Recent advances in hyperspectral image processing

ZHANG Liangpei* and DU Bo

"State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; "School of Computer Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

(Received 10 June 2012; final version received 20 July 2012)

Hyperspectral images (HSI) provide a new way to exploit the internal physical composition of the land scene. The basic platform for acquiring HSI data-sets are airborne or spaceborne spectral imaging. Retrieving useful information from hyperspectral images can be grouped into four categories. (1) Classification: Hyperspectral images provide so much spectral and spatial information that remotely sensed image classification has become a complex task. (2) Endmember extraction and spectral unmixing: Among images, only HSI have a complete model to represent the internal structure of each pixel where the endmembers are the elements. Identification of endmembers from HSI thus becomes the foremost step in interpretation of each pixel. With proper endmembers, the corresponding abundances can also be exactly calculated. (3) Target detection: Another practical problem is how to determine the existence of certain resolved or full pixel objects from a complex background. Constructing a reliable rule for separating target signals from all the other background signals, even in the case of low target occurrence and high spectral variation, comprises the key to this problem. (4) Change detection: Although change detection is not a new problem, detecting changes from hyperspectral images has brought new challenges, since the spectral bands are so many, accurate band-to-band correspondences and minor changes in subclass land objects can be depicted in HSI. In this paper, the basic theory and the most canonical works are discussed, along with the most recent advances in each aspect of hyperspectral image processing.

Keywords: hyperspectral images; classification; spectral unmixing; endmembers extraction; target detection; hyperspectral change detection

1. Introduction

Environmental protection and ecological pressures have dramatically accelerated the development of remote sensing techniques. Remote observation from space/airborne platforms is more powerful than traditional ways of collecting data for large-scale land coverages requiring short revisiting periods. One basis of remotely sensed observation is that spectral analysis is proven effective in distinguishing different objects, especially those land cover types with visually indiscriminate appearances. Even very similar objects will have obviously different spectral stamps at a very fine spectral resolution. Spectral imaging provides a high spectral resolution from hyperspectral images by decomposing the reflected sonar radiance into large number of bands with minor spectral resolutions, so that the spectra of the different land objects present a nearly continuous shape. The spectral shape reveals the internal physical features of each material and is more reliable than visual feature classification. Moreover, spatial information, including shape and texture, can also be taken into consideration. By combining both the spectral and spatial information, hyperspectral images promise the ability to classify different land objects types and even subclass types. No other technology besides hyperspectral images provides such a capacity (1).

Hyperspectral images contain overwhelming information about different kinds of land objects. Several obstacles, however, have to be conquered in hyperspectral image processing. Hyperspectral images that have three important differences from conventional multispectral and pan images, such as high spectral resolution; continuous spectra; and mapping and spectra, are provided. In other words, a high-dimensional signal is the first challenge in hyperspectral image analysis. Actually, the spectral dimension can be over 1000. Thus, the number of samples may not always be larger than the number of the image dimension. Other problems originate from the limited spatial resolution, the spatial resolution is less than the size of most land object types, and so mixed pixels are unavoidable. In this case, conventional pure pixel-based methods are no longer suitable. For example, a matched filter would not expose incorrect results due to the complex spectral signature from mixed land covers. Figure 1 formulates the complex spectral variation phenomenon (2).

Exploiting the underlying information in hyperspectral images has drawn great interest in the remote sensing...
fields, due to its superiority in portraying so many physically different land types. Depending on different applications in different domains, several domains have been developed.

(1) First, like the multiple spectral remote sensing images, classifying different types of land cover is critical and very informative since the subclassification becomes possible from the plentiful spectral information contained in hyperspectral images. Classification methods for hyperspectral images benefit from both spectral and spatial features, which makes the spatial/spectral information fusion-based methods very competitive.

(2) Second, one notable difference in hyperspectral image processing is that the typical elemental spectral signatures can compose the pixels’ spectra. This is due to the fact that the spectral resolution is high enough to find these most representative spectral signatures, or called “endmembers” (3); the elements in the mixed spectrum model. Endmember extraction, in this case, is another subject in hyperspectral image processing. With proper endmembers, the quantitative analysis from the inspection of mixed pixels is made possible. Typical methods include least squares and its constrained variants (4). This unmixing processing following endmember extraction decomposes the image into a series of abundance maps, each map suggesting the spatial distribution of each kind of land cover.

(3) The third aspect addresses the extraction of certain land objects of interest or detecting targets. Target detection from hyperspectral images is of great significance, since targets are assumed to be of low presence probability so that they are not easy to separate by conventional classification methods. They might be detectable with a generalized likelihood ratio test (GLRT). Subspace and physical models have also shown promising performance for target detection.

(4) The last issue is change detection from multi-time HSI data. It means that the different HSI data from different times on the same scene can be compared to calculate any possible land cover change information. The fine spectral representation capability seems to be superior for describing possible minor changes in land cover. However, change detection may be the least mature among the four aspects mentioned. The numerous bands make the strict band-to-band matching difficult, while spectral correction between two images may be also an obstacle.

Section 2 details the four parts of hyperspectral image processing.

2. Hyperspectral classification

An interesting and important research problem is discerning precisely and automatically different land-cover types in hyperspectral images. Generally, each pixel in a hyperspectral image consists of hundreds or even thousands of bands. This makes the discrimination among pixels of different land-cover types a high-dimensional data classification problem. As pointed out in Ref. (5), a corresponding large number of training samples are required for achieving satisfactory accuracy in this kind of problem. However, in hyperspectral image classification, the number of available training samples might be very limited. A great deal of research has aimed to find more efficient ways to overcome this problem. Recent research methodologies to solve the problems can be considered from three aspects: the first one is utilizing dimension reduction to reduce the number of features; the second one is utilizing more robust classifiers to overcome the small sample problem; and the third one is optimizing the way of sampling.

2.1. Dimension reduction

Dimension reduction algorithms have been proposed as a most direct way to achieve this problem, which can be classified into two kinds: feature extraction and band selection. Linear discriminant analysis (LDA) (6) is the most popular and practical feature extraction methods in classification application, since Fisher’s criterion can effectively model the separability. However, generally, it is not a proper feature extraction method for the classification of hyperspectral data, due to several shortcomings of the Fisher’s criterion. Many adjustments have been proposed to solve the problems of LDA by designing new scatter matrices to measure the class separability. Among these methods, nonparametric weighted feature extraction (7) and double nearest proportion feature extraction (8) are creative extensions designed to overcome the various shortcomings attaining better performance. As another direction in feature extraction methods, the manifold learning algorithms (9, 10) effectively generate a discriminant subspace, constructing a model to maximize the distance between different-class samples and minimize the distance between same-class samples.

Band selection processing is the most prevailing method to reduce feature number. The simplest suboptimal search strategy employs sequential forward selection and sequential backward selection techniques (11), which achieve the best feature subset with a prefixed number of features by adding to or removing from the current feature subset one feature at a time. Besides these two methods, many new search algorithms have been devised for dimensionality reduction, keeping pace with the rapid development in soft computing. Representative among
these new search algorithms, a genetic algorithm approach has been applied to feature selection (12). Recently, Zhong et al. treat dimensionality reduction as an optimization problem and utilize a clonal selection algorithm which searches for an optimum solution of dimensionality reduction with fewer features in a feature space (13). Experimental results demonstrate that these proposed algorithms outperform other algorithms and hence provide effective new options for dimensionality reduction in hyperspectral remote sensing imagery.

