ISPRS Journal of Photogrammetry and Remote Sensing 108 (2015) 138–150

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Review Article

Maximum margin metric learning based target detection for hyperspectral images

Yanni Dong [a], Liangpei Zhang [a], Lefei Zhang [b], Bo Du [b,*]

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

ARTICLE INFO

Article history:
Received 15 April 2015
Received in revised form 6 June 2015
Accepted 15 July 2015

Keywords:
Target detection
Hyperspectral image (HSI)
Metric learning
Subspace detectors

ABSTRACT

Target detection is one of the most important problems in hyperspectral image (HSI) processing. However, the classical algorithms depend on the specific statistical hypothesis test, and the algorithms may only perform well under certain conditions, e.g., the adaptive matched subspace detector algorithm assumes that the background covariance matrices do not include the target signatures, which seldom happens in the real world. How to develop a proper metric for measuring the separability between targets and backgrounds becomes the key in target detection. This paper proposes an efficient maximum margin metric learning (MMML) based target detection algorithm, which aims at exploring the limited samples in metric learning and transfers the metric learning problem for hyperspectral target detection into a maximum margin problem which can be optimized via a cutting plane method, and maximally separates the target samples from the background ones. The extensive experimental results with different HSIs demonstrate that the proposed method outperforms both the state-of-the-art target detection algorithms and the other classical metric learning methods.

1. Introduction

The main task of hyperspectral target detection is to decide whether a target of interest is present or not in each pixel of the hyperspectral image (HSI) by exploring the spectral signatures (Manolakis, 2003; Bioucas-Dias et al., 2013). Target detection is also one of the most important tasks in HSI analysis (Manolakis et al., 2001; Manolakis and Shaw, 2002; Nasrabadi, 2014). There have been a number of target detection methods proposed for HSIs. One of the earliest methods, presented by Harsanyi (1993), is based on a linear operator that minimizes the total energy in the HSI and constrains the response of the operator of the signature of interest to be a constant. This method is the so-called "constrained energy minimization" (CEM) (Du et al., 2003), and it is especially suitable in the case of the image background being complicated or very hard to characterize, because it only needs the knowledge of the desired target signature (endmember) (Chang, 2005). In addition, Harsanyi also presented orthogonal subspace projection (OSP) (Chang, 2005; Matteoli et al., 2011; Harsanyi and Chang, 1994; Chang et al., 1999), which results from the theory of least squares and has been further developed in the sensor array processing community. It is believed that this method is the first approach to separate the desired target signatures from the undesired target signatures by eliminating the undesired target signatures prior to the detection of the desired target signatures, so as to improve the signal detectability (Manolakis et al., 2003). Furthermore, the CEM algorithm has been extended into the target-constrained interference-minimized filter (TCIMF) (Ren and Chang, 2000). The CEM algorithm can accomplish the following two tasks: detection of the desired target and minimization of the interfering noise, whereas the TCIMF algorithm can also achieve elimination of the undesired targets (Chang, 2005).

Another kind of approach uses a statistical hypothesis test to differentiate the pixels containing the desired targets from those pixels that only contain the background spectra. Two good examples of this approach are the adaptive matched subspace detector (AMSD) (Manolakis et al., 2001; Kwon and Nasrabadi, 2006) and the adaptive cosine/coherence estimator (ACE) (Kraut et al., 2001; Kraut and Scharf, 1999). The AMSD algorithm is a typical structured background detector, in which the background spaces are modeled by a linear subspace approach to describe the pixels in the HSI (Ientilucci and Schott, 2005). We can also use endmembers to define the background when the noise variance is unknown. In this case, the hyperspectral data can be used to estimate the noise variance, and we then form the detector by generalized likelihood ratio (GLR) theory (Fowler and Du, 2012; Kim and Hero, 2001). The ACE algorithm is an unstructured background detector, which assumes

* Corresponding author.

http://dx.doi.org/10.1016/j.isprsjprs.2015.07.003
0924-2716/

that the additive noise has been included in the background, and it is one of the most powerful target detectors for HSIs at this point in time (Manolakis and Shaw, 2002). Hybrid sub-pixel target detection methods take advantage of the structured and unstructured models and have performed well in the circumstance of weak targets in a complex background (Broadwater and Chellappa, 2007).

Beyond the above classical target detectors which are based on signal detection theory, recent work has introduced machine learning theory into the HSI processing field (Yang and Shi, 2014; Zhang et al., 2015). For example, the linear subspace mixture models can be easily extended into the nonlinear domain by mapping the input data into a potentially infinite-dimensional feature space, which can be efficiently implemented by kernel methods such as the kernel matched subspace detectors (Kwon and Nasrabadi, 2006) and kernel OSP (Capobianco et al., 2009). The kernel-based methods try to find a proper measurement for the target detection. Another important approach drawing great interest is distance metric learning in the machine learning domain, which also constructs a distance measurement to separate the probable targets from the backgrounds. Distance metric learning methods have shown promising performance in many tasks, including classification (Weinberger et al., 2005; Jiao et al., 2012), recognition (Chopra et al., 2005; Imani and Ghassemian, 2015), and retrieval (Hoi et al., 2006; Mountrakis et al., 2011). As we know, the purpose of classification is to assign all the pixels into various classes, while target detection is designed to search for the targets of interest, and can be viewed as a binary classification problem. However, there are enormous differences between target detection and classification, in both the pixel number of the desired targets/classes and the performance evaluation (Manolakis et al., 2001). The key obstacle that needs to be addressed in target detection is the training data, especially the desired target signatures, which are very limited in number in hyperspectral target detection, and they are not enough for proper metric learning. In order to overcome this problem and guarantee the meaning of the algorithm, some constraints should be considered, e.g., pairwise constraints with "similar" or "dissimilar" labels, a non-negative constraint, a positive semi-definite constraint, and so on.

A number of methods have been previously presented to learn the distance metrics from data. Xing et al. (2002) focused on the problem of learning a distance metric to increase the accuracy of nearest neighbor algorithms, but their method is slow and requires the solving of a semi-definite programming problem. Another method (Schultz and Joachims, 2004), defines the relevant distance constraints and uses the Frobenius inner product as a regularizer. However, this method is less general than the methods using the full Mahalanobis matrix, because it assumes that the metric matrix is diagonal. The large margin nearest neighbor (LMNN) method (Weinberger et al., 2005; Shen et al., 2010) minimizes the sum of the distance of the pairs of points so that the neighbors are in the same class. Neighborhood component analysis (NCA) (Goldberger et al., 2004) maximizes a random variant of the leave-one-out k-nearest neighbor (KNN) score on the training samples.

However, to date, few studies have explored distance metric learning in hyperspectral target detection. Since we know that metric learning has been successfully used in many different domains, we attempt here to construct a type of distance metric learning for target detection. In this manuscript, we develop a metric learning method based on the maximum margin to automatically learn a Mahalanobis distance metric from the training samples. This distance metric takes the form of a large number of pairwise constraints with "similar" or "dissimilar" labels, and we use the Frobenius norm of the Mahalanobis metric matrix as a regularizer (Xiong et al., 2012; Franc and Sonnenburg, 2008; Joachims, 2006). We then combine the idea of maximum margin metric learning (MMML) and the conventional target detection algorithms to improve the separability between the target and background pixels for target detection. The contributions of this paper can be summarized as:

(1) The MMML algorithm utilizes the maximum margin framework