Neurocomputing 108 (2013) 103–110

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Rare signal component extraction based on kernel methods for anomaly detection in hyperspectral imagery

Yanfeng Gu [a,*], Lin Zhang [b]

[a] School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
[b] Division of Adults and Graduates Studies, Eastern Nazarence College, MA, USA

A R T I C L E I N F O

Article history:
Received 6 September 2012
Received in revised form
25 October 2012
Accepted 12 November 2012
Communicated by Lixian Zhang
Available online 3 January 2013

Keywords:
Hyperspectral imagery
Anomaly detection
Independent component analysis
Kernel methods (KM)
High-order statistics
Singular value decomposition (SVD)

A B S T R A C T

Anomaly detection is one of hot research topics in hyperspectral remote sensing. For this task, RX detector (RXD) is a benchmark method. Unfortunately, Gaussian distribution assumption adopted by RXD cannot be well satisfied in hyperspectral images due to high dimensionality of data and complicated correlation between spectral bands. In this paper, we address this problem and propose an algorithm called rare signal component extraction (RSCE), aiming at finding a subspace where the Gaussian assumption is well obeyed and improving detection performance of RXD. RSCE algorithm first utilizes kernel singular value decomposition (KSVD) to construct a kernel-based whitening operator, and then, carries out kernel-based whitening on hyperspectral data. After that, RSCE algorithm is to extract and determine a singular signal subspace by means of independent component analysis in reproducing kernel Hilbert space (RKHS) space and singularity measure. Numerical experiments were conducted on two real hyperspectral datasets. The experimental results show that the proposed RSCE algorithm greatly improves the detection performance of RXD and outperforms other state-of-the-art methods.

1. Introduction

Hyperspectral imaging provides fine spectral resolution for good description of ground materials. By means of this merit of hyperspectral imaging, anomaly detection, as a class of crucial techniques promoting rapid applications of hyperspectral images, has become a hot topic recently. In most natural scenes, research of rare and small targets is a relevant issue as they usually are man-made. In hyperspectral imagery, anomaly detection aims at finding rare pixels having different spectral signatures from the surrounding background clutters. Main characteristic of anomaly detection consists in that it carries out detection without *priori spectra* about target materials. In addition, it is generally assumed that the rare anomalies are viewed as be embedded within background clutter with a very low signal-to-clutter ratio (SCR), namely, the anomalies are mixed with background clutter even at subpixel-level ([8–10,13–16,18,20,21]).

Recently, RX detector (RXD), which was proposed by Reed and Yu [19], has been considered as a benchmark method for anomaly detection in multispectral and hyperspectral imaging. RXD is derived from the generalized-likelihood ratio test (GLRT) and an adaptive detector with constant false alarm rate (CFAR) what we would pursue in anomaly detection. In RXD, it is assumed that both of the spectra of the anomalous targets and the covariance of the background clutters are unknown. Furthermore, the background clutter is assumed as being obeyed to unimodal Gaussian distribution. Thus, Mahalanobis distance between the pixel to be detected and the background mean vector is compared to a threshold to detect anomalies. At the same time, the mean vector and covariance matrix of the background clutter are estimated by using those pixels surrounding the pixel. After RXD was proposed, some modified RXD-based detectors were investigated one after the other. For example, Chang et al. proposed a number of modified RXD, such as normalized RXD (NRXD), correlation matrix based NRXD (CNRXD) [3].

