bands are removed.

Figure 1 Five AVIRIS reflectance signatures



4.1 How large the size for a target to be considered as an anomaly?

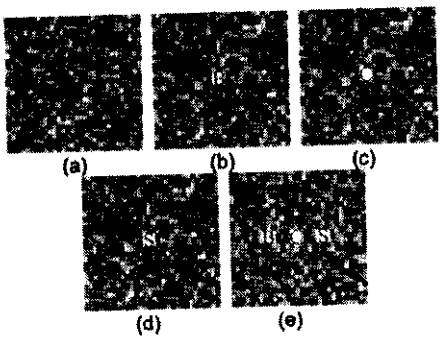
The first issue to be investigated is "how large size can an anomaly be?". Obviously, a target that can be identified visually should not be considered as an anomaly. With this interpretation, the size of an anomaly must be relatively small. The issue is how large the size is for a pixel to be considered as an anomaly. Interestingly, it has to do with the ratio of the entire image size to the size of an anomaly, β rather than the size of an anomaly alone. According to our extensive experiments, the β can be empirically shown to be approximately 100, in which case, the size of anomaly can be determined by:

$$n_{\mathrm{ACAD}} = \left[\frac{N}{100}\right]$$

where N is the total number of pixels of the image and $[x]$ is the largest integer $\leq x$.

To make our analysis simple, we simulated a synthetic background image shown in Figure 2(a), which has size of 30 × 30 pixels made up by 50 percent grass and 50 percent red soil plus a Gaussian noise with signal-to-noise ratio (SNR) 30:1 as defined in Harsanyi and Chang (1994). Three panels of 3 × 3 pixels that were simulated by 100 percent blackbrush, 100 percent creosote leaves and 100 percent sagebrush, respectively , were used for experiments.

Figure 2 (a) A synthetic image of 30 × 30 pixels; (b) 3 × 3 blackbrush at the center; (c) 3 × 3 creosote leaves at the center; (d) 3 × 3 sagebrush at the center; and (e) three 3 × 3 panels of blackbrush, creosote leaves and sagebrush implanted at the center of (a)



These three 3 × 3 panels were implanted at the center of the synthetic image in Figure 2(a) separately in Figure 2(b)-(d) and all together at the center in Figure 2(a) with three pixels apart in Figure 2(e). It should be noted that the panel pixels were implanted in a way that they were used to replace their corresponding background pixels.

Figures 3-6 are anomaly detection results of applying RXF, R-RXF, CRXF, DWEST with inner/outer windows specified by 5/11 (i.e. inner and outer windows were specified by the size of 5 × 5 and 11 × 11 pixels, respectively) and ACAD to Figure 2(b)-(d), respectively , with the corresponding ACAD-generated anomaly maps shown in Figure 7. It should be noted that the dual windows 5/11 for DWEST were selected based our experiments that gave the best possible performance.

As we can see, RXF and R-RXF performed very similarly in terms of visual detection. Interestingly, when CRXF was implemented, the first panel pixel was clearly detected. Afterwards, the amounts of detected subsequent panel pixels were decreased significantly and vanished eventually. This cause was due to the real-time processing of the causal sample correlation matrix where the first detected target dominated subsequent detected targets. However, this problem was remedied by ACAD shown in figures labeled by (e) in Figures 3-6 where all the panel pixels were detected with substantial abundance fractions. It was also found that DWEST did not perform as well as it was claimed (Kwon *et al.*, 2003) where the blurring effects around three separate 3 × 3 panels in Figure 6(d) were caused by the use of dual windows. Comparing the results in Figures 3-6, it is clearly demonstrated that ACAD significantly improved all the other four anomaly detectors, RXF, R-RXF, CRXF and DWEST.

Next we increased the size of the image in Figure 2(a) from 30 × 30 to 35 × 35 pixels while retaining the same size for the three separate 3 × 3 panels in Figure 2(e) as shown in Figure 8(a).

Compared to Figures 6(a)-(e) to 8(b)-(f), there were no visible differences. Interestingly, this was no longer true if the size of the image in Figure 2(e) was decreased to 25 × 25 pixels shown in Figure 9(a) with the same three separate 3 × 3 panels in Figure 2(e) implanted in the image background. Figure 9(b)-(g) shows anomaly detection results of applying RXF, R-RXF, CRXF, DWEST with 5/11

Figure 3 Anomaly detection of RXF, R-RXF, CRXF, DWEST and ACAD in Figure 2(b)

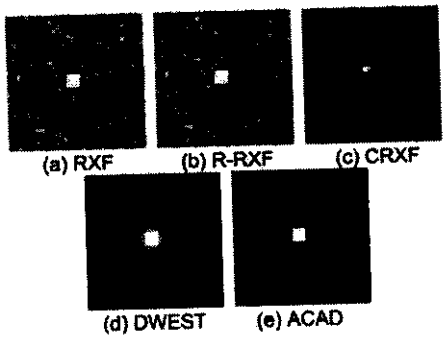


Figure 4 Anomaly detection of RXF, R-RXF, CRXF, DWEST and ACAD in Figure 2(c)

