Anomaly Detection for Hypaerspectral Imagery Using Analytical Fusion and RX

Guo-Liang Zhang, Chun-Ling Yang

School of Electrical Engineering and Automation
Harbin Institute of Technology
92 Xi Da Zhi Road, Harbin, 150001, China
zglhit@hit.edu.cn

ABSTRACT. *Anomaly detection is attractive for the analysis of hyperpectral imagery. This paper describes an expanded anomaly detection algorithm for small targets in hyperspectral imagery. As a variant of the well known multivariate anomaly detector called RX algorithm, the approach called the dimension reduction RX algorithm (DRRX) is proposed. The analytical fusion strategy is incorporated into the RX algorithm to lead to the DRRX algprithm. Experimental results are presented for the proposed DRRX and the classical constant false alarm rate (CFAR) RX algorithm for detecting small anomalies in hyperspectral imagery. The results show that the proposed DRRX algorithm outperforms the classical RX for detecting small targets in hyperspectral imagery.*
Keywords: Hyperspectral imagery; anomaly detection; analytical fusion; RX algorithm

1. **Introduction.** Hyperspectral imaging sensors provide image data containing both spatial and spectral information, and the information of hyperspectral imagery is used to perform target detection and recognition tasks [1]. In recent years, many target detection techniques in hyperspectral imagery have been widely investigated and have proven valuable in many applications including search-and-rescue operations, border surveillance, and mine detection. Anomaly detection methods with no a priori knowledge represent a current field of scientific research for the hyperspectral imagery and pattern recognition communities [2]. From a practical point of view, these methods could be useful tools for image analysts working with hyperspectral imagery. An anomaly detector enables one to detect targets whose signatures are spectrally distinct from their surroundings. In general, such anomalous targets are relatively small compared to the image background and only occur in the image with low probabilities. A constant false alarm rate (CFAR) detector was developed by Reed and Yu and called RX algorithm, which has shown success in anomaly detection for multispectral and hyperspectral images [3, 4]. Many scholars have studied this algorithm and several variants of the RX algorithm have been presented, such as NRXD, MRXD and CRXD [5], dual-window-RX [6], regularized-RX [7], kernel-RX [8], regularized-kernel-RX [9] and wavelet-RX [10]. However, the above proposed algorithms based on RX may have good performance for some certain scenarios.

It should be noted that the classical RX algorithm is computational cost, which needs to estimate the inversion of background covariance matrix. The covariance matrix is estimated from a small number of very high-dimensional data samples, therefore, its inverse is often unstable. In addition, the hyperspectral image with too many bands will cause Hough effect. In order to improve the RX algorithm, dimension reduction is critical.

Some dimension reduction methods have been proposed [11, 12, 13, 14]. It is noted that image fusion is a relatively new research field at the leading edge of data fusion technology. Image fusion can combine multiple images of the same scene to generate a new composite image with better quality. The fused images can provide a better interpretation of the scene than each of the single sensor image can do. Data fusion technique has been used to sharpen the hyperspectral images [15]. However, the fusion strategy in [15] may not meet the needs of target detection, which is sensitive to the image fusion techniques. Depending on the application and fusion method, research in the imagery fusion can be classified in two broad categories:image based representational fusion and video based analytical fusion [16]. Image based representational fusion, which has long history, mainly uses multi-resolution techniques to obtain best representation of the data. The pyramid approaches and the wavelet-based method have popularity in this area. However, the analytical fusion for specific issues are very recent. This type of fusion methodology is required to enhance the capabilities of automatic detection and tracking system for surveillance purpose. Cvejic N has shown that the averaging technique outperforms the Laplacian Pyramid and the Discrete Wavelet Transform in the application of target tracking by fusing the visible and infrared image at pixel-level fusion in [17]. This results demonstrated that the approaches which are used in the representational fusion applications can not be immediately employed in the analytical fusion area. The distance between the target and background pixel distributions is large in the analytical fusion, which befits target detection. The interested reader is refered to [18] about the analytical fusion. According to the above discussion, the analytical fusion approach will be investigated for detecting targets from hyperspectral images in this paper.

The rest of this paper is organized as follows. In section 2, the weighted average algorithm for analytical fusion is investigated. In section 3, we introduce the RX algorithm and integrate the weighted average method with RX to lead to an anomaly detector, which is called DRRX. Section 4 presented the simulation results of RX and DRRX to demonstrate the effectiveness of the proposed algorithm. Conclusions are finally drawn in Section 5.