2.2. Robust classifiers

One possible approach to handle the high-dimensional nature of hyperspectral data-sets is to consider the geometrical properties rather than the statistical properties of the classes. Good classification performance is demonstrated by support vector machines (SVMs) using spectral signatures as input features (14). Variations of the SVM-based algorithms have also been proposed to improve the classification accuracy. These variations include semisupervised learning which exploits both labeled and unlabeled samples (15), postprocessing of the individually labeled samples based on certain decision rules (16), and incorporating spatial information directly in the SVM kernels (17, 18). Recently, multinomial logistic regression (MLR) (19) has been proposed to be an alternative approach to deal with ill-posed problems. MLR has the advantage of learning the class probability distributions themselves. Semisupervised variations can also be found in Ref. (20).

In recent years, artificial intelligence theory has been widely used in the remote sensing field. Zhong et al. first applied artificial immune systems to unsupervised and supervised classification of multi/hyperspectral remote sensing images (21–23). Experimental results suggest that these artificial immune classifiers for remote sensing imagery can yield better results than traditional classification algorithms, such as the maximum likelihood classifier. Furthermore, they proposed adaptive artificial immune network to overcome the inherent complexity of current AIN models. As a novel branch of computational intelligence, DNA computing has also been used in the hyperspectral land cover classification, which has the strong computing and matching capability to discriminate the tiny differences in DNA strands by DNA encoding and matching in the molecule layer (24). Experimental results demonstrate that this algorithm presents excellent processing efficiency superior to a SVM. Recently, sparse data representation has been extended to hyperspectral land cover classification, based on the observation that a hyperspectral pixel can be sparsely represented by a linear combination of a few training samples from a structured dictionary. Experimental results show that sparse representation method outperforms the classical supervised classifier SVMs in most cases.

2.3. Optimizing the way of sampling

From the aspect of sampling for the land cover classification, two popular machine learning approaches for dealing with this problem are semisupervised learning and active learning. Semisupervised algorithms incorporate the unlabeled data into the classifier training phase to obtain better decision boundaries. Essentially, three different classes of semi-supervised learning algorithms are encountered in the literature: (1) techniques based on expectation–maximization algorithms (25); (2) low density separation algorithms (26); and (3) graph-based methods (27).

In contrast, active learning assumes the existence of a rudimentary learner trained with a small amount of labeled data. The learner has access to both the unlabeled data and a “teacher.” The learner then selects an unlabeled data point and obtains its label from the teacher. The goal of the active learner is to select the most “informative” data points so as to accurately learn from the fewest such additionally labeled data points.

Several active learning methods have been proposed so far. They may be grouped into three different classes: (1) active learning methods relies on SVM specificities, which samples the candidates lying within the margin of the current SVM by computing their distance to the dividing hyperplane (28); (2) active learning methods relies on the estimation of the posterior probability distribution function of the classes and label the samples with high uncertainty (29); and (3) active methods based on the query by committee paradigm. The algorithm selects the samples where the disagreement between the classifiers is maximal (30).

Another active research direction is how to incorporate the spatial information in the land cover classification. This type of processing has been approached in the past from various points of view. For instance, several possibilities are discussed in (1) for the refinement of results obtained by spectral-based techniques in multi-spectral imaging through a second step based on spatial context. This simple operation separates spatial from spectral information, and thus the two types of information are not treated simultaneously. Thus, many methods have proposed to effectively combine the spectral and the spatial features to generate better classification result and can be grouped into three classes: (1) the first class of techniques is to extract spatial features by specific operation, such as Wavelet transform, gray level co-occurrence matrix (31), structural and shape features (32), and so on. And then construct corresponding model to integrate the different sources of features. For example, Zhang et al. use patch alignment framework to linearly combine multiple features in the optimal way and obtain a unified low-dimensional representation of these multiple features for subsequent classification (33); (2) the second class of techniques is based on the tensor, Zhang L et represent the image object as a multifeature tensor that encodes spectral–textural information (34); and (3) the third class of the techniques is based on the
basic assumption that, in a hyperspectral image, it is very likely that two neighboring pixels will have the class same label (35).

3. Hyper spectral unmixing

Spectral unmixing refers to decompose a single mixed spectrum into several elementary spectra and their corresponding abundances simultaneously. In this way, the image interpretation can be deepened into a subpixel level, which is very useful for accurate land object mapping and the physically spectra inversion. Hyperspectral image has a large number of spectral bands and its spectral resolution is also very fine, so that it is able to reveal the diagnostic feature of different materials, which makes hyperspectral spectral unmixing much more meaningful than multispectral images’ spectral unmixing. Hyperspectral spectral unmixing has become one of the hottest topics in the information extraction field by remote sensed images.

The pixels exist brobably in the hyperspectral images, and spectral unmixing is the effective approach to solve this problem to realize the subpixel classification. Spectral unmixing is an important task for remotely sensed hyperspectral data exploitation (36). It expresses each (possibly mixed) pixel of the hyperspectral image as a combination of spectrally pure substances (called endmembers) weighted by their corresponding abundances (37). To investigate the spectral unmixing, we firstly study the spectral mixture model.

3.1. Linear and nonlinear spectral mixture model

The linear mixture model (LMM) has been widely used due to its simple physical interpretation and tractable estimation process. Consequently, many linear unmixing algorithms achieve excellent performance and are well-understood.

The linear spectral mixture Model (LSMM) assumes that the spectral response in each pixel is a linear combination of endmember spectra, with the weights being proportions. Let \( y \) denotes the \( l \) vector for one pixel in the image, \( l \) is the number of the bands, the mathematical formulation of LSMM is:

\[
y = \sum_{i=1}^{l} M_i a_i + e = M a + n
\]

where \( p \) is the number of the endmembers, \( M \) is the endmember matrix or source matrix, \( a \) is the abundance vector for the pixel, and \( n \) represents the error term. There are two constraints for the abundance vector: abundance non-negativity constraint (ANC) and abundance sum-to-one constraint (ASC).

There are many situations in which it may not be appropriate and could be advantageously replaced by a nonlinear one. Nonlinear mixing models provide an interesting alternative for overcoming the inherent limitations of the LMM. Nonlinear unmixing has generated considerable interest among researchers and different methods have been proposed to account for nonlinear effects. Let \( y \) denotes the \( l \) vector for one pixel in the image, \( l \) is the number of the bands, the mathematical formulation of the nonlinear spectral mixture model (NSMM) is:

\[
y = g \left( \sum_{i=1}^{l} a_i m_i \right) + n = g(Ma) + n \quad (2)
\]

where \( l \) is the number of the endmembers, \( M \) is the endmember matrix or source matrix, \( a \) is the abundance vector for the pixel, \( n \) represents the error term, and \( g \) is an appropriate nonlinear function.

(1) Direct nonlinear model. This kind of model has certain physical interpretation and suit for a certain scene, Somers et al. (38) and Nascimento and Bioucas-Dias (39) have extended the collection of endmembers by adding artificial cross-terms of pure spectral signatures to model light scattering effects on different materials.

(2) Neural network. A novel approach based on neural networks for the extraction of pixel abundances from hyperspectral data is developed (40). This work expands over previous efforts in the literature focused on using neural networks for nonlinear unmixing purposes (41, 42).

(3) Postnonlinear mixing models (PNMMs). PNMMs are flexible generalizations of the standard LMMs that have been introduced (43, 44) for source separation problems. The main advantage of PNMMs is that they can accurately model much different nonlinearity. A generalized bilinear model algorithm for nonlinear unmixing of hyperspectral images are proposed (45).

(4) Kernel-based model. Another kind of nonlinear mixed model is based on kernel methods (4), which turns LSMM into NSMM through kernel function.