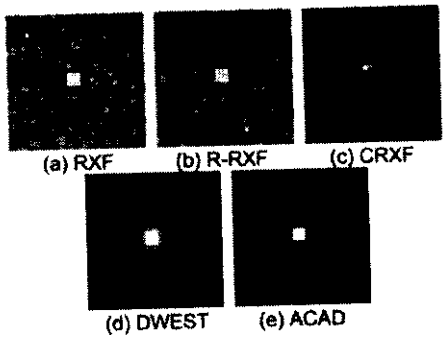


Figure 5 Anomaly detection of RXF, R-RXF, CRXF, DWEST and ACAD in Figure 2(d)

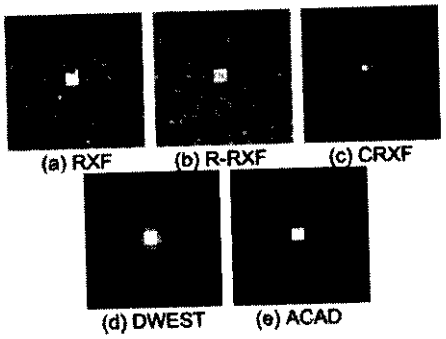


Figure 6 Anomaly detection of RXF, R-RXF, CRXF, DWEST and ACAD in Figure 2(e)

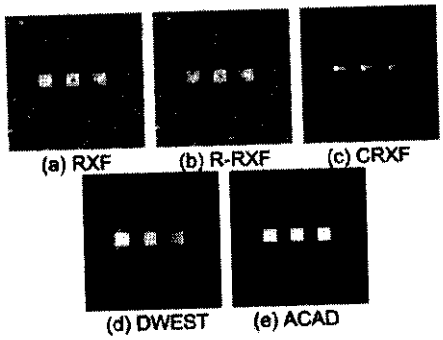


Figure 7 Anomaly map generated by ACAD for Figure 2(b)-(e)

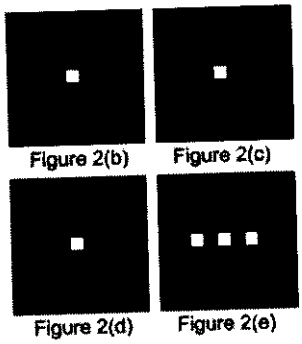
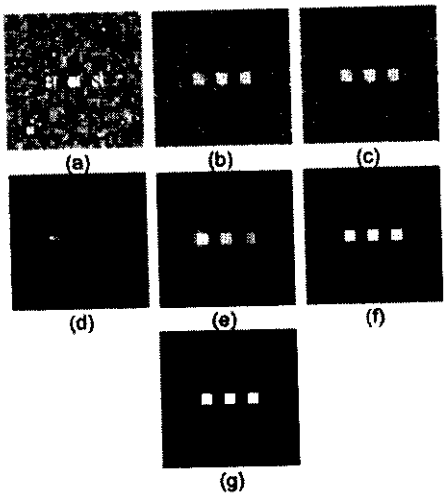


Figure 8 (a) A synthetic image of 35 × 35 pixels with three separate 3 × 3 panels at the center; (b) RXF; (c) R-RXF; (d) CRXF; (e) DWEST with 5/11; (f) ACAD; and (g) ACAD-generated anomaly map



dual windows and ACAD to Figure 8(a) with ACAD generated anomaly map in Figure 9(g). As shown in Figure 9(b)-(d) RXF, R-RXF and CRXF had difficulty with detecting all the three separate 3 × 3 panel pixels, while ACAD did not have any problem at all. As for DWEST, the result remained largely unchanged compared to those in Figures 6(d) and 8(e).

On the other hand, if the image size was fixed at 30 × 30, but the size of the three separate 3 × 3 panels was increased from 3 × 3 to 5 × 5 as shown in Figure 10(a). Figure 10(b)-(g) shows anomaly detection results of applying RXF, R-RXF, CRXF, DWEST with 5/11 dual windows and ACAD to Figure 10(a) with ACAD anomaly map in Figure 10(g).

Once again, Figure 10(f) and (g) shows that ACAD detected all 5 × 5 panel pixels very effectively, but both RXF

Figure 9 (a) A synthetic image of 25 × 25 pixels with three separate 3 × 3 panels at the center; (b) RXF; (c) R-RXF; (d) CRXF; (e) 5/11 DWEST; (f) ACAD; and (g) ACAD-generated anomaly map

