Pattern Recognition 68 (2017) 24–37

Contents lists available at ScienceDirect

Pattern Recognition

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

The matched subspace detector with interaction effects

Ziyu Wang [a,b], Jing-Hao Xue [b,*]

[a] Department of Security and Crime Science, University College London, London WC1E 6BT, UK
[b] Department of Statistical Science, University College London, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 25 July 2016
Revised 19 October 2016
Accepted 1 March 2017
Available online 2 March 2017

Keywords:
Matched subspace detector (MSD)
Linear mixing model (LMM)
Bilinear mixing model (BMM)
Interaction effects
Target detection
Hyperspectral image (HSI)

ABSTRACT

This paper aims to propose a new hyperspectral target-detection method termed the matched subspace detector with interaction effects (MSDinter). The MSDinter introduces "interaction effects" terms into the popular matched subspace detector (MSD), from regression analysis in multivariate statistics and the bilinear mixing model in hyperspectral unmixing. In this way, the interaction between the target and the surrounding background, which should have but not yet been considered by the MSD, is modelled and estimated, such that superior performance of target detection can be achieved. Besides deriving the MSDinter methodologically, we also demonstrate its superiority empirically using two hyperspectral imaging datasets.

1. Introduction

Hyperpsectral target detection aims to detect small objects from the background of a hyperspectral image (HSI) by the use of known target spectra. The number of target pixels is relatively very small compared with the total number of pixels in an HSI, e.g. only a few target pixels in millions of pixels. Typical applications of the HSI target detection include the detection of specific terrain features, minerals and crops for resource management, the detection of military vehicles and aeroplanes for defence, etc. Comprehensive overviews and gentle tutorials of the HSI target detection can be found in [1–4].

Target detection algorithms are typically derived from the binary hypothesis model, which consists of two competing hypotheses: the H_0 (absence of target) hypothesis and the H_1 (presence of target) hypothesis. The likelihood ratio or the generalised likelihood ratio (GLR) of functions of target and background can be used to construct a detector.

Some well-known detectors have been successfully applied to the HSI target detection, including the matched subspace detector (MSD) [5], the orthogonal subspace projection detector (OSP) [6], the spectral matched filter (SMF) [7,8], the adaptive coherence/cosine detectors (ACEs) [9,10] and the constrained energy minimization (CEM) [11]. Kwon et al. [12] also extend the MSD, OSP, SMF and ACEs to their corresponding kernel versions based on the kernel-based learning theory. Several methods have been developed based on the CEM specifically [13–15]. Yang et al. [13] utilise an inequality constraint on the output detector to solve the spectral variability problems, instead of the equal constraint on the CEM. A hierarchical structure of CEM [14] is proposed, which suppresses the backgrounds while preserving the target spectra to boost the performance of CEM. In a very recent work, Yang et al. [15] use total variation to constrain the spatial smoothness and show a promising detection performance when only one single target spectrum is available for training.

Sparse representation (SR)-based algorithms have also been applied to the HSI target detection [16–21]. Chen et al. [16] propose a sparsity-based target detection (STD), linearly modelling a test pixel by the training background samples and the training target samples. Zhang et al. [17] propose an SR-based binary hypothesis model (SRBBH), which is in the similar fashion of the binary hypothesis model of the MSD. The kernel versions of the STD and SRBBH can be found in [18] and [19], respectively. Detailed reviews of SR algorithms for the HSI classification and detection can be found in [20,21].

The assumption of these well-known detectors [5–10,16,17] is the linear mixing model (LMM) [22]. The LMM assumes that the spectrum of a mixed pixel can be represented as a linear combination of component spectra (endmembers). The weight (abundance) of each endmember spectrum is proportional to the fraction of the pixel area covered by the endmember. If there are p spectral bands, the p-variate spectrum $\mathbf{x} = [x_1, \ldots, x_p]^T$ of a mixed pixel can be ex-

* Corresponding author.
E-mail addresses: ziyu.wang.12@ucl.ac.uk (Z. Wang), jinghao.xue@ucl.ac.uk (J.-H. Xue).