Actually, the anomaly detection of hyperspectral images can be thought as a design of filtering anomalies of interest from discrete spatial–spectral signals. There are some important factors we should consider like rationality and stability of model, and nonlinear physical phenomena ([24–26]). When one uses these detectors to deal with hyperspectral data, there are two main factors which limit the performance of RXD and the modified RXD, i.e., rationality of the Gaussian assumption and high dimensionality of the data. First, the assumption of the background clutter being obeyed to unimodal Gaussian distribution is better satisfied, the RXD is more effective. However, the unimodal Gaussian assumption cannot be satisfied generally in real case because the background clutters are mixed each with other or with anomalies at pixel or subpixel scale. This fact leads to poor false alarm performance [14]. If there are multiple classes of

* Corresponding author. Tel.: +86 451 86403020; fax: +86 451 86413583.
E-mail address: guyf@hit.edu.cn (Y. Gu).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.11.016

background clutters in test region, it is more difficult to properly represent the underlying distribution by means of the unimodal Gaussian distribution [2]. To overcome the limitation of unimodal Gaussian distribution, researchers [22] proposed linear mixture of Gaussian models to more properly characterize nonhomogeneous multicomponent scenes. The parameters of the Gaussian mixture model can be estimated using the stochastic expectation maximization (SEM) approach [17]. However, it is badly limited by the requirement for knowing the number of Gaussian models. Second, the high dimensionality of hyperspectral images is another factor badly limiting the performance of the conventional RXD. This factor makes covariance estimation of the background clutter more difficult and unreliable. Principal component analysis (PCA), which is one of the most conventional feature extraction and dimensionality reduction algorithms in many pattern classification applications, is able to pre-process the hyperspectral images and is helpful to improve RXD. Furthermore, several independent component analysis (ICA)-based algorithms have been proposed in Ref. [7]. Compared with PCA, the merit of ICA stems from the use of high-order statistics. By means of ICA, those algorithms built anomaly detection map combining the thresholded independent components (ICs) with high kurtosis. Unfortunately, for hyperspectral images, the high dimensionality indicates complicated correlation between spectral bands [4]. As a result, those linear algorithms, like PCA and ICA, are not always satisfied to detect anomalies in hyperspectral images.

Along with the wide applications of kernel methods (KM), nonlinear detection algorithms have also been developed in hyperspectral remote sensing. In the kernel-based detectors, high-order correlation between spectral bands is implicitly exploited by using kernel function. In this way, one is able to capture the high-order correlation structures in input data set and get nonlinear decision boundaries. Kwon and Nasrabidi proposed a nonlinear kernelized version of the conventional RXD by extending the RXD from the original input space to a high-dimensional feature space on the whole dimensionality of input data [12]. This detection algorithm, called Kernel RX, demonstrated superior performance over the conventional RXD on real hyperspectral data. Kernel RX algorithm is a whole kernelized algorithm in feature space. As a result, the kernelization operation improves separability of anomalies of interest from background clutters, but a different role of the kernelized components in feature space is not distinguished. Banerjee et al. proposed a method for anomaly detection in hyperspectral images based on support vector data description (SVDD), which is a kernel method for modeling the support of a distribution [2]. The SVDD-based detection is a nonparametric algorithm, which possesses several attractive merits, such as sparsity, good generalization, and use of kernel. By means of the use of kernel for modeling the support of the nontrivial multimodal distributions, the SVDD-based algorithm kept away from obstacle brought by mathematically unimodal Gaussian assumption. In fact, The SVDD-based algorithm models the probability density function (PDF) though one class support vector machine (SVM). For this reason, there is no surely guarantee that all local samples come from same single background clutter rather than multiple background clutters. As a consequence, the detection performance of SVDD-based detection will have a decline while the number of training samples is inadequate or multiple background clutters surround the pixel under test. Recently, Gu et al. [7] have proposed a selective kernel principal component analysis (SKPCA)-based feature extraction algorithm for anomaly detection in hyperspectral images, in which KPCA is used to capture the higher order correlation structure of input data, and local kurtosis is used as a measure to select the most singular component for anomaly detection. In fact, the data variance is weakly affected by the rare anomalous signals and is mainly controlled by the background clutters. Therefore, the techniques, like PCA and KPCA, estimate the signal subspace addressing mostly the background clutters and ignoring the presence of rare target pixels [1].