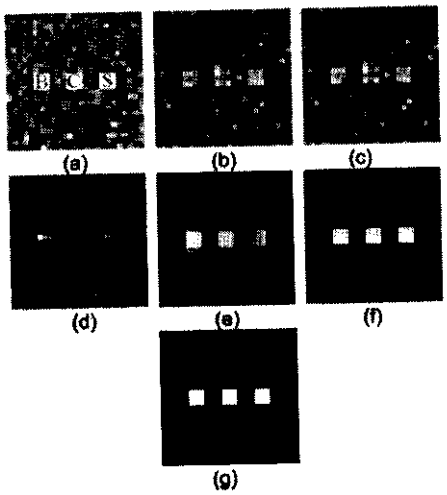
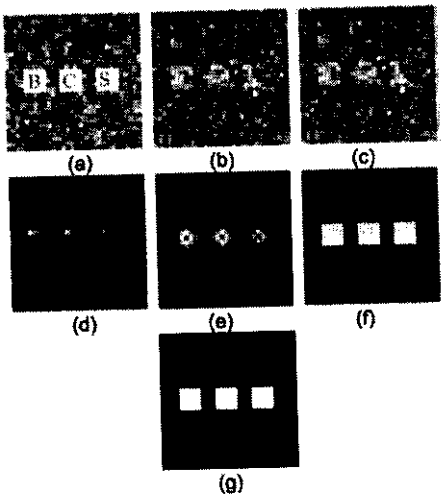


Figure 10 (a) A synthetic image of 30 × 30 pixels with three separate 5 × 5 panels at the center; (b) RXF; (c) R-RXF; (d) CRXF; (e) 5/11 DWEST; (f) ACAD; and (g) ACAD-generated anomaly map



and R-RXF in Figure 10(b) and (c) had even greater difficulty with detecting these three separate panel pixels. On the other hand, CRXF in Figure 10(d) did improve performance slightly compared to the result in Figure 9(d). Interestingly, DWEST performed worse than it did in Figure 9(e) where it did not detect the center pixels in each of three 5 × 5 panels due to an inappropriate use of the dual window 5/11. The above four experiments simply demonstrated that the effectiveness of RXF, R-RXF, CRXF and DWEST was closely related to the size of anomalies relative to the size of the entire image. Additionally, DWEST was also very sensitive to the size of dual windows. However, ACAD was very robust throughout all the experiments and was the best among all the evaluated anomaly detectors. According to our experimental study, RXF, R-RXF and CRXF can be only effective if the ratio of the image size to the size of an anomaly to greater or equal to a certain number, such as 100:1 as demonstrated in Figures 3-10, but may fail if the ratio lower than 100:1 as shown in Figures 9 and 10 with the image size-to-anomaly size ratios, 625:1 ≅ 64.44:1 and 36:1, respectively.

It is worth noting that the performance of ACAD is determined by the parameter, n_{ACAD}. In the above experiments, the n_{ACAD} was chosen to be:

$$n_{\mathrm{ACAD}} = \left\lceil \frac{N}{100} \right\rceil$$

which can be determined, but not necessarily by the ratio of image size to anomaly size, β set to 100:1. Interestingly, once it was chosen properly, ACAD performs very robustly regardless of the ratio of image size to anomaly size, β was below 100:1 as demonstrated in Figures 9 and 10. This implies that the performance of ACAD is not determined by the relative size of the entire image to the anomaly, but rather by the number of pixels considered in n_{ACAD}. In this paper, the n_{ACAD} was set to:

$$n_{\mathrm{ACAD}} = \left\lceil \frac{N}{100} \right\rceil$$

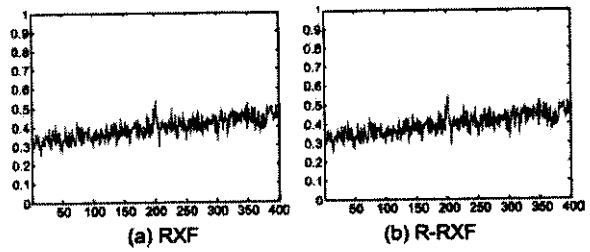
4.2 How does an anomaly respond to its surroundings?
This section investigates the issue of how an anomaly responds to its surrounding pixels. To simplify simulations, only three signatures in Figure 1, creosote leaves, dry grass and red soil were used to simulate 401 mixed pixels, $\{\mathbf{r}_{i-1}\}_{i-1}^{401}$ as follows. We start the first pixel with 100 percent red soil and 0 percent dry grass, then began to increase 0.25 percent dry grass and decrease 0.25 percent red soil every pixel until the 401st pixel which contained 100 percent dry grass. We then added creosote leaves, denoted by t_0, to pixel numbers 198-202 at abundance fractions 10 percent while reducing the abundance of red soil and dry grass evenly. For example, after addition of creosote leaves, the resulting pixel 200 contained 10 percent creosote leaves, 45 percent red soil and 45 percent dry grass. White Gaussian noise was also added to each pixel to achieve a 30:1 SNR. Figure 11(a) and (b) shows anomaly detection of five creosote leaves pixels $\{\mathbf{r}_{i-1}\}_{i-198}^{202}$ by $\delta^{\mathrm{RXF}}(\mathbf{r})$ and $\delta^{\mathrm{R-RXF}}(\mathbf{r})$ where both $\delta^{\mathrm{RXF}}(\mathbf{r})$ and $\delta^{\mathrm{R-RXF}}(\mathbf{r})$ performed very similarly in detection of creosote leaves.

Since both CRXF and ACAD were implemented causally, Figure 12(a) and (b) shows progressive detection results of CRXF and ACAD as pixel numbers starting from 197 to 203 with $\mathbf{R}(\mathbf{r}_k)$ and $\hat{\mathbf{R}}(\mathbf{r}_k)$ defined in equations (2) and (5), respectively.