Generally, based on the spectral models, the spectral unmixing procedure includes three main steps: (1) estimation of the number of endmembers in a scene; (2) automatic identification of the spectral signatures of these endmembers; and (3) estimation of the endmember abundances in each pixel of the scene. Over the last years, several algorithms have been developed for each part of the chain. Figure 1 shows the spectral unmixing procedure. In Figure 2, given the hyperspectral data-set: (1) dimensionality reduction improves algorithm performance and data storage, (2) extract the endmember signatures from the data-set, and (3) given the identified endmembers, the inversion step gets the abundance maps while solving a constrained optimization problem.

Spectral unmixing can be grouped in two categories, the one is endmember extraction, then based on endmember estimate abundance and the other is blind unmixing, which can extracts endmember and estimates abundance at the same time.
3.2. Endmember extraction algorithms

During the past few decades, a great number of endmember extraction algorithms (EEAa) have been proposed. EEA are grouped into automatic methods and interactive methods. The interactive methods of selecting endmembers find it hard to acquire the whole endmember spectra; they also consume more processing time compared to the automatic methods. Over recent decades, several algorithms have been developed for automatic or semiautomatic extraction of spectral endmembers. Based on different criterion, the EEAs can be categorized into different classes. Among the endmember extraction methods, according to the assumption as to whether pure pixels exist in the image, endmember extraction methods can be classified into two groups: endmember identification methods and endmember generation methods.

3.2.1. Endmember identification methods

Endmember identification is based on the assumption that the input data-set contains at least one pure pixel for each distinct material present in the scene; therefore, a search procedure that aims at finding the most spectrally pure signatures in the input scene is feasible. Some of the classic approaches, such as the pixel purity index (46–48), N-finder algorithm (49), vertex component analysis (50), simplex growing algorithm (47, 51), convex cone analysis (52), sequential maximum angle convex cone (53), Iterative error analysis (IEA) (54), orthogonal subspace projection (OSP) (55, 56), and automatic morphological endmember extraction (18, 36, 57), are endmember identification approaches. However, the assumption behind these algorithms may not be true in practical applications.

3.2.2. Endmember generation methods

Therefore, endmember generation algorithms have been proposed, under the assumption that pure signatures are not present in the input data. This group is combining some transform or matrix factorization with some constraints. Such algorithms include minimum volume transform (58), minimum volume simplex analysis (59, 60), constrained nonnegative matrix factorization (61), minimum volume constrained non-negative matrix factorization (MVC-NMF) (62), and simplex identification via split augmented Lagrangian (63). A new hybrid automatic endmember extraction algorithm (64) inspired by IEA and the OSP method, which integrates the spatial and spectral information, considers the correlation and similarity between endmembers based on a local window. Spatial preprocess endmember extraction (65) considers the spatially homogeneous areas to enhance endmember extraction accuracy.

3.3. Hyperspectral blind unmixing algorithms

Based on LMM, traditional methods for unmixing always implemented the endmember extraction step first then the inversion step, therefore, the accuracy of the endmember extraction results will affect the accuracy of the estimation for the abundances. To avoid this, there are many hyperspectral unmixing approaches which implement endmember extraction and inversion simultaneously, which can be called as blind unmixing.

Blind unmixing approaches are generally based on two theories: (1) independent component analysis (ICA) and (2) NMF.

3.3.1. Hyperspectral blind unmixing based on ICA

ICA extracts independent sources from mixed observations with the assumption that the sources are statistically independent. The basic model is as follows:

$$ X = AS + N $$

where $X$ is the mixed observations, $A$ is the mixed coefficient, $S$ is the original independent sources, and $N$ is the noise.
ICA can be used for hyperspectral unmixing, with the assumption that the abundances are statistically independent vectors; however, the ASC, which is the basis of LMM, obviously contradicts this assumption. A solution to this problem was recently proposed through the combination of ICA and Bayesian positive source separation (66), which provides a promising way of separating physically meaningful spectral signals. Chang et al. (67) also proposed a ICA-based abundance quantification algorithm (ICA-AQA), which is a high-order statistics-based technique. ICA-AQA avoids the disadvantage of ICA for unmixing since ICA cannot be implemented as a constrained method due to its independency assumption. Nascimento and Bioucas-Dias (68) proposed a dependent component analysis, which assumes the abundances are Dirichlet distribution to avoid the fact that ICA does not satisfy the ASC and ANC.

3.3.2. Hyperspectral blind unmixing based on NMF

NMF (69, 70) decomposes a high-dimensional data-set into two nonnegative matrices: one consisting of “basis vectors” and the other of “coefficient vectors”:

\[
\min_{A \geq 0, S \geq 0} f(A, S) = \|X - AS\|_F
\]

where \(X\) is the high-dimensional data-set, \(A\) is the basis vectors, and \(S\) is the coefficient vectors.

When applied to hyperspectral unmixing, NMF does not require the pure-pixel occurrence assumption. Because of this advantage, NMF has drawn much attention in the field of hyperspectral imagery. Unfortunately, due to the nonconvexity of the objective function in NMF, the algorithm may fall into many local minima. To avoid this, an improved model was proposed to introduce auxiliary constraints into the NMF algorithm, according to different applications:

\[
\min_{A \geq 0, S \geq 0} G(A, S) = f(A, S) + \lambda_1 g_1(A) + \lambda_2 g_2(S)
\]

where \(f(A, S)\) is defined by (2), \(g_1(A)\) is a function constraining the spectral matrix \(A\), which is constructed by the spectral properties, and \(g_2(S)\) is a function constraining the abundance matrix \(S\), which represents the ground objects’ distribution properties.

In the hyperspectral unmixing field, according to the properties of the signatures and the ground distribution, different constrained algorithms have been developed for the spectral unmixing applications. One approach for spectral unmixing is minimum dispersion constrained NMF (71), in which the variance of the spectral matrix is employed to constrain the recovered spectra as flat as possible, but preserving the possible spectral singularities (peaks and sharp variations). Miao and Qi (62) proposed an approach for the endmember extraction from hyperspectral data, named MVC-NMF. The traditional NMF drives the estimation of endmembers by moving outwards through the data cloud, considered as the “external force,” while the MVC-NMF algorithm introduces a minimum volume constraint as the “internal force” which tries to force the endmembers to circumscribe the data cloud. The true endmembers are obtained by balancing the tradeoff between the two forces. Another approach for hyperspectral unmixing, named piecewise smoothness NMF with a sparseness constraint (72), was proposed by Jia and Qian, which imposes both piecewise smoothness and sparseness constraints on NMF. Liu proposed an approach named abundance separation and smoothness constrained NMF (73), which introduces two constraints, namely, the abundance separation constraint and abundance smoothness constraint, into the basic NMF. Wang et al. (74) introduced an endmember dissimilarity property into the NMF algorithm that aims at searching a set of vectors with the least similarity so as to find the global optimal solution to basic NMF and to improve the physical meaning of the resulting endmembers.

4. Target detection

A great deal of research has aimed to find more efficient ways to perform the target detection application. Recent research methods can be classified into two kinds: full-pixel targets methods and subpixel targets methods.

4.1. Full-pixel targets-based detection methods

These methods are based on the hypothesis that all the pixels in the images do not contain mixed spectra. Full-pixel targets-based detection performance is mainly determined by the variability of target and background spectra. These methods can be separated into two subgroups: target and background with known statistics and anomaly detection.

4.1.1. Target and background with known statistics

Statistical methods are very important in target detection and statistical methods based on normal probability models, especially, are widely used in target detection. In these methods, target and background spectra are regarded as random vectors with multivariate normal distributions:

\[
H_0 : x \sim N(\mu_0, \Sigma_0), \text{ target absent} \\
H_1 : x \sim N(\mu_1, \Sigma_1), \text{ target absent}
\]

Different distributions under the binary hypotheses make different methods. The fisher’s linear discrimination (75), constrained energy minimization (CEM) (76), spectral angle mapper (SAM) (77), etc. are typical examples of target and background with known statistics-based methods.