In this paper, we investigate the anomaly detection in hyperspectral images and present a rare signal component extraction algorithm called RSCE to improve the detection performance of RXD. In the RSCE algorithm, kernel-based singular value decomposition (KSVD) is utilized to build kernel subspace for background clutters and capture the *nonlinearly higher order correlation* of input hyperspectral images. Then, a kernel whitening operator, which is derived from the kernel subspace, is performed on the original data. After that, we are able to perform ICA to get the kernel-based nonlinear ICs. Finally, the rare signal component (RSC) can be found from the kernel-based ICs by means of singularity measure.

For all these descriptions, the main works of this paper can be determined in the following: (1) the *conformation* of KSVD-based whitening operator, which enables one to effectively reduce the data dimensionality and rapidly capture *nonlinearly higher order correlation* of the input hyperspectral dataset; (2) the *separable extraction* for kernel-based ICs, which is realized by KSVD and linear FastICA and overcome the computational limitation from directly kernel-independent component analysis (KICA); (3) the *determination* of the RSC by the singularity measure from kernel ICs, which makes the unimodal Gaussian distribution rationalized and is more suitable for rare signal detection than kernel-based principal components (KPC).

The rest of this paper is organized as follows. In Section 2, we provide a brief introduction to RXD, which is used to implement the final detection in the RSCE algorithm. In Section 3, we describe information about the RSCE algorithm in detail, including key procedures such as KSVD-based whitening, ICs extraction, and singularity measures for RSC. In addition, realization outline of the RSCE algorithm is given in this section. Experiments and result analyses based on two real hyperspectral datasets are shown in Section 4, and conclusions are drawn in Section 5.

2. *RX* algorithm

In the conventional RXD, two competing hypotheses to be distinguished are given by

$$\mathbf{H}_0 : \mathbf{z} = \mathbf{n} \text{ (target absent)}$$

$$\mathbf{H}_1 : \mathbf{z} = \alpha\mathbf{t} + \mathbf{n} \text{ (target present)} \quad (1)$$

where $\mathbf{z}$ is an input spectral signal to be detected and consists of l spectral bands denoted by $\mathbf{z} = [z_1 z_2 \cdots z_l]^T$, $\alpha > 0$, $\mathbf{n}$ is a vector that represents process of the background clutters and noise, and $\mathbf{t}$ is the spectral signature of the anomalous target given by $\mathbf{t} = [t_1 t_2 \cdots t_l]^T$.

In RXD, the target signature $\mathbf{t}$ and background covariance Γ_B are assumed to be unknown. The model assumes that the data arise from two normal probability density functions with the same covariance matrix but different means. Under $\mathbf{H}_0$, the data (background clutters) are modeled as $N(0,\Gamma_B)$, and under $\mathbf{H}_1$ the data are modeled as $N(\mathbf{t},\Gamma_B)$. In most cases, local unimodal Gaussian distribution is adopted in the conventional RXD. Generally, estimating the mean and covariance of background clutter is conducted within a double concentric window. The inner window is used to protect the probably anomalous information, whose size should be equal to the size of the target of interest in the test scene. In contrast to the inner window, the spectral samples in the outer window are used to estimate the mean $\hat{\boldsymbol{\mu}}_B$ and covariance $\hat{\Gamma}_B$ of background clutter. Based on the description mentioned above, expression of the conventional RXD is given by

$$RX(\mathbf{z}) = \left(\mathbf{z} - \hat{\boldsymbol{\mu}}_B\right)^T \left(\hat{\Gamma}_B\right)^{-1} \left(\mathbf{z} - \hat{\boldsymbol{\mu}}_B\right) \quad (2)$$

where $\hat{\Gamma}_B$ is the background covariance matrix estimated from the reference data of the background clutters in the outer window, and $\hat{\boldsymbol{\mu}}_B$ is the estimated sample mean of the background clutters. Let η be a threshold, if $RX(\mathbf{z}) \geq \eta$, then assumption with target present is adopted, otherwise assumption with target absent.