As shown in Figure 12(a), CRXF did not detect the creosote leaves until it reached the pixel 198 which is the first creosote leaves pixel and missed the remaining four creosote leaves pixels from 199 to 202 afterwards. This makes sense because after the first creosote leaves pixel was detected at 198, the spectral signature of the creosote leaves was not considered as an anomalous signature any more due to the fact that the following four pixels appeared to have the same signature. Interestingly, as shown in Figure 12(b), ACAD detected all the five creosote leaves pixels $\{\mathbf{r}_{i-1}\}_{i-198}^{202}$ once a creosote leaves pixel was detected and removed subsequently. This experiment demonstrated the significant impact of detected anomalies on subsequent anomaly detection. It should be noted that as the processed pixel $\mathbf{r}$ reached the last pixel, i.e. $\mathbf{r} = 401$, $\mathbf{R}(401) = \boldsymbol{R}$. In this case, $\delta^{\mathrm{CRXF}}(401)$ is the same as $\delta^{\mathrm{R-RXF}}(401)$. It should be noted that since no spatial correlation exists among simulated pixels, DWEST was not implemented.

4.3 How sensitive an anomaly to noise?
This section investigates the noise sensitivity to anomaly detection. In doing so, we used the same image in Figure 2(e) with additive Gaussian noise of SNR = 20:1, 10:1 and 5:1, respectively. Figure 13 shows the results of RXF, R-RXF, CRXF, DWEST with 5/11 dual windows and ACAD obtained with SNR = 20:1, 10:1 and 5:1, respectively.

Figure 11 Detection results of RXF and R-RXF



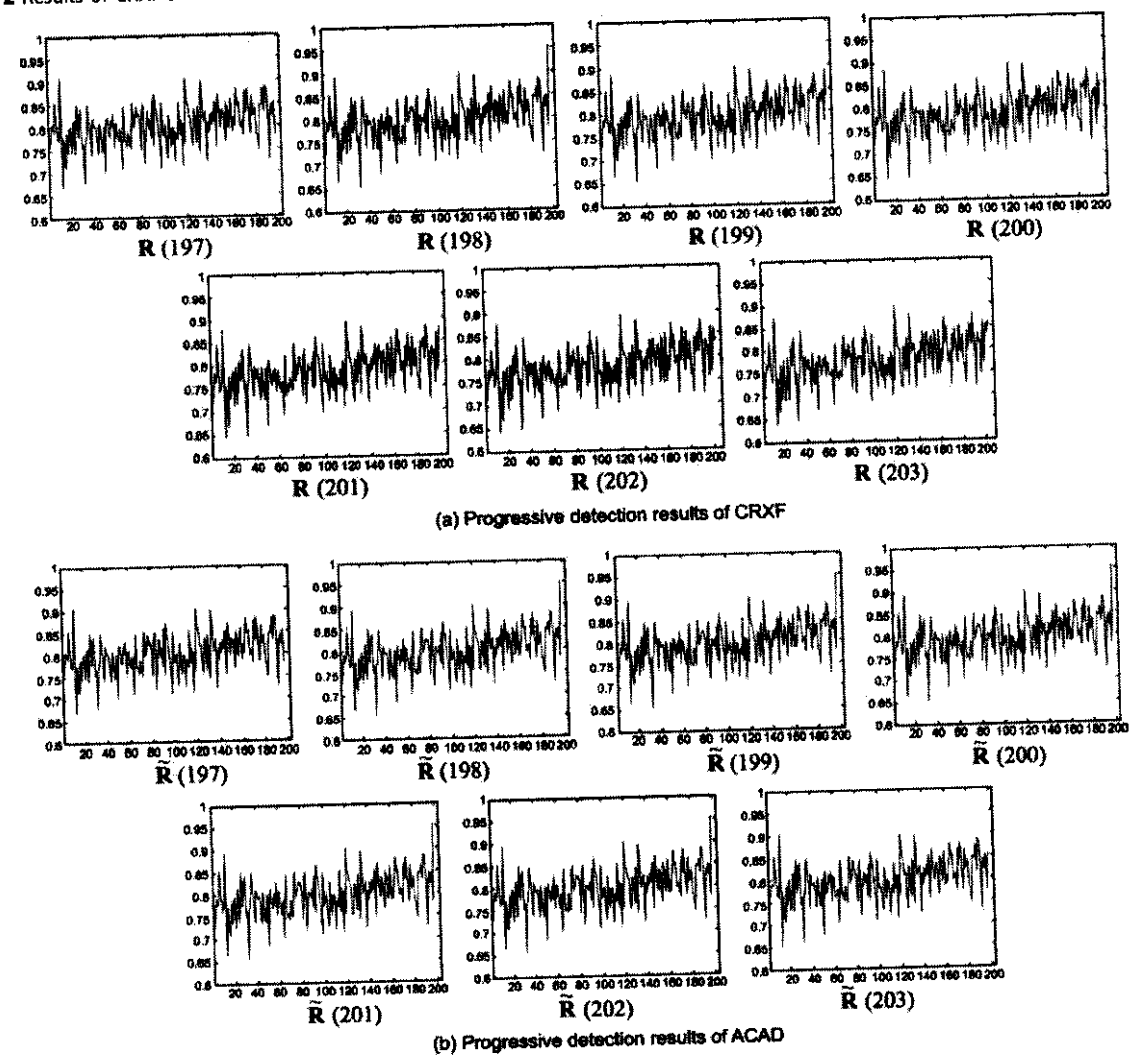
As shown, the noise also had significant impact on performance of RXF, but relatively less effect on CRXF, but had less effect on ACAD and almost no effect on DWEST. This experiment concluded that noise did not have as much effect as the ratio of the image size to anomaly size did to RXF, R-RXF and CRXF.