The fisher’s linear discrimination is based on the hypothesis that the target and background classes follow
multivariate normal distributions with different mean vectors and the same covariance matrices. And the likelihood-ratio detector takes the form of a linear processor:

\[
y = D(x) = k \Gamma^{-1}(\mu_1 - \mu_0)x
\]

where \( k \) is a normalization constant.

The CEM algorithm is to design a linear filter which can minimize the total energy of the output of the linear filter and confine the target spectra of the output of the linear filter to 1.

\[
y(n) = c^T x(n)
\]

\[
E = \frac{1}{N} \sum_{n=1}^{N} y^T(n) = c^T \left[ \frac{1}{N} \sum_{n=1}^{N} x(n)x^T(n) \right] c = c^T Rc
\]

The matrix \( R \) is the estimated correlation (not covariance) matrix of the cube. The CEM filter is given by:

\[
c_{\text{CEM}} = \frac{R^{-1}\mu_1}{\mu_1^T R^{-1} \mu_1}
\]

The SAM is the cosine of the angle between the test and target spectra and is always between zero and one.

### 4.1.2. Anomaly detection

In many situations of practical interest, there is no sufficient a priori information to specify the statistics of the target class. The method used under these circumstances is known as the anomaly detector. The hyperspectral data are considered to have a certain distribution and the anomalies are the pixels that do not fit the distribution. RX and its extended algorithms (78–81) are efficient methods for anomaly detection.

RX algorithm is the Mahalanobis distance of the test pixel spectrum from the mean of the background class. We consider only the case \( \Gamma_0 = \Gamma_1 \Gamma \), which leads to a mathematically tractable problem.

\[
y = D(x) = (x - \mu_0)^T \Gamma^{-1}(x - \mu_0)
\]

For simplicity, we usually estimate the parameter \( \mu_0 \) using the whole image data.

The main improvement of RX is using the local calculation instead of global calculation. Local anomaly detectors are characterized by the use of processing windows that are passed over every pixel in an image to find anomalies. Reed and Yu (82) fostered the local window scheme in the RX detector

\[
y = D(x) = (x - \mu)^T \times \left( \frac{N}{N+1} C + \frac{1}{N+1}(x - \mu)(x - \mu)^T \right)^{-1}(x - \mu)
\]

where \( N \) is the number of pixel vectors in the processing window, \( c \) is the window covariance matrix, and \( \mu \) is the window mean vector.

Recently, several new algorithms were introduced into full target detection, such as kernel-based methods and random-selection-based anomaly detector. Kwon and Nasrabadi use kernel methods to implicitly transform the data into a higher dimensional feature space. The RX detector is then applied in this higher dimension. On this basis, kernel methods are introduced into other detectors. Jiao and Chang (83) investigates a kernel-based CEM which employs various kernels to expand the original data space to a higher dimensional feature space that CEM can be operated on. Wang et al. (84) proposes a local subspace based nonlinear OSP method for target detection in order to optimize the background subspace and suppress false alarms better.

In addition, Du and Zhang (81) proposes an anomaly detection method based on the random selection of background pixels. Pixels are randomly selected from the image scene to represent the background statistics; the random selections are performed a sufficient number of times; blocked adaptive computationally efficient outlier nominators are used to detect anomalies each time after a proper subset of background pixels is selected; finally, a fusion procedure is employed to avoid contamination of the background statistics by anomaly pixels. In addition, the real-time implementation of the random selection based anomaly detection is also developed by random selection from updating data and QR decomposition.

### 4.2. Subpixel targets-based detection methods

Due to the complex conditions in the field and the limitation of spatial resolution of the hyperspectral data, mixed pixels occur widely in hyperspectral data (85). Subpixel targets occupy only part of the pixel area. The remaining part is filled with one or more materials, which will be collectively referred to as background. Subpixel targets-based detection performance is affected by background interference as well as by the variability of target and background spectra. The choice of the mathematical model used to describe the variability of target and background spectra leads to different families of subpixel target detection methods. The variability of the target spectral signature is always described using a subspace model. Therefore, the choice of the background model leads to two different classes of subpixel target detection methods. The most widely used spectral mixing model is the linear mixing model.

#### 4.2.1. Unstructured background models

Unstructured background models assume that the additive noise has been included in the background, which in turn is modeled by a multivariate normal distribution with mean zero and covariance matrix, such as those shown in Figure 3. Since HSI data have a nonzero mean,
we usually remove the estimated mean from the data cube and the target subspace vectors to comply with this model. The competing hypotheses are:

\[ H_0 : x = v, \text{ target absent} \]
\[ H_1 : x = Sa + v, \text{ target present} \]

The Kelly detector (86, 87), the adaptive matched filter (AMF) detector (88, 89), and the adaptive coherence/cosine estimator (ACE) detector (90, 91) are typical methods for unstructured background models-based detection methods.

The Kelly detector is obtained by using the generalized likelihood ratio (GLR) approach. A key assumption in the derivation for the Kelly detector is that the covariance matrix of the background is the same under the two hypotheses. While the ACE detector uses a more appropriate hypotheses that the background has the same covariance structure under both hypotheses, but with different variances as the amount of background covered area is different under the two hypotheses for subpixel targets.

For targets with amplitude variability, we have \( p = 1 \) and the target subspace \( S \) is specified by the direction of a single vector \( s \). Then the formulas for the previous GLR detectors are simplified to:

\[ y = D(x) = \frac{(s^T \Gamma^{-1} s)^2}{(s^T \Gamma^{-1} s)(\phi_1 + \phi_2 x^T \Gamma^{-1} x)} \]  

where \( \phi_1 = N, \phi_2 \) presents the Kelly detector, \( \phi_1 = 0, \phi_2 \) presents the ACE, and \( \phi_1 = N, \phi_2 \) presents the AMF.

### 4.2.2. Structured background models

Structured background models assume that the background variability is modeled using a subspace model.

\[ H_0 : x = Ba_{b0} + w, \text{ target absent} \]
\[ H_1 : x = Sa + Ba_{b1} + w, \text{ target present} \]

The target endmembers matrix has to be specified by the user, the background endmembers matrix is determined from the data cube and \( w \sim N(0, \sigma_w^2) \). Figure 4 provides a geometrical illustration of the structured background model.

The adaptive subspace detector (ASD) (92, 93) and the OSP (94) are examples of this kind of method.

The only source of randomness for the ASD is the additive noise with unknown variance. The unknown model parameters, such as the abundance coefficient matrices and the noise variance, are computed by using the maximum likelihood estimation method.

\[ y = D(x) = \frac{x^T(P_p - P_{Ze})x}{x^T P_{Ze}x} \]  

OSP can remove the influence of the background and can maximize the remaining signal considering various

---

Figure 3. Illustration of signal model for subspace targets in a homogeneous normal background.

Figure 4. Illustration of structured background model.
kinds of noise, such as white Gaussian noise, non-Gaussian white noise and non-white noise, etc.

$$y = D_{OSP}(x) = s^TP_1x$$

(16)

Recently, several new algorithms were introduced into subpixel target detection, such as noise model based methods. For subpixel targets-based detection methods (95) consider the problem of detecting a subspace signal in white Gaussian noise when the noise power may be different under the null hypothesis – where it is assumed to be known – and the alternative hypothesis. Then they derive the GLRT for the problem.