2. **Weighted Average Algorithm.** In the image fusion applications, weighted average algorithm (WAA) is one of fundamental approaches. WAA has very high popularity because it can be performed in spatial and transform domain. In practical applications, the weighted average algorithm in spatial domain is the simplest image fusion method. It takes the weighted average of the source images pixel by pixel. The WAA can be presented as

$$IMG_o(i,j) = \sum_{i=1}^{L} w_l \cdot IMG_i(i,j), \quad \sum_{i=1}^{L} w_l = 1 \tag{1}$$

where IMG_l and IMG_o refer to the pixel values of the source image and fusion image at the location (i, j) respectively. w_l denotes the weight for each source image l.

This approach has been performed in many applications [19]. However, the critical technique of weighted average algorithm is the choice of weighted average coefficients. The weighted average algorithm should have the ability to determine the weights for the source images adaptively. According to the pixel in different bands, the weighted average coefficient is often represented as

$$w_l = \frac{IMG_l(i,j)}{\sum_{i=1}^{L} IMG_i(i,j)} \tag{2}$$

A special case is to make the weights equal, that means taking the average of the source images pixel by pixel. (i.e., $w_l=1/L$, $l = 1, 2, \ldots, L$). But this operation may lose significant information. In addition, correlation between different bands can be taken into account to determine the weights. For instance, when the two images fusion case is concerned, source of image should be specified nearly equal weights. When the low correlation between the source images appears, source image with significant features will be assigned a larger weight than the other. Actually, the calculation of correlation coefficients is computational cost. The weighted average algorithm includes several advantages. It is simple and intuitive. It can improve the reliability of detection via fusing complementary data. It has been proved that this approach has good performance in the analytical fusion for target detection and tracking.

3. **Anomaly Detectors.** In most applications, a priori knowledge about the desired target is unknown. In such cases, detectors should be designed to search anomalous spectra from the local background. This problem is typically formulated as a binary hypothesis test. The two hypotheses are background only or background plus target. The anomaly detectors are adaptive and designed using generalized likelihood ratio test (GLRT) approach. That means the unknown parameters (for example signal intensity and background covariance matrix) can be estimated from the acquired data. Different statistical models used for the background lead to different anomaly detection algorithms. The benchmark anomaly detection algorithm is RX algorithm.

3.1. **RX Algorithm.** The RX algorithm is an adaptive constant false alarm rate (CFAR) detector, which is derived from GLRT and a Gaussian multivariate distribution model. This algorithm can be mathematically expressed as

$$\delta_{RX}(r) = (r - \mu)^T \Sigma^{-1} (r - \mu) \begin{array}{cc} \geq \eta & H_1 \\ < \eta & H_0 \end{array} \qquad (3)$$

where r is the spectral pixel vector, μ represents the mean spectral vector within the background (the mean of each spectral band), Σ is the estimated background covariance matrix, and η is the threshold, which can be obtained according to the desired false alarm probability. H_0 means the sense is background noise only, and H_1 shows background noise and target in the sense. RX detector cues all targets in the scene that exhibit any significant spectral difference with respect to the local background. Its very simple to implement and has been widely used to solve the anomaly detection in multispectral and hyperspectral imaging. Although many variants of the RX algorithm have been presented, one limitation of RX algorithm should be highlighted: the estimation and inversion of large covariance matrices is a time-consuming process.

3.2. **Dimension Reduction RX Algorithm.** In order to reduce computational cost of the RX algorithm, efficient detection techniques are necessary. In this section, the analytical fusion method is explored to be incorporated into the RX algorithm. As mentioned previously, the weighted average algorithm has good performance in target detection applications. The RX has expensive computation for the hyperspectral images with high dimensionality. Thus the weighted average algorithm is first utilized to fuse the hyperspectral images. This operation is essentially dimension reduction, which fused hyperspectrla imagery into multi-band images (i.e.several bands). Then the multi-band images are processed by the RX algorithm to detect desired targets. There is no doubt that the proposed fusion algorithm can reduce computational cost of the RX algorithm, and make the RX algorithm perform more efficiently. Concretely the detection procedure of this novel and efficient algorithm is introduced in the following descriptions.

