Pattern Recognition 47 (2014) 344–358

Contents lists available at ScienceDirect

Pattern Recognition

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

Target detection based on a dynamic subspace

Bo Du [a,*], Liangpei Zhang [b]

[a] School of Computer Science, Wuhan University, PR China
[b] The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, PR China

ARTICLE INFO

Article history:
Received 26 February 2013
Received in revised form
28 May 2013
Accepted 5 July 2013
Available online 24 July 2013

Keywords:
Target detection
Hyperspectral images
Subspace learning

ABSTRACT

For hyperspectral target detection, it is usually the case that only part of the targets pixels can be used as target signatures, so can we use them to construct the most proper background subspace for detecting all the probable targets? In this paper, a dynamic subspace detection (DSD) method which establishes a multiple detection framework is proposed. In each detection procedure, blocks of pixels are calculated by the random selection and the succeeding detection performance distribution analysis. Manifold analysis is further used to eliminate the probable anomalous pixels and purify the subspace datasets, and the remaining pixels construct the subspace for each detection procedure. The final detection results are then enhanced by the fusion of target occurrence frequencies in all the detection procedures. Experiments with both synthetic and real hyperspectral images (HSI) evaluate the validation of our proposed DSD method by using several different state-of-the-art methods as the basic detectors. With several other single detectors and multiple detection methods as comparable methods, improved receiver operating characteristic curves and better separability between targets and backgrounds by the DSD methods are illustrated. The DSD methods also perform well with the covariance-based detectors, showing their efficiency in selecting covariance information for detection.

1. Introduction

Targets in the remote sensing domain refer to ground objects of special interest [1–3]. For example, vehicles on a road comprise the targets when the extraction of their accurate positions is the main task. The spectral differences between target and non-target backgrounds is the foremost feature in target detection by hyperspectral remote sensing imagery, which makes it a two-class classification problem. However, targets in remote sensing images usually occupy only a small fraction of the whole image, with each having a limited size. As a result, the minimization strategy for the misclassification error cannot be used in target detection, otherwise all the targets would be labeled as background [2,3].

Remote sensing target detection methods contrarily focus on maximizing the probability of detection with a certain constant false alarm rate, which originates from the signal estimation and detection theory in the communications field [4–6]. Thus, the target pixel is considered as the signal of interest to be detected. In this way, the problem is transformed into a signal processing procedure. This is the main idea behind the signal detection based methods, including the finite impulse response filter, likelihood ratio test, hypothesis testing, and so on [1,2,7–9]. Other methods exploit the linear spectral mixture [10,11], assuming that each pixel consists of different endmembers. "Endmember" refers to a pure pixel of the land object in the image and presents a unique spectral feature. The difference between different endmembers is the key to interpreting hyperspectral images. This approach is widely used in spectral unmixing from hyperspectral images [12, 13]. However, the distinction is that unmixing decomposes the scene into all the constituent materials in their proportions, whereas target detection should give a more or less binary indication of the presence of a single material or class of interest.

In spite of their different theoretic origins, many detectors can be considered as subspace-based detectors as the information about targets and backgrounds is reserved in the subspace. Two important ways of constructing subspaces shall now be considered. One way is to use the different endmembers' spectra as a basis to compose the target subspace and background subspace, respectively. In other words, the composite units in the spectral mixture model are used. The other approach is to selectively choose bands from the image as a subset to make up the subspace, which is actually a band selection procedure. However, both approaches have the following drawbacks. They both use the same subspace background construction method for different imagery; that is, they choose the whole dataset as background. The resulting subspace will cause a drift to the inverse direction of the main direction of the subspace in the feature space [15], or a contaminated background subspace [16–18]. Research has been undertaken to prove that different pixels in

* Corresponding author. Tel.: +86 13871461059.
E-mail address: remoteking@whu.edu.cn (B. Du).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.07.005

different regions usually contain different backgrounds [19], due to the discontinuous and inhomogeneous nature of real-world landscape composition.

Other researchers have reported that the spectrally similar pixels are actually on the same patch in the whole manifold feature space [20–25], which is another clue as to the local subspace construction. Furthermore, a locally constructed subspace provides an efficient way to avoid the limitation of the data distribution [26–28]. A typical example is the Gaussian distribution assumption, which is the one most widely used in most detection methods, although, in many cases, the assumption is not reasonable [29–31, 33]. In fact, with the local subspace, only the linear relationship between the pixel under observation and the neighborhood pixels is used. Therefore, no assumption about the data distribution is taken into consideration.

Above all, targets pixels usually reside in parts of the feature space, so a rigid global subspace may not be able to model different pixels very well. The corresponding background subspace should be elaborately constructed from certain pixels of the image dataset. The problem then lies in the fact that with only certain of the target pixels for training in the image can we find the most suitable background subspace for the exact and robust detection of the rest of the target pixels. Several studies have been undertaken on dynamic subspace construction for a classification problem [34–36]. After dynamic subspace formulation, a fusion strategy is usually undertaken. Some work on the fusion method for hyperspectral target detection has already been undertaken [37–40].