Besides all above, different thinking was introduced into target detection. Du et al. managed to improve target detection performance in terms of the data. They used spectral transformation methods followed by noise adaptive principal component analysis (NAPCA) for the data. Then $k$-nearest neighbor clustering is applied to the NAPCA-transformed feature subspace to achieve detection (96, 97) used random projection for dimensionality reduction of hyperspectral imagery with a goal of target detection.

5. Hyperspectral change detection

Change detection is one of the earliest applications for remote sensing. With the development of hyperspectral technology, many researchers attempt to apply multitemporal hyperspectral images for change detection. Recently, the method can be grouped into four classes: anomaly change detection, model-based change detection, image transformation, and others.

5.1. Anomaly change detection

The difference image is obtained by subtracting one image to another image, pixel by pixel, and band by band. Since the illumination and environmental parameters generally change between the observations, the general changes exist everywhere in the whole difference image. Comparatively, the small target change is anomalous. Thus, anomaly detection algorithm can be applied to detect small target change.

Figure 5 shows the basic diagram of anomaly change detection. First the “predictor” is applied to suppress the general changes. And then subtraction is performed to obtain residual. Finally, the “detector” is used to detect anomaly change. Eismann et al. (98) discussed research on the performance of several anomaly change detection methods with respect to the presence of diurnal and seasonal variances. Theiler et al. (99) derived a class of algorithms by modeling the data with elliptically contoured distribution. Theiler et al. (100) applied local coregistration adjustment to improve the robustness to misregistration. The “detector” can be updated to improve the performance in anomaly changes (101).

5.2. Model-based change detection

Hyperspectral imagery contains abundant spectral information. Generally, changes in the imagery typically correspond to changes in material reflectance. Meola et al. (102) used a physical model describing the sensor-reaching radiance and formulated change detection as a statistical hypothesis.

Figure 6 shows the diagram of physical sensor-reaching radiance model with shadow terms (102). This physical model can be formulated as follows:

$$L[m] = \rho[m] \odot (\bar{s}[m] \tau \odot L_s + \beta[m] \tau \odot L_d) + L_p + n[m]$$

(17)

The model is applied to simulate the radiance value of each image. With multitemporal images of little change, the material spectra are reconstructed and the spectral changes are detected. In this model, shadow coefficients are included in the physical model to reduce false alarms caused by the shadow. Recently, Meola et al.
(103) also applied model-based approach to airborne VNIR/SWIR hyperspectral imagery to detect real material spectral changes.

5.3. Image transformation

A hyperspectral image can be regarded as a high-dimensional data-set. Generally, dimension reduction can be first performed to get low-dimensional dataset. Then, change detection methods for multispectral image can be applied. Marpu et al. proposed a hyperspectral change detection approach with feature reduction and iteratively reweighted multivariate alteration detection (IRMAD) (104). It is feasible because MAD transformation is invariant to affine transformation. But, these kinds of methods do not make use of the high-dimensional information. Plentiful change information is included in the change vectors of hyperspectral difference images. Wu et al. (105) proposed a straightforward method which is based on the ICA on hyperspectral difference image. In this method the linear spectral mixture is considered. The hyperspectral difference image can also be considered as a mixture of the endmembers from the multitemporal images and the difference abundance. Therefore, ICA is performed to separate the independent endmembers from the observations. In the abundance image, the abundance values of each endmember indicate the changes in each land-cover. The ICA change detection method is shown in Figure 7.

Du et al. used kurtosis and distinct change vector to extract change information from hyperspectral difference image directly (106). Nielsen (107) wanted to add regularization in IRMAD to be applied in hyperspectral images.

5.4. Others

Besides the difference image, the compound image, obtained by stacking the multitemporal hyperspectral images into a higher dimensional data-set, can also be applied to detect changes. Wu et al. (108) applied a tar-

Figure 7. The diagram of ICA change detection.

Figure 8. The diagram of targeted change detection.
get detection algorithm on the stacked hyperspectral image to detect targeted change. In the stacked data-set, every change type has its special stacked spectral signature combined by the spectral signatures of materials before and after change. In the targeted change problem, the objective is to detect one type of change and does not consider other changes. So the target detection algorithm can be performed on the stacked image to solve the targeted change problem. Experiments indicate that it is effective, straightforward and easy to apply. A diagram is shown in Figure 8.

Advanced theories and algorithms are used in hyperspectral change detection. Nielsen et al. (109) tried to apply sparse principle component analysis in hyperspectral change detection. Du (110) applied tensor factorization on stacked multitemporal hyperspectral images to keep the spatial, spectral, and temporal structures in the original images. Liu et al. (111) used hierarchical spectral analysis and built different levels of a data-set with different spectral resolutions. Also, other advanced methods can be applied in this study area and a developing direction for hyperspectral change detection research.

6. Conclusions
Combining spectral and spatial information has been an issue addressed since the development of multiple spectral remotely sensed images. Although the spectral features revealed by multiple spectral images are actually coarse. Spectral imaging in a real sense achieves this aim by the spectral splitting technology. Furthermore, with dramatically increasing spatial and spectral resolution, HSI covers so large a quantity of information that it depicts the land-cover much more powerfully than any other remote sensing technology. From the four main aspects of hyperspectral images processing, both the most classic and advanced work were introduced. From this work, directions in current research trends are revealed: (1) Finding the means to extract optimal samples for hyperspectral classification will further enhance classification accuracy. More robust classifiers with fused spectra-spatial features will be the research focus. (2) Solving the nonlinear spectral model-based spectral unmixing problem when the linear spectral model no longer performs in a complex scene. Developing new nonlinear models with a more confidential physical meaning would conquer the current obstacles in the linear spectral model. Further research on flexible endmembers processing into the geometric model to express the hyperspectral data-set may be of essence. (3) As to the target detection from hyperspectral images, means to properly express both the backgrounds and the mixed target pixels comprise the key to construct promising new detectors. (4) Hyperspectral change detection may be influenced by radiance correction and geometric distortion. Furthermore, full exploitation of spectral information in minor changes mapping is another important issue that needs further research.

Acknowledgments
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 Grant 61102128, HuBei Province Natural Science Foundation under Grant No. 2011CDB455, China’s Post-doctoral Science Foundation under 211–180,788, and the Fundamental Research Funds for the Central Universities under 211-274633.

Notes on contributors
Zhang Liangpei is currently a “Chang-Jiang Scholar” chair professor appointed by the Ministry of Education, China. He has published more than 200 research papers and five patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Du Bo is currently a lecturer with the School of Computer, Wuhan University, Wuhan, China. His major research interests include pattern recognition, hyperspectral image processing, and signal processing.