We assume that the acquired hyperspectral imagery with L bands will be processed. Generally speaking, the wavelength ranges from near infrared to far infrared according to the atmospheric window. For example, we divided the acquired hyperspectral imagery into three parts. Part one is the near infrared scope with L_n bands, part two presents the mid-infrared range with L_m bands, and part three shows the far infrared area with L_f bands. Obviously the summation of L_n, L_m, and L_f is L. Then the imagery of every part is fused into one band image using the weighted average algorithm. This operation assures that hundreds of bands hyperspectral imagery can be reduced into three bands multispectral imagery. This fusion stage can be expressed by the WAA.

$$IMG_n(i,j) = \sum_{i=1}^{L} w_{l_n} \cdot IMG_{l_n}(i,j), \quad \sum_{i=1}^{L} w_{l_n} = 1$$

$$IMG_m(i,j) = \sum_{i=1}^{L} w_{l_m} \cdot IMG_{l_m}(i,j), \quad \sum_{i=1}^{L} w_{l_m} = 1 \tag{4}$$

$$IMG_f(i,j) = \sum_{i=1}^{L} w_{l_f} \cdot IMG_{l_f}(i,j), \quad \sum_{i=1}^{L} w_{l_f} = 1$$

where w_{l_n}, w_{l_m}, w_{l_f} can be calculated by the equation (2). It is desirable to interpret why we divide the bands into three parts according to the wavelength. The emissivity of material varies slowly in adjacent bands, that is, the spectral characteristics of materials are considered to be approximate in continuous bands. In the demarcations of wavelength range, the intensity of target and background may be mutational. Therefore the assigned weights in different wavelengths accords with the essential fundamentals. It can be easily extended into only one kind of wavelength ranges (such as near infrared), which are divided into several parts. The idea behind this operation is dimension reduction. If the spectral characteristics of target are known or partial known, the accuracy of divided wavelength ranges will be higher. Furthermore the detection performance will be better. The criterion for divide the bands will be introduced later.

It is assumed that the raw hyperspectral images with L bands are divided into P parts according to the criterion. The images in each part are fused by using the WAA, i.e. the raw hyperspectral images are reduced to multispectral imgaes with P bands. We have $R(i,j)=[IMG_1(i,j), IMG_2(i,j), \ldots, IMG_P(i,j)]^T$ and $\mu = [\mu_1, \mu_2, \ldots, \mu_P]^T$. μ_p is the mean of the pth fused image. The dimension reduction RX algorithm can be given by

$$\delta_{DRRX}(R) = (R-\mu)^T \Sigma^{-1} (R-\mu) \begin{array}{cc} \geq \lambda & H_1 \\ < \lambda & H_0 \end{array} \tag{5}$$

where λ is the detection threshold, the denotation DRRX means dimension reduction RX. The dimension of the background covariance matrix becomes low after the analytical fusion stage. So the computational cost is reduced. However, its obvious that there are L^2 (L is all the bands) parameters need to be calculated in the conventional RX algorithm. If the symmetry of the background covariance matrix is considered, at least $L(L+1)/2$ parameters will be calculated. To speed up the calculation of RX algorithm further, the variant of the background covariance matrix is presented.

We process the multispectral imagery with P bands after the analytical fusion. The background covariance matrix Σ is commonly estimated using maximum likelihood estimator (MLE). Thus we have $\Sigma = (XX^T)/N$. X is the matrix of residual clutter noise-only processes, N is a scale presents the pixel samples. We denote $M = XX^T$, thus $\Sigma = M/N$.

Actually, the matrix M can be expressed as

$$\mathbf{M} = \begin{bmatrix} R_{11}(X) & R_{12}(X) & \cdots & R_{1P}(X) \\ R_{21}(X) & R_{22}(X) & \cdots & R_{2P}(X) \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ R_{P1}(X) & R_{P2}(X) & \ldots & R_{PP}(X) \end{bmatrix} \qquad (6)$$

It should be noted that the matrix M has two important characteristics. (1) $R_{ij} = R_{ji}, (i,j) \in [1, P]$; (2) $R_{ii} > R_{ij}, (i \neq j)$. Thus the matrix M is symmetrical and diagonally dominant. According to our research, the source images have strong correlation in the same wavelength range, and weak correlation in the different wavelength range. Thus M can be approximated by M^1 without diminishing the detection performance.

$$\mathbf{M} \approx \mathbf{M^1} = \begin{bmatrix} R_{11}(X) & 0 & \cdots & 0 \\ 0 & R_{22}(X) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & R_{PP}(X) \end{bmatrix} \qquad (7)$$