4.4 How different anomalies to be detected?
This section demonstrates that the detectability of anomalies does not depend on whether the anomalous pixels are pure or mixed; rather their actual spectral signatures. In other words, the spectral signature of an anomaly is immaterial as long as it has a spectrally distinct spectral signature and "sufficiently" different from its surrounding background.

Two scenarios were conducted to address this issue. One scenario was designed to see if the anomaly detectors are sensitive to pure pixels. In this scenario, we replaced the three panels made up of pure pixels in Figure 2(e) with three mixed-pixel panels made by 50 percent blackbrush + 50 percent creosote leaves, 50 percent blackbrush + 50 percent sagebrush, 50 percent creosote leaves + 50 percent sagebrush as shown in Figure 14(a)-(f) are the anomaly detection results of applying RXF, R-RXF, CRXF, DWEST and ACAD to Figure 14(a) with ACAD generated anomaly map in Figure 14(g). The results obtained in Figure 14(b)-(f) were very comparable to those in Figure 9. This implied that anomaly detection has nothing to do with if an anomaly is pure or mixed.

The second scenario was designed to see how distinct anomalies have impact on RXF, R-RXF, CRXF, DWEST and ACAD. The experiments conducted previously demonstrated that RXF, R-RXF, CRXF, DWEST and ACAD could detect anomalies regardless of whether they are of the same type or distinct types, and the effectiveness of their anomaly detection ability is closely related to the spectral signatures of pixels surrounding an anomaly. This is particularly evident for the case of CRXF where once the first target pixel was detected, subsequent target pixels with the same type of the signature could be not detected unless the previously detected target pixels were removed as the way performed by ACAD. In this scenario, we created four sets of panels. The first panel is a 1 × 3 panel of three pixels concatenated together and all were made by the same blackbrush signature. This first three-pixel panel was then implanted in Figure 2(a) at the center. The second panel is also a 1 × 3 panel of three pixels concatenated together. Unlike the first three-pixel panel made by a single pure signature, the three pixels in the second three-pixel panel were made by three different pure signatures, blackbrush, creosote leaves, sagebrush, respectively. The second panel was then implanted five-pixel right below the first three-pixel panel as a second row in Figure 2(a). Similarly, the third panel was a 1 × 6 panel made up of six pixels and was also implanted five-pixel below the second panel as a third row in Figure 2(a) where the first two pixels in the six-pixel panel are made by blackbrush concatenated by next two pixels made by creosote leaves and the last two concatenated pixels made by sagebrush. The fourth panel is a 2 × 6 panel consisting of 12 pixels implanted five-pixel below the third panel as a fourth row in Figure 2(a) where the first 2 × 2 panel in the 2 × 6 panel was made by blackbrush concatenated by another 2 × 2 panel made by creosote leaves, then concatenated by a third 2 × 2 panel made by sagebrush. These four sets of panels, a 1 × 3 panel

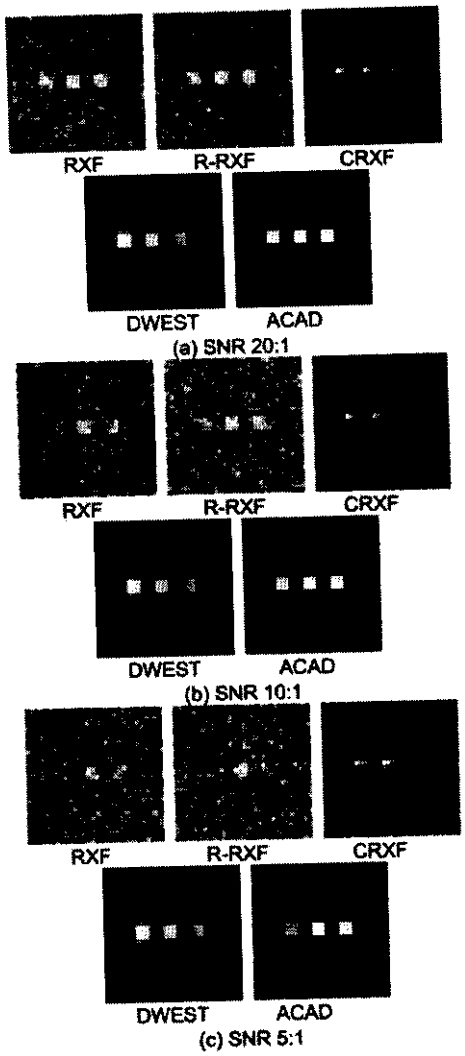
Figure 12 Results of CRXF and ACAD



of three pixel with the same spectral signature, 1 × 3 panel of three pixels with three distinct signatures, a 1 × 6 panel of six pixels with three distinct signatures and the 2 × 6 panel of twelve pixels with three distinct signatures were arranged in the first, second, third and fourth rows as shown in Figure 15(a).