References
(1) Landgrebe, D.A. Signal Theory Methods in Multispectral Remote Sensing; Wiley-Interscience: New York, NY, 2003.
(2) Heinz, D.C.; Chang, C.-I. Fully Constrained Least Squares Linear Mixture Analysis for Material Quantification in Hyperspectral Imagery. IEEE Trans. Geosci. Remote Sens. 2001, 39 (3), 529–545.
(3) Keshava, N.; Mustard, J.F. Spectral Unmixing. IEEE Signal Process. Mag. 2002, 19 (1), 44–57.
(4) Broadwater, J.; Chellappa, R.; Banerjee, A.; Burlina, P. Kernel Fully Constrained Least Squares Abundance Estimates. Geoscience and Remote Sensing Symposium, IGARSS; IEEE: Barcelona, 2007; pp. 4041–4044.
(5) Landgrebe, D. Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data. Inform. Process. Remote Sens. 1999, 82 (3), C38.
(6) Mika, S.; Ratsch, G.; Weston, J.; Scholkopf, B.; Mullers, K.R. Fisher Discriminant Analysis with kernels; IEEE: Madison, WI, 1999; pp. 41–48.
(7) Kuo, B.C.; Landgrebe, D.A. Nonparametric Weighted Feature Extraction for Classification. IEEE Trans. Geosci. Remote Sens. 2004, 42, 1096–1105.
(8) Huang, H.Y.; Kuo, B.C. Double Nearest Proportion Feature Extraction for Hyperspectral-Image Classification. IEEE Trans. Geosci. Remote Sens. 2010, 48, 4034–4046.
(9) Zhang, T.; Tao, D.; Yang, J. Discriminative Locality Alignment. European Conference on Computer Vision, Marseille, Oct 12–18 2008; pp. 725–738.
(10) He, X.; Niyogi, P. Locality Preserving Projections (l pp). Proceedings of the NIPS, Advances in Neural Information Processing Systems: Saul, L., Weiss, Y., Bottou, L., Eds; MIT Press: Vancouver, 2004; p. 103.
(11) Webb, A.R.; Cepsey, K.D.; Cawley, G. Statistical Pattern Recognition; Wiley: Chichester, 2011.
(12) Siedlecki, W.; Sklansky, J.A. Note on Genetic Algorithms for Large-Scale Feature Selection. Pattern Recogn. Lett. 1999, 10, 335–347.
(13) Zhong, Y.; Zhang, L. A New Fuzzy Clustering Algorithm Based on Clonal Selection for Land Cover Classification. Math. Probl. Eng. 2011, 1–24. doi: 10.1155/2011/708459.
(14) Gualtieri, J.A.; Crompt, R.F. Support Vector Machines for Hyperspectral Remote Sensing Classification, 27th AIPR Workshop: Advances in Computer-Assisted Recognition; Citeeseer, 1999; Merickso Robert J. Ed, Vol. 3584, pp. 221–232.

(15) Bruzzone, L.; Chi, M.; Marconcini, M.A. Novel Transductive SVM for Semisupervised Classification of Remote-Sensing Images. IEEE Trans. Geosci. Remote Sens. 2006, 44, 3363–3373.

(16) Bovolo, F.; Bruzzone, L.; Marconcini M.; Persello, C. A Novel Context-Sensitive SVM for Classification of Remote Sensing Images; Technical Report. 2006, DIT-06-040.

(17) Camps-valls, G.; Gomez-chova, L.; Mu oz-mar, J.; Vila-frances, J.; Calpe-Maravilla. J. Composite Kernels for Hyperspectral Image Classification. IEEE Geosci. Remote Sens. Lett. 2006, 3, 93–97.

(18) Plaza, A.; Benediktsson, J.A.; Bordman, J.W.; Brazilian; J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A. Recent Advances in Techniques for Hyperspectral Image Processing. Remote Sens. Environ. 2009, 113, S110–S1122.

(19) Böhning, D. Multinomial Logistic Regression Algorithm. Ann. Inst. Statist. Math. 1992, 44, 197–200.

(20) Krishnapuram, B.; Wiliams, K.; Xue Y.; Hartemink, A.; Carin, L.; Figueiredo, M.A.T. On Semi-Supervised Classification. Advances in Neural Information Processing Systems (NIPS). MIT press, 2005, 48(17).

(21) Pal, M. Artificial Immune-Based Supervised Classifier for Land-Cover Classification. Int. J. Remote Sens. 2008, 29, 2273–2291.

(22) Zhong, Y.; Zhang, L.; Huang, B.; Li, P. An Unsupervised Artificial Immune Classifier for Multi/Hyperspectral Remote Sensing Imagery. IEEE Trans. Geosci. Remote Sens. 2006, 44, 420–431.

(23) Zhang, L.; Zhong, Y.; Huang, B.; Li, P. A Resource Limited Artificial Immune System Algorithm for Supervised Classification of Multi/Hyper-Spectral Remote Sensing Imagery. Int. J. Remote Sens. 2007, 28, 1665–1686.

(24) Jiao, H.; Zhong, Y.; Zhang, L. Artificial DNA Computing-Based Spectral Encoding and Matching Algorithm for Hyperspectral Remote Sensing Data. IEEE Trans. Geosci. Remote Sens. 2012, 49, 1–20.

(25) Dempster, A.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. J. Roy. Statist. Soc. Ser. B (Methodological) 1977, 39, 1–38.

(26) Vapnik, V.N. An Overview of Statistical Learning Theory. IEEE Trans. Neural Networks 1999, 10, 988–999.

(27) Chung, F.K. Spectral Graph Theory. CBMS Regional Conference Series in Mathematics; American Mathematical Society, 1997; Vol. 92.

(28) Schohn, G.; Cohn, D. Less is More: Active Learning with Support Vector Machines. Proceedings of the Seventh International Conference on Machine Learning (ICML), San Francisco, CA; Citeeseer, 2000; pp. 839–846.

(29) Kullback, S.; Leibler, R.A. On Information and Sufficiency. Ann. Math. Stat. 1951, 22, 79–86.

(30) Freund, Y.; Seung, H.S.; Shamir, E.; Tishby, N. Selective Sampling Using the Query by Committee Algorithm. Mach. Learn. 1997, 28, 133–168.

(31) Corresponding, A.; Hirsch, J.; Weber, C. The Utility of Texture Analysis to Improve Per-Pixel Classification for High to Very High Spatial Resolution Imagery. Int. J. Remote Sens. 2005, 26, 733–745.

(32) Zhang, L.; Du, B.; Zhong, Y. Hybrid Detectors Based on Selective Endmembers. IEEE Trans. Geosci. Remote Sens. 2010, 48 (6), 2633–2646.

(33) Zhang, L.; Zhang, L.; Tao, D.; Huang, X. On Combining Multiple Features for Hyperspectral Remote Sensing Image Classification. IEEE Trans. Geosci. Remote Sens. 2012, 50, 879–893.

(34) Zhang, L.; Tao, D.; Huang, X.A. Multifeature Tensor for Remote-Sensing Target Recognition. IEEE Geosci. Remote Sens. Lett. 2011, 8, 374–378.

(35) Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J. Spectral-Spatial Classification of Hyperspectral Imagery Based on Partition Clustering Technique. IEEE Trans. Geosci. Remote Sens. 2009, 47, 2973–2987.

(36) Plaza, A.; Du, Q.; Bioucas-Dias, J.; Jia, X.; Kruse, F. Foreword to the Special Issue on Spectral Unmixing of Remotely Sensed Data. IEEE Trans. Geosci. Remote Sens. Remote Sens. 2011, 49 (11), 4103–4110.

(37) Bioucas-Dias, M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 2012, 5, 354–379.

(38) Somers, B.; Cools, K.; Delalieux, S.; Stuckens, J.; Zande, D.V.; Verstraeten, W.W.; Coppin, P. Nonlinear Hyperspectral Mixture Analysis for Tree Cover Estimates in Orchards. Remote Sens. Environ. 2009, 113 (6), 1183–1193.

(39) Nascimento, J.M.P.; Bioucas-Dias, J.M. Nonlinear Mixture Model for Hyperspectral Unmixing, Proc. SPIE 2009, 747701.1–747701.8.

(40) Licciardi, G.A.; Del Frate, F. Pixel Unmixing in Hyperspectral Data by Means of Neural Networks. IEEE Trans. Geosci. Remote Sens. 2011, 49 (11), 4163–4172.

(41) Plaza, J.; Plaza, A.; Perez, R.; Martinez. P. On the Use of Small Training Sets for Neural Network-Based Characterization of Mixed Pixels in Remotely Sensed Hyperspectral Images. Pattern Recognit. 2009, 42 (11), 3032–3045.