Then the background covariance matrix Σ can be expressed by M^1. We have $\Sigma \approx \Sigma_{dd} = (M^1/N)$, where dd denotes the diagonally dominant matrix. In the matrix Σ_{dd}, only P parameters need to be calculated, which can further reduce the computational cost. Thus the final DRRX algorithm can be obtained by equation (8), where γ is the detection threshold.

$$\delta_{DRRX}(R) = (R - \mu)^T \Sigma_{dd}^{-1} (R - \mu) \begin{array}{l} \geq \gamma \quad H_1 \\ < \gamma \quad H_0 \end{array} \qquad (8)$$

The procedures of DRRX algorithm can be outlined as follows via the above analysis.
Step 1: Acquire the real hyperspectral images. If it's difficult to obtain the real data completely, the simulation data based on the partial real data can be used.
Step 2: Divide the hyperspectral images into P parts. To make sure that the images lies in each part have high spectral similarity, while the images from different part have low spectral similarity.
Step 3: The analytical fusion approach based on the weighted average algorithm is employed to fuse the images in each part. The raw hyperspectral images with hundreds of bands can be reduced to multispectral images with P bands.
Step 4: Preprocessing the fused images by means of the whitening procedure.
Step 5: Run the DRRX algorithm, which is presented as equation (8).
Step 6: Output the detection results.

In the procedures of DRRX algorithm, dividing the bands in step 2 is difficult to implement. This operation should be based on the real spectral characteristics which are unknown. A rule for bands divided is proposed based on spectral correlation coefficients in the paper. The $P - 1$ larger gradients can be found by estimating the two adjacent bands over the entire wavelength. The spectral correlation coefficient can be defined as

$$C(A, B) = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (A - \bar{A})(B - \bar{B})}{\sqrt{\sum_{i=1}^{I} \sum_{j=1}^{J} (A - \bar{A})^2 \sum_{i=1}^{I} \sum_{j=1}^{J} (B - \bar{B})^2}} \qquad (9)$$

where A and B are the source images at adjacent wavelength. The image size is $I \times J$ pixels. $\bar{A}$ and $\bar{B}$ are the global mean of A and B respectively. $C(A, B)$ is the spectral correlation coefficient. Its noted that partial priori or full priori of the spectral can make the detection performance better. In step 4, the whitening procedure is to suppress the clutter background and make the residual noise to be Gaussian White noise with zero mean.

4. **Simulation Results Analysis.** In this section, we conducted a comparative analysis between RX algorithm and DRRX algorithm to demonstrate their relative performance using a series of computer simulation. Firstly, we presented the detection performance of RX algorithm in theory. In figure 1, the curves of the probability of detection versus the generalized signal to noise ratio (GSNR), which is defined in [3], for a fixed false alarm probability $P_{FA}=10^{-5}$ and various values of N and J are shown. N is the samples and J is the available bands. The curves show that for a fixed N, the CFAR detector has a lower detection probability if more bands are used to provide the same GSNR. This is because the number of unknown parameters in the background covariance matrix M increases as J gets larger. Thus we can deduce that the analytical fusion algorithm, which reduces the bands of images, can improve the detection performance of the RX algorithm.
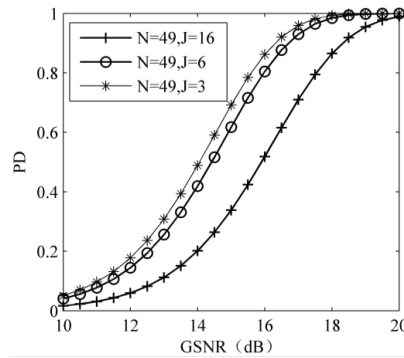

FIGURE 1. P_D versus GSNR for $N = 49$, $P_{FA}=10^{-5}$.