Figure 15(b)-(f) shows the results of RXF, R-RXF, CRXF, DWEST and ACAD. As we can see from Figure 15, ACAD was still the best among all the evaluated detectors. Interestingly, DWEST was the worst detector. Once again this was also due to its sensitivity to its used dual windows. While RXF and R-RXF performed relatively the same and better than CRXF, an observation on CRXF in Figure 15(d) is worthwhile. It should be noted that CRXF was carried out causally in real time. Once the blackbrush was detected in the first pixel of the first 1 × 3 panel in the first row, the blackbrush was then largely discarded afterwards. This was witnessed in detection of subsequent panel detection. Most noticeably was anomaly detection of the second panel in the second row which was made up in order by three distinct signatures, blackbrush, creosote leaves, sagebrush. CRXF discarded the first pixel in the second panel which was the blackbrush and went ahead to detected the second signature, creosote leaves instead as shown by the brightest pixel in the second panel. The ability of CRXF in anomaly detection of the third and fourth rows was reduced due to effects caused by the detected anomalies in previous rows. This interesting experiment demonstrated that CRXF performed differently from RXF and R-RXF when anomalies were distinct. This may be an advantage of CRXF over RXF and R-RXF. However, whether or not anomalies are distinct seems to have no impact at all on ACAD. The only effect resulting from a sequence processing of anomaly detection on ACAD was that the gray scale values of earlier detected anomalies were suppressed by the subsequent detected anomalies. This was mainly due to causality required for real-time processing, which was also true for CRXF and will also be demonstrated in the following sections of real image experiments with real time implementation. Nevertheless, such visual assessment does not have impact on anomaly detection because ACAD

Figure 13 Results of RXF, R-RXF, CRXF, DWEST and ACAD with SNR = 20:1, 10:1 and 5:1



can build and update its anomaly library once a new anomaly was detected while the detection process is taking place. A similar process was first developed for CRXF in Chang and Chiang (2002) and Chang (2003).

Finally, a concluding remark is worthwhile. In order to address the issues of interest, the simulations presented here were custom-designed and relatively simple to simulate various scenarios. They can be used as a base to simulate more sophisticated synthetic images for further exploration.

5. Real hyperspectral image experiments

In this section, we conducted real hyperspectral image experiments to substantiate the utility of ACAD in real hyperspectral data exploitation. The image scene to be used for our experiments is shown in Figure 16(a) (band 80) which was considered in Kwon *et al.* (2003). It is a 210-band HYDICE image scene of size 90 × 33 pixels with five vehicles parked vertically in a large grass field along the tree line. The

Figure 14 (a) A synthetic image with mixed pixel targets; (b) RXF; (c) R-RXF; (d) CRXF; (e) DWEST; (f) ACAD; and (g) ACAD-generated anomaly map

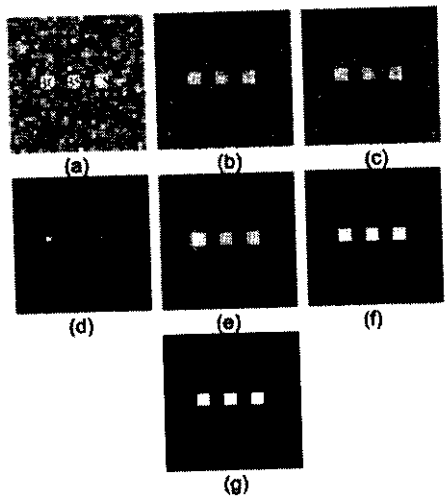
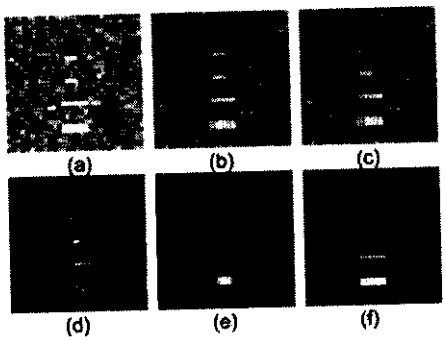


Figure 15 (a) Two 1 × 3 panels, a 1 × 6 panel and the 2 × 6 panel implanted in Figure 2(a); (b) RXF; (c) R-RXF; (d) CRXF; (e) DWEST; and (f) ACAD



ground truth map of the five vehicles is provided in Figure 16(b) where the pixels marked by red are the center pixels of the vehicles and the pixels marked by yellow are considered as vehicle pixels mixed with background pixels. The spectral resolution and spatial resolution are 10 nm and 1.56 m, respectively.