(42) Plaza, J.; Plaza, A. Spectral Mixture Analysis of Hyperspectral Scenes Using Intelligently Selected Training Samples. IEEE Geosci. Remote Sens. Lett. 2010, 7 (2), 371–375.

(43) Jutten, C.; Karhunen, J. Advances in Nonlinear Blind Source Separation, 4th International Symposium Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan, April, 2003; pp. 245–256.

(44) Babaie-Zadeh, M.; Jutten, C.; Nayebi, K. Separating Convolutative Post Non-linear Mixtures. Proceedings of the 3rd Workshop on Independent Component Analysis and Signal Separation (ICA2001), San Diego, CA, 2001; pp. 138–143.

(45) Halmiri, A.; Altman, Y.; Dobigeon, N.; Tourneret, J.-Y. Nonlinear Unmixing of Hyperspectral Images Using a Generalized Bilinear Model. IEEE Trans. Geosci. Remote Sens. 2011, 49 (11), 4153–4162.

(46) Boardman, J.; Kruse, F.; Green, R. Mapping Target Signatures via Partial Unmixing of AVIRIS Data. Proceedings of the Summaries JPL Airborne Earth Science Workshop, Pasadena, CA, January 23, 1995; pp. 23–26.

(47) Chang, C.-I.; Plaza, A. A Fast Iterative Algorithm for Implementation of Pixel Purity Index. IEEE Geosci. Remote Sens. Lett. 2006, 3 (1), 63–67.

(48) Chang, C.-I.; Wu, C.-C.; Chen, H.-M. Random Pixel Purity Index. IEEE. Geosci. Remote Sens. Lett. 2010, 7 (2), 324–328.
(49) Winter, M.E. Fast autonomous Spectral End-Member Determination in Hyperspectral Data. Proceedings of the Thirteenth International Conference on Applied Geologic Remote Sensing, Vancouver, BC, Canada, 1999; pp. 337–344.

(50) Nascimento, J.M.P.; Dias, J.M.B. Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* 2005, 43 (4), 898–910.

(51) Wu, C.-C.; Lo, C.-S.; Chang, C.-I. Improved Process for Use of a Simplex Growing Algorithm for Endmember Extraction. *IEEE Geosci. Remote Sens. Lett.* 2009, 6 (3), 523–527.

(52) Ifarraguerri, A.; Chang, C.-I. Multispectral and Hyperspectral Image Analysis with Convex Cones. *IEEE Trans. Geosci. Remote Sens.* 1999, 37 (2), 756–770.

(53) Gruninger, J.A.; Ratkowsk, J.I.; Hoke, M.L. The Sequential Maximum Angle Convex Cone (SMACC) Endmember Model. *Proceedings SPIE. Algorithms for Multispectral and Hyper-spectral and Ultraspectral Imagery*, Vol. 5425-1, Orlando, FL, Apr, 2004.

(54) Neville, R.A.; Staenz, K.; Szeredi, T.; Lefebvre, J.; Haufl, P. Automatic Endmember Extraction from Hyperspectral Data for Mineral Exploration. *Proceedings of Fourth International Airborne Remote Sensing Conference and Exhibition/21st Canadian Symposium on Remote Sensing*, Ottawa, Canada, 1999; pp. 21–24.

(55) Harsanyi, J.C.; Chang, C.I. Detection of Low Probability Subpixel Targets in Hyperspectral Image Sequences with Unknown Backgrounds. *IEEE Trans. Geosci. Remote Sens.* 1994, 32, 779–785.

(56) Chang, C.-I. Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis. *IEEE Trans. Geosci. Remote Sens.* 2008, 43 (3), 502–518.

(57) Plaza, A.; Martinez, P.; Perez, R.; Plaza, J. Spatial/Spectral Endmember Extraction by Multidimensional Morphological Operations. *IEEE Trans. Geosci. Remote Sens.* 2002, 40 (9), 756–770.

(58) Craig, M. Minimum-Volume Transforms for Remotely Sensed Data. *IEEE Trans. Geosci. Remote Sens.* 1994, 32 (3), 542–552.

(59) Li, J.; Bioucas-Dias, J. Minimum Volume Simplex Analysis: A Fast Algorithm to Unmix Hyperspectral Data. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Vol. 4, Boston, MA, 2008; pp. 2369–2371.

(60) Chan, T.-H.; Chi, C.-Y.; Huang, Y.-M.; Ma, W.-K. A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing. *IEEE Trans. Signal Process* 2009, 57 (11), 4418–4432.

(61) Parra, L.C.; Sajada, P.; Du, S. Recovery of Constituent Spectra Using Non-Negative Matrix Factorization. *Proc. SPIE 2003*, 5207 (1), 321–331.

(62) Miao, L.; Qi, H. Endmember Extraction from Highly Mixed Data Using Minimum Volume Constrained Non-negative Matrix Factorization. *IEEE Trans. Geosci. Remote Sens.* 2007, 45 (3), 765–777.

(63) Bioucas-Dias, M. A Variable Splitting Augmented Lagrangian Approach to Linear Spectral Unmixing. *Proceedings of the 1st WHISPERS*, Aug, 2009, pp. 1–4.

(64) Li, H.; Zhang, L. A Hybrid Automatic Endmember Extraction Algorithm Based on a Local Window. *IEEE Trans. Geosci. Remote Sens.* 2011, 49 (11), 4223–4238.

(65) Zortea, M.; Plaza, A. Spatial Preprocessing for Endmember Extraction. *IEEE Trans. Geosci. Remote Sens.* 2009, 47 (8), 2679–2693.

(66) Moussaoui, S.; Hauksdottir, H.; Schmidt, F.; Jutten, C.; Chanussot, J.; Bric, D.; Doute, S.; Benediktsson, J.A. On the Decomposition of Mars Hyperspectral Data by ICA and Bayesian Positive Source Separation. *Neurocomputing* 2008, 71 (10–12), 2194–2208.

(67) Chang, C.-I.; Wu, C.-C.; Liu, W.; Ouyang, Y.-C. New Growing Method for Simplex-Based Endmember Extraction Algorithm. *IEEE Trans. Geosci. Remote Sens.* 2006, 44 (10), 2804–2819.

(68) Nascimento, J.M.P.; Bioucas-Dias, J.M. Dependent Component Analysis: A Hyperspectral Unmixing Algorithm. *Lect. Notes Comput. Sci.* 2007, 4478, 612–619.

(69) Lee, D.D.; Seung, H.S. Algorithms for Non-Negative Matrix Factorization. *Adv. Neural Inf. Process. Syst.* 2000, 13, 556–562.

(70) Paatero, P.; Tapper, U. Positive Matrix Factorization: A Non-Negative Factor Model with Optimal Utilization of Error. *Environmetrics* 1994, 5, 111–126.

(71) Hück, A.; Guillaume, M.; Blanc-Talon, J. Minimum Dispersion Constrained Nonnegative Matrix Factorization to Unmix Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* 2010, 48, 2590–2602.

(72) Jia, S.; Qian, Y.T. Constrained Nonnegative Matrix Factorization for Hyperspectral Unmixing. *IEEE Trans. Geosci. Remote Sens.* 2009, 47, 161–173.

(73) Liu, X.S.; Xia, W.; Wang, B.; Zhang, L.M. An Approach Based on Constrained Nonnegative Matrix Factorization to Unmix Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 757–770.

(74) Wang, N.; Du, B.; Zhang, L. An Endmember Dissimilarity Based Nonnegative Matrix Factorization Method for Hyperspectral Unmixing. 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), IEEE, Shanghai, June 6–9, 2012, pp. 1–4.