http://dx.doi.org/10.1016/j.patcog.2017.03.002

pressed as a mixture of K endmembers $\mathbf{m}_k$ with additive noise:

$$\mathbf{x} = \Sigma_{k=1}^{K} a_k \mathbf{m}_k + \mathbf{n} = \mathbf{M}\mathbf{a} + \mathbf{n}, \tag{1}$$

where $\mathbf{M}$ is a $p \times K$ matrix whose columns are the K endmember spectra $\mathbf{m}_k = [m_{k,1}, \ldots, m_{k,p}]^T$ for $k = 1, \ldots, K$, respectively; $\mathbf{a} = [a_1, \ldots, a_K]$ is the fraction abundance vector; and $\mathbf{n} = [n_1, \ldots, n_p]^T$ represents the additive Gaussian white noise. Physical considerations dictate that the abundances have to satisfy 1) the non-negative constraint, i.e. $a_k \geq 0$, and 2) the sum-to-one constraint, i.e. $\Sigma_{k=1}^{K} a_k = 1$. Although the non-negative constraint and the sum-to-one constraint are quite meaningful, they are not always enforced because it significantly complicates the solving of detection problems. As explained in [22] and as usually the case, we can relax both constraints in target detection.

For the HSI target detection, the underlying physical assumption of the LMM is that each incident photon interacts with one earth surface component only and the reflected spectra do not mix before entering the sensor. Therefore, adopting the LMM in [5–10,16,17] assumes that the target spectral signature in the scene remains linearly mixed with the surrounding background spectra after entering the sensor. However this is not true in practice, since the target spectral signatures captured by the hyperspectral sensor can appear significantly different from the true underlying spectrum. The exhibited target spectrum may be contaminated by the *interaction effect* of its true underlying spectrum and its surrounding environments. The reasons can be, but not limited to, that the sensor picks up the signal from multiple scattering of photons and as a result, the abundance vector of targets will be dependent on the characteristics of their surrounding background.

To cope with multiple scattering problems and to model interaction effects, the bilinear mixing model (BMM) has been proposed in the hyperspectral analysis, particularly for the unmixing applications [23–28]. Nascimento et al. [23] and Fan et al. [24] address the HSI unmixing problem by taking into account of the second-order scattering interaction between endmembers, referred to as "Nascimento model" and "Fan model" hereafter, respectively. The two models are distinguished by different sum-to-one constraints imposed on the abundances. Halimi et al. [25] propose a generalised bilinear model (GBM) to unmix an HSI pixel and solve the problem by a hierarchical Bayesian algorithm. Practical analysis [26–28] also demonstrate impacts of different orders of interactions in real HSI mixing problems, such as tree cover estimates in orchards. It shows that the second-order interaction has the most significant effect of nonlinear mixing and the higher order interactions can be neglected. On top of the BMM, Heylen et al. [29] derive a multilinear mixing model (MLM) which extends the BMM to an infinite orders of interactions. Experimental studies in [23–29] have been carried out and shown superior performance of the above-mentioned nonlinear mixing models to conventional linear mixing models.

In this paper, to account for the effect of interaction between the target and their surrounding background on the target spectral signature captured by the sensor, we propose to introduce interaction effects into the models for the HSI target detection. Specifically, we propose a new model, termed the matched subspace detector with interaction effects (MSDinter), by introducing the terms that describe the interaction effects between the target and its surrounding background. To our knowledge, such model is the first one proposed for the HSI target detection. The proposed MSDinter model is able to capture better the target-background mixing effects within pixel spectrum and therefore can improve the performance of target detection.

2. The matched subspace detector

The matched subspace detector (MSD) [5] is a popular algorithm which explores the idea of the LMM binary hypothesis model (4). The task is to determine if a test pixel $\mathbf{x}$ contains materials characterised by exemplar target spectral signatures, i.e. whether the test pixel can be represented by a linear combination of target spectral signatures and background spectral signatures. In the MSD, the target spectral signatures and background spectral signatures are represented by the bases of a target subspace and the bases of a background subspace, respectively. The underlying assumption of the MSD in the HSI target detection is that each basis vector of these subspaces represents an endmember, which follows the assumption in the LMM (1).