Figure 17(a)-(e) shows anomaly detection results of RXF, R-RXF, CRXF, DWEST with 7/13 dual windows used in Kwon *et al.* (2003) and ACAD, respectively , along with ACAD-generated anomaly map shown in Figure 17(f).

As mentioned previously, due to causality processed by CRXF and ACAD in real time the gray scale values of the top four vehicles were suppressed by the gray scale values of the bottom vehicle in Figure 17(c) and (e). As a result, the bottom vehicle was visually brighter than the other four vehicles. However, it did not affect their detection performance as will be shown in Figure 18. Additionally, despite the fact that RXF, R-RXF and CRXF seemed to perform comparably according to the visual inspection of Figure 17, CRXF performed better than RXF and R-RXF. Similarly it was also true for ACAD.

Figure 16 Five-vehicle HYDICE image scene (a) five-vehicle image scene; and (b) ground truth map of five vehicles

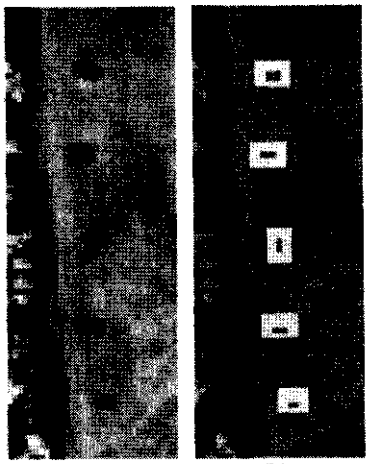


Figure 17 Anomaly detection results (a) RXF; (b) R-RXF; (c) CRXF; (d) DWEST with 7/13 dual windows; (e) ACAD; and (f) ACAD-generated anomaly map

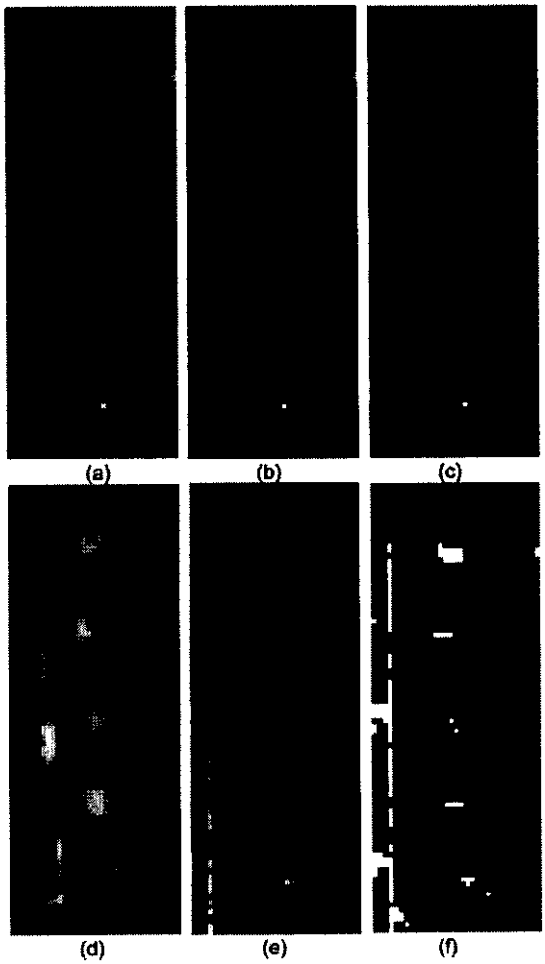
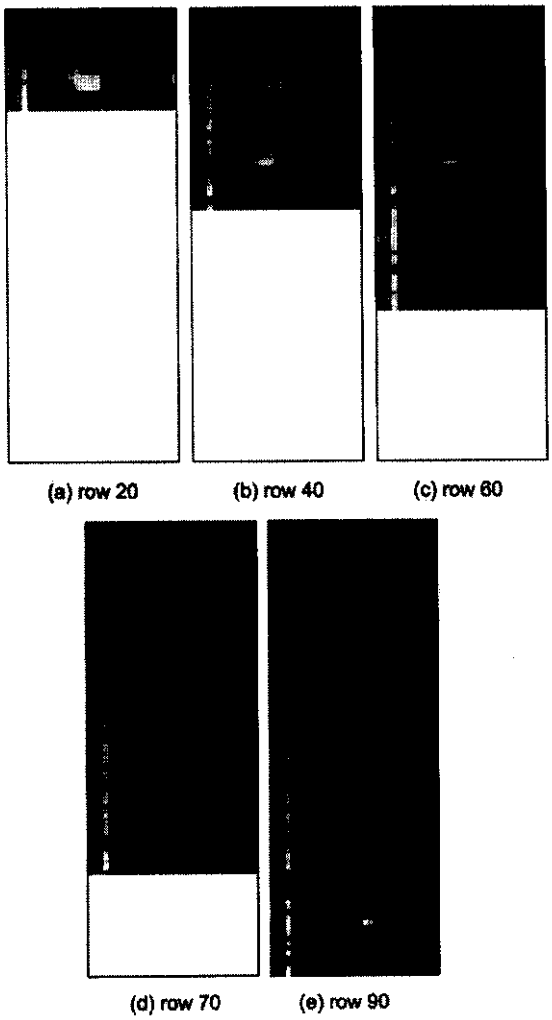