(75) Fisher, R.A. The Use of Multiple Measures in Taxonomic Problems. *Ann. Eugenics* 1936, 7, 179–188.

(76) Farrand, W.H.; Harsanyi, J.C. Mapping the Distribution of Mine Tailings in the Coeur d’Alene River Valley, Idaho, through the Use of a Constrained Energy Minimization Technique. *Rem. Sens. Environ.* 1997, 59 (1), 64–76.

(77) Schowengerdt, R.A. Remote Sensing: Models and Methods for Image Processing; Academic: San Diego, CA, 1997.

(78) Yu, X.; Reed, I.S. Comparative Performance Analysis of Adaptive Multispectral Detectors. *IEEE Trans. Signal Process* 1993, 41 (8), 2639–2656.

(79) Kwon, H.; Nasrabadi, N.M. Kernel RX-Algorithm: A Nonlinear Anomaly Detector for Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* 2005, 43 (2), 1309–1320.

(80) Hruschka, E.R.; Ebecken, N.F.F. Rule Extraction from Neural Networks: Modified RX Algorithm. *Proceedings of International Joint Conference on Neural Networks; IEEE*: Washington, DC, 1999; Hyvarinen, A.; Karhunen, J.; Oja, E. Independent Component Analysis; Wiley: New York, NY, 2001.

(81) Du, B.; Zhang, L. Random-Selection-Based Anomaly Detector for Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* 2011, 49 (5), 1578–1589.

(82) Reed, I.S.; Yu, X. Adaptive Multiple-Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution. *IEEE Proc. Acoust. Speech, Signal Process.* 1990, 38, 1760–1770.

(83) Jiao, X.; Chang, C.-I. Kernel-Based Constrained Energy Minimization (K-CEM). Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV. *Proc. SPIE 2008*, 6966, 69661S.
(84) Wang, T.; Du, B.; Zhang, L. A Local Subspace Based Non-linear Target Detector. 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), IEEE, Shanghai, June 6–9, 2012, pp. 1–4.
(85) Zhang, L.P.; Li, D.R. Study of the Spectral Mixture Model of Soil and Vegetation in Poyang Lake Area, China. Int. J. Remote Sens. 1998, 9 (11), 2077–2084.
(86) Kelly, E.J. An Adaptive Detection Algorithm. IEEE Trans. Aerosp. Electron. Syst. 1986, 22 (1), 115–127.
(87) Kelly, E.J. Adaptive Detection in Non-Stationary Interference, Part III; Technical Report 761; Lincoln Laboratory, DTIC# ADA-185622; 1986.
(88) Robey, F.C.; Fuhrmann, D.R.; Kelly, E.I.; Nitzberg, R. A CFAR Adaptive Matched Filter Detector. IEEE Trans. Aerosp. Electron. Syst. 1992, 28 (1), 208–216.
(89) Chen, W.-S.; Reed, I.S. A New CFAR Detection Test for Radar. Dig. Signal Process. 1991, 1 (4), 198–214.
(90) Kraut, S.; Scharf, L.L. The CFAR Adaptive Sub-Space Detector is a Scale-Invariant GLRT. IEEE Trans. Signal Process. 1999, 47, 2538–2541.
(91) Kraut, S.; Scharf, L.L.; McWhorter, L.T. Adaptive Subspace Detectors. IEEE Trans. Signal Process. 2001, 49 (1), 1–16.
(92) Scharf, L.L.; Friedlander, B. Matched Sub-Space Detectors. IEEE Trans. Signal Process. 1994, 42, 2146–2157.
(93) Kay, S.M. Fundamentals of Statistical Signal Processing: Prentice Hall: Englewood Cliffs, NJ, 1998.
(94) Harsanyi, J.C.; Chang, C.-I. Hyperspectral Image Classification and Dimensionality Reduction: An Orthogonal Subspace Projection Approach. IEEE Trans. Geosci. Remote Sens. 1994, 32 (4), 779–785.
(95) Francois, V.; Olivier, B.; Cédric, R. Matched Subspace Detection with Hypothesis Dependent Noise Power. IEEE Trans. Signal Process. 2008, 56 (11), 5713–5718.
(96) Du, Q.; Wei, W.; May, D.; Younan, N.H. Noise-Adjusted Principal Component Analysis for Buried Radioactive Target Detection and Classification. IEEE Trans. Nucl. Sci. 2010, 57 (6), 3760–3767.
(97) Du, Q.; Fowler, J.E.; Ma, B. Random-Projection-Based Dimensionality Reduction and Decision Fusion for Hyperspectral Target Detection. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, Vancouver, July 24–29, 2011, pp. 1790–1793.
(98) Eismann, M.T.; Meola, J.; Hardie, R.C. Hyperspectral Change Detection in the Presence of Diurnal and Seasonal Variations. IEEE Trans. Geosci. Remote Sens. 2008, 46 (1), 237–249.
(99) Theiler, J.; Scovel, C.; Wohlbeg, B.; Foy, B.R. Elliptically Contoured Distributions for Anomalous Change Detection in Hyperspectral Imagery. IEEE Geosci. Remote Sens. Lett. 2010, 7 (2), 271–275.
(100) Theiler, J.; Wohlbeg, B. Local Coregistration Adjustment for Anomalous Change Detection. IEEE Trans. Geosci. Remote Sens. 2012, 99, 1–10.
(101) Brisebarre, G.; Guillaume, M.; Huck, A.; Denise, L. Hyperspectral Change Detection by Direction Pursuit. 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS); Shanghai, 2012.
(102) Meola, J.; Eismann, M.T.; Moses, R.L.; Ash, J.N. Detecting Changes in Hyperspectral Imagery Using a Model-Based Approach. IEEE Trans. Geosci. Remote Sens. 2011, 49 (7), 2647–2661.
(103) Meola, J.; Eismann, M.T.; Moses, R.L.; Ash, J.N. Application of Model-Based Change Detection to Airborne VNIR/SWIR Hyperspectral Imagery. IEEE Trans. Geosci. Remote Sens. 2012, 99, 1–14.
(104) Marpu, P.; Gamba, P.; Benediktsson, J.A. Hyperspectral Change Detection with IR-MAD and Initial Change Mask. 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS); Shanghai, 2012.
(105) Wu, C.; Du, B.; Zhang, L. Hyperspectral Change Detection Based on Independent Component Analysis. J. Remote Sens. 2012, 3 (16), 545–561.
(106) Du, Q.; Younan, N.; King, R. Change Analysis for Hyperspectral Imagery. 2007 International Workshop on the Analysis of Multi-temporal Remote Sensing Images (MultiTemp 2007); Provinciehuis Leuven, Belgium, 2007.
(107) Allan, A. Nielsen, The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. IEEE Trans. Image Process. 2007, 16 (2), 463–478.
(108) Wu, C.; Zhang, L.; Du, B. Targeted Change Detection for Stacked Multi-temporal Hyperspectral Image. 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS); Shanghai, 2012.
(109) Nielsen, A.A.; Larsen, R.; Vestergaard, J.S. Sparse Principal Component Analysis in Hyperspectral Change Detection. Image and Signal Processing for Remote Sensing XVII; Prague, MN, October 25, 2011, Vol. 8180, pp. 81800S.1–81800S.6.
(110) Du, Q.A. New Method for Change Analysis of Multi-temporal Hyperspectral Images. 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS); Shanghai, 2012.
(111) Liu, S.; Bruzzone, L.; Bovolo, F.; Du, P. Unsupervised Hierarchical Spectral Analysis for Change Detection in Hyperspectral Images. 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS); Shanghai, 2012.