We designed the simulation hyperspectral data based on the real data from Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). The simulation data has 30 bands imagery, whose first band and third band are shown in figure 2. Images of this scene range from near infrared to long-wave infrared. The image size is 120×120 pixels. Twelve small targets with different GSNR are embedded into the images.
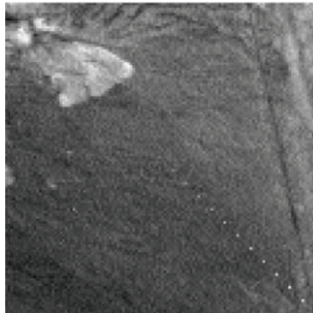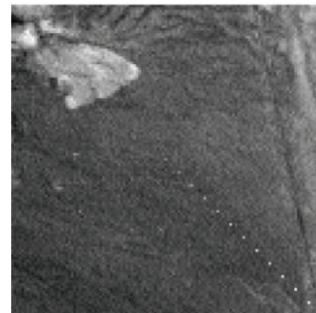
(a) Band 1   (b) Band 3

FIGURE 2. Hyperspectral imagery: (a) Band 1. (b) Band 3.

The spectral correlation coefficient between band 1 and band 3 is 0.97. This result demonstrates that the assumption about approximate spectral characteristics of materials in continuous bands is correct. According to the criterion for dividing bands, we have $P = 4$ in the paper. The detection results for RX and DRRX are presented in figure 3.

In order to compare the detection performance of DRRX with that of RX, the curves of receiver operating characteristic (ROC) are presented in figure 4, where PD is the probability of detection, Pfa denotes the probability of false alarm. PD can be defined
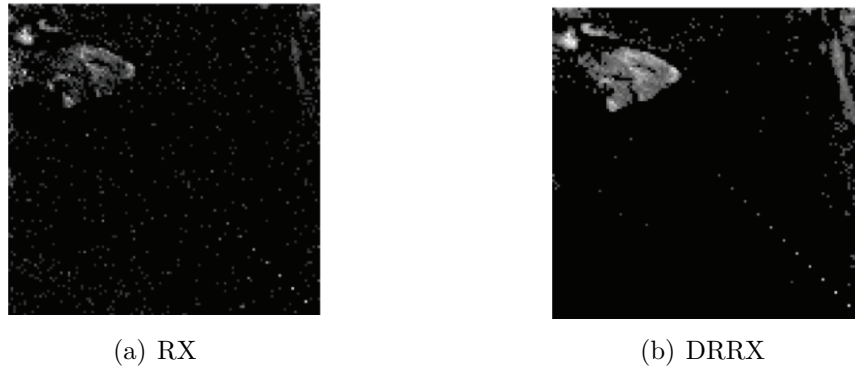
(a) RX   (b) DRRX

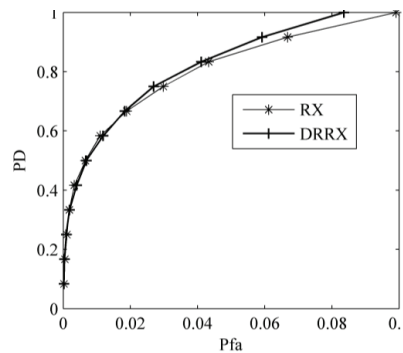FIGURE 3. Detection results of RX and DRRX: (a) RX result. (b) DRRX result.

FIGURE 4. Comparative of ROC curves of DRRX and RX.

as PD=N_d/N_t, and Pfa=N_f/N_b [20]. N_d is the detected pixels from target; N_t denotes all the pixels of the targets; N_f is the false alarms; N_b refers to the pixels of background. The ROC curves show that the performance of DRRX and that of RX are almost the same when the probability of false alarm is lower. That is because RX and DRRX both use the local multivariate Gaussian model and less interfere affects the performance. Meanwhile, the ROC curves demonstrate that the performance of DRRX is better than that of RX when the probability of false alarm increases. That's because the analytical fusion approach as a smoothing filter has reduced a part of clutter noise.

Finally we performed the RX algorithm and DRRX algorithm in DSP 6000. The time consuming of RX is 17434996815 clock cycles. And DRRX is 15492814777 clock cycles. This experiment demonstrates the effectiveness of DRRX, which can be performed in real-time process.

5. **Conclusions.** Anomaly detection is very important for hyperspectral imagery analysis. However, there is no one method that is good in all situations for all cases. In this paper we have introduced a novel approach of anomaly detection for hyperspectral imagery called dimension reduction RX (DRRX). DRRX algorithm is established by employing the analytical fusion approach and RX algorithm. We have demonstrated that the proposed DRRX algorithm outperforms the RX algorithm. Actually, the anomaly detection based on local Gaussian model has another limitation, which is often inadequate for the representation of the background distribution, especially in non-homogeneous conditions. This property may lead to excessive false alarms in applications. Thus the anomaly detection algorithms and analytical fusion methods should be further exploited.