Figure 18 Real-time progressive anomaly detection by ACAD



Obviously, DWEST had shown blurring artifacts around the five vehicles which were caused by the use of dual windows. It could not identify the spatial locations of the five vehicles as also shown in Kwon *et al.* (2003). It is worth noting that all the five anomaly detectors detected the tree line with various gray scale values. This made sense since the tree line could be considered as anomalies with their spectral signatures distinct from their neighborhood.

6. Real-time implementation of ACAD

One of advantages resulting from ACAD is its ability in real-time implementation. Since the sample correlation matrix $\mathbf{R}(\mathbf{r}_k)$ can be expressed as:

$$\begin{aligned}\mathbf{R}(\mathbf{r}_k) &= ((k-1)/k)\left[\frac{1}{k-1}\sum_{i=1}^{k-1}\mathbf{r}_i\mathbf{r}_i^{\mathrm{T}}\right] + (1/k)\mathbf{r}_k\mathbf{r}_k^{\mathrm{T}} \\ &= ((k-1)/k)\mathbf{R}(\mathbf{r}_{k-1}) + (1/k)\mathbf{r}_k\mathbf{r}_k^{\mathrm{T}}\end{aligned} \tag{7}$$

where $\mathbf{R}(\mathbf{r}_k)$ can be updated by the incoming pixel $\mathbf{r}_k$ and the one-step past sample correlation matrix $\mathbf{R}(\mathbf{r}_{k-1})$. Using the

QR decomposition the computation of the inverse of $\mathbf{R}(\mathbf{r}_k)$ can be easily implemented in systolic arrays to realize real-time implementation (Hsueh, 2004). Using a similar approach, the sample correlation $\hat{\mathbf{R}}(\mathbf{r}_k)$ implemented in ACAD can be also realized in real time. A detailed description of real-time implementation of ACAD can be found in Hsueh (2004). Therefore, in this section, we only demonstrate experiments that show the real-time processing of ACAD operating on the image scene in Figure 16(a). Figure 18 shows real-time progressive anomaly detection carried out by ACAD.

As another example, real-time processing of a 15-panel HYDICE image scene by CRXF and ACAD can be found in Hsueh and Chang (2004) and Hsueh (2004).

7. Conclusions

Anomaly detection has been investigated extensively in the past. However, it seems that its characterization has not been addressed. This paper explores several issues to characterize anomalies for detection, which are:

- How large size for a target to be considered as an anomaly?
- How an anomaly responds to its proximity?
- How sensitive for an anomaly to noise?
- How different anomalies to be detected?

In order to address these issues, this paper develops a new ACAD that improves commonly used anomaly detectors including the RXF (Reed and Yu, 1990), the CRXF (Chang and Chiang, 2002; Chang, 2003) and an adaptive anomaly detector, DWEST (Kwon *et al.*, 2003). Most importantly, ACAD has advantages in several aspects. It can be implemented causally and processed in real time. It detects various anomalies regardless of whether they are of the same type or distinct types. In other words, it can detect anomalies regardless of their spectral signatures are similar or distinct. It also produces an anomaly map, which can spatially locate the detected anomalies in the original image. Such a map can be used for target classification and discrimination for future data processing. Additionally, this paper also explores an interesting finding that the ratio of the image size to the size of an anomaly is crucial and determines if a pixel is an anomaly. Furthermore, it also shows that anomaly detection is sensitive to noise, but does not have as much impact as the image size to anomaly size ratio does to the anomaly detection. Finally, it is worth noting that a field programmable gate arrays design for ACAD has been developed in Hsueh (2004) which provide feasibility of ACAD to be implemented in real-time.

References

Chang, C-I. (2003), *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic Publishers, New York, NY.

Chang, C-I. (2005), "Orthogonal subspace projection revisited: a comprehensive study and analysis", *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 43 No. 3, pp. 502-18.

Chang, C-I. and Chiang, S-S. (2002), "Anomaly detection and classification for hyperspectral imagery", *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 40 No. 6, pp. 1314-25.

Harsanyi, J.C. and Chang, C-I. (1994), "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach", *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 32 No. 4, pp. 779-85.

Hsueh, M. (2004), *Adaptive Causal Anomaly Detection*, MS thesis, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD, August.

Hsueh, M. and Chang, C-I. (2004), "Adaptive causal anomaly detection for hyperspectral imagery", paper presented at IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK.

Kwon, H., Der, S.Z. and Nasrabadi, N.M. (2003), "Adaptive anomaly detection using subspace separation for hyperspectral imagery", *Optical Engineering*, Vol. 42 No. 11, pp. 3342-51.

Reed, I.S. and Yu, X. (1990), "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 38 No. 10, pp. 1760-70.