When a target pixel presents, the spectrum of an observed pixel can be decomposed into two components under the LMM assumption, as

$$\mathbf{x} = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \tag{2}$$

where $\mathbf{T} = [\mathbf{t}_1, \ldots, \mathbf{t}_{r_t}]$ is a $p \times r_t$ matrix representing the target subspace, and $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_{r_b}]$ is a $p \times r_b$ matrix representing the background subspace; $\mathbf{T}$ is derived from a training target matrix $\mathbf{M}_{\boldsymbol{T}} \in \mathbb{R}^{p\times N_t}$ whose columns are the N_t target spectra $\mathbf{M}_{\boldsymbol{T}}(\cdot,\ n_t)$ for $n_t = 1, \ldots, N_t$, respectively; $\mathbf{B}$ is derived from a training background matrix $\mathbf{M}_{\boldsymbol{B}} \in \mathbb{R}^{p\times N_b}$ whose columns are the N_b background spectra $\mathbf{M}_{\boldsymbol{B}}(\cdot,\ n_b)$ for $n_b = 1, \ldots, N_b$, respectively; $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are the corresponding abundance vectors of the subspace $\mathbf{T}$ and the subspace $\mathbf{B}$, respectively; and $\mathbf{n}$ is the additive Gaussian white noise.

When the target is absent, the spectrum of the observed pixel is adequately described by

$$\mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \tag{3}$$

which is a reduced order model. Therefore, to decide whether a given target is present or not, we can fit the full model and the reduced model to the test pixel spectrum and check which model provides a better fitting according to certain criterion. Formulated as a binary hypothesis test, the detection problem becomes a decision between the two competing hypotheses H_0 and H_1,

$$\begin{aligned} &H_0 : \mathbf{x} = \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \ \text{target absent}, \\ &H_1 : \mathbf{x} = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\beta} + \mathbf{n}, \ \text{target present}. \end{aligned} \tag{4}$$

Model (4) is defined as the MSD model. Using the generalised likelihood ratio test (GLRT) [3], the output detector of the MSD model is given by

$$D_{\text{MSD}}(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{P}_B^{\perp} \mathbf{x}}{\mathbf{x}^T \mathbf{P}_V^{\perp} \mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \nu, \tag{5}$$

where $\mathbf{P}_B^{\perp} = \mathbf{I} - \mathbf{P}_B$ with $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ being the projection matrix onto the column space of $\mathbf{B}$; and $\mathbf{P}_V^{\perp} = \mathbf{I} - \mathbf{P}_V$ with $\mathbf{P}_V = \mathbf{V}(\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T$ being the projection matrix onto the column space of $\mathbf{V}$, where $\mathbf{V}$ is a $p \times (r_t + r_b)$ concatenated matrix of $\mathbf{T}$ and $\mathbf{B}$, i.e. $\mathbf{V} = [\mathbf{T}, \mathbf{B}]$.

The value of $D_{\text{MSD}}(\mathbf{x})$ is compared to a threshold ν to make a final decision of which hypothesis should be rejected for test pixel $\mathbf{x}$. In general, any set of orthogonal basis vectors that spans the corresponding subspace can be used as the column vectors of $\mathbf{B}$ and $\mathbf{T}$. In this paper, the significant eigenvectors (normalised by the square roots of their corresponding eigenvalues) of the background and target covariance matrices $\mathbf{C}_b$ and $\mathbf{C}_t$ are used to create the column vectors of $\mathbf{B}$ and $\mathbf{T}$, respectively.

3. The matched subspace detector with interaction effects (MSDinter)

The linear model (2) in the MSD assumes that the abundance vector $\boldsymbol{\gamma}$ of the target subspace $\mathbf{T}$ in composing a target pixel $\mathbf{x}$ will not change if the characteristics of the background change. Specifically, the effect of one-unit change of $\mathbf{T}$ on $\mathbf{x}$ is the marginal effect