In this paper, we take a dynamic selection strategy as the method for constructing the subspace. However, since the spectral resolution is one of the main advantages for hyperspectral images to be able to differentiate a target of interest from the spectrally similar background objects, the subspace construction manner in previous papers, which choose a subset of bands from the complete set of bands, known as band selection, is not used. In contrast, the subspace is not selected randomly from the different bands, but constructed by the randomly selected pixels. Iterative procedures are then carried out to evaluate the performance of the dynamic subspaces to detect the training target pixels and, finally, to obtain the most suitable one for each detection procedure. The criterion is to determine the proper number of pixel blocks for the subspace construction for the detector in each detection procedure. We then choose from the image dataset those pixel blocks under this number that present the optimal detection performance for each detection procedure. Furthermore, the manifold patch structure is also taken into consideration in optimizing the subspaces by eliminating the anomalous pixels on the manifold feature space. Then, with the remaining pixels, multiple detection procedures based on the corresponding subspaces are undertaken independently, and fused afterwards to obtain a robust detection result. Our contributions in this manuscript can be summarized as

(1) Given limited training target pixels, our method tries to find the most suitable pixels to construct the subspace for the detection, so as to have enough discriminative ability to separate the individual targets.
(2) With two nested performance analysis procedures, our method is able to choose the most informative pixels for the construction of the detector in a certain detection procedure.
(3) By constructing a multiple-detector strategy, our method is believed to be robust with regard to the complex backgrounds in the image scene. Despite the possible contamination of targets in the formulation of the particular detection procedure, target pixels can still be determined by a final fusion of all the detection procedures.
(4) The proposed dynamic subspace detection (DSD) theory is applicable to a covariance-based detector. It is also a useful background statistics estimation and selection criteria, and it provides a standard framework for optimized detection methods employing any basic state-of-the-art detector.

The remainder of this paper is organized as follows. Section 2 formulates the proposed DSD framework. Section 3 describes the experiments used to test our proposed method and presents the results of these experiments in comparison with other state-of-the-art detection methods. Finally, Section 4 summarizes the paper.

2. The dynamic subspace detection (DSD) method

In this paper, the background pixels' dataset is computed from the whole dataset. The choosing criterion is to ensure that the corresponding subspace can better augment the separability between target and non-target pixels in the detection procedure. Separability is the key to judging the performance of target detectors. It refers to the ability to separate a target of interest from the background by a certain detection method. It can be measured by the statistics of the target and background values after detection. A promising method should be able to suppress the background into a comparably low-value range and extrude the target pixels into a high-value range. The majority of the target pixels' values and the majority of the background pixels' values are expected to be distributed in a diverse range, or a gap between them is preferred. The target subspace is fixed as the number of target pixels is so low. Therefore, the manufacture of the background subspace is the key step. The following estimations are done on a whole single hyperspectral image. With the training target pixels, we want to choose the most suitable background pixels for constructing the background subspace.

2.1. Determination of the blocks for each subspace construction

Due to the rarity of the target pixels, the target subspace is actually made up of the mean spectra of the target training samples, or the signatures from the spectral library. Meanwhile, the background subspace is dynamically chosen from the whole dataset. In the dynamic subspace classifiers, the number of bands used to construct each subspace has to be determined [36], whereas in our method, the number of pixels chosen to construct the background subspace should be determined.

Unlike dynamic subspace classification (DSC) [35,36], which chooses bands from raw high-dimension samples by randomly projecting them into a subspace where all the samples have a zero constant in the unselected dimension, a method of assembly is used in our method. A subset of pixels are randomly selected from the whole dataset and assembled to form a subspace for a certain detector, such as the adaptive matched subspace detector (AMSD), adaptive cosine estimator (ACE), and so on [23]. This procedure is based on the assumption that a subset, instead of the whole dataset, may be more appropriate and reasonable for a linear subspace [33], which is consistent with the "locally linear, globally non-linear" assumption in manifold learning methods [20,41–45].

In order to obtain typical background pixels for the subspace construction, and to increase the probability of hitting pure pixels, a block structure is used as the basic choosing unit, instead of a single pixel. The whole image is segmented into many blocks, and a block is defined as a square with a size of $floor(\sqrt{L+1})\times floor(\sqrt{L+1})$, where L is the band number, and $floor()$ refers to the integral function. Each pixel corresponds to a certain block and lies in the center of the block, so the blocks are overlapping. The purpose is to ensure that the detector is composed of the more representative pixels. The number of blocks is determined by an evaluation of the importance of the block datasets with different numbers of blocks. A vector probability mass function, termed $\boldsymbol{Cd}$,