<div align="center">

REFERENCES

</div>

[1] A. Plaza, J. A. Benediktsson, J. W. Boardman, and et al. Recent advances in techniques for hyper-spectral image processing, *Remote Sensing of Environment*, vol. 113, pp. S110–S122, 2009.

[2] S. Matteoli, M. Diani, and G. Corsini, A tutorial overview of anomaly detection in hyperspectral images, *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 5-28, 2010.

[3] I. Reed, and X. Yu, Adaptive Multi-Band CFAR Detection of An Optical Pattern With Unknown Spectral Distribution, *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760-1770, 1990.

[4] X. Yu, L. E. Hoff, I. S. Reed, A. M. Chen, and L. B. Stotts, Automatic target detection and recognition in multiband imagery: a unified ML detection and estimation approach, *IEEE Trans. Image Processing*, vol. 6, no. 1, pp. 143-156, 1997.

[5] C. I. Chang, and S. S. Chiang, Anomaly detection and classification for hyperspectral imagery, *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 6, pp. 1314-1325, 2002.

[6] H. Kwon, S. Z. Der, and N. M. Nasrabadi, Adaptive anomaly detection using subspace separation for hyperspectral, *Optical Engineering*, vol. 42, no. 11, pp. 3342-3351, 2003.

[7] N. M. Nasrabadi, Regularization for spectral matched filter and RX anomaly detector, *Proc. of SPIE, the International Society for Optical Engineering*, vol. 6966, pp. 1-12, 2008.

[8] J. M. Molero, E. M. Garzón, I. García, and A. Plaza, Anomaly detection based on a parallel kernel RX algorithm for multicore platforms, *Journal of Applied Remote Sensing*, vol. 6, no. 1, pp. 1-10, 2012.

[9] Z. W. Shi, J. Wu, S. Yang, and Z. G. Jiang, RX and its variants for anomaly detection in hyper-spectral images, *Infrared and Laser Engineering*, vol. 41, no. 3, pp. 796-802, 2012.

[10] A. Mehmood, and N. M. Nasrabadi, Wavelet-RX anomaly detection for dual-band forward-looking infrared imagery, *Applied Optics*, vol. 49, no.24, pp. 4621-4632, 2010.

[11] N. Gillis, and R. J. Plemmons, Dimensionality reduction, classification, and spectral mixture analysis using non-negative under approximation, *Optical Engineering*, vol. 50, no. 2, pp. 1-16, 2012.

[12] K. Burgers, Y. Fessehatsion, S. Rahmani, and J. Y. Seo, A comparative analysis of dimension reduction algorithms on hyperspectral data, University of California, Los Angeles, United States, Report, 2009.

[13] J. Wang, and C. I. Chang, Independent Component Analysis-Based Dimensionality Reduction With Applications in Hyperspectral Image Analysis, *IEEE Trans. Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1586-1600, 2006.

[14] B. Mojaradi, H. Abrishami-Moghaddam, M. J. V. Zoej, and R. P. W. Duin, Dimensionality reduction of hyperspectral data via spectral feature extraction, *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2091-2105, 2009.

[15] M. Moeller, T. Wittman, and A. L. Bertozzi, A variational approach to hyperspectral image fusion, *Pro. of SPIE* vol. 7334, pp. 73341-73341E-10, 2009.

[16] P. Kumar, A. Mittal, and P. Kumar, Study of robust and intelligent surveillance in visible and multi-modal framework, *Informatica*, vol. 32, pp. 63C-77, 2008.

[17] N. Cvejic, S. G. Nikolov, H. D. Knowles, A. Loza, and et. la, The effect of pixel-level fusion on object tracking in multi-sensor surveillance video, *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-7, 2007.

[18] X. W. Zhang, Y. N. Zhang, Z. Guo, and et al. Advances and perspective on motion detection fusion in visual and thermal framework, *Journal of Infrared and Millimeter Waves*, vol. 30, no. 4, pp. 354-360, 2011.

[19] R. C. Luo, and C. C. Chang, Multisensor fusion and integration: a review on approaches and its applications in mechatronics, *IEEE Trans. Industrial Informatics*, vol. 8, no. 1, pp. 49-60, 2012.

[20] J. Kerekes, Receiver operating characteristic curve confidence intervals and regions, *IEEE Trans. Geosci. Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 251-255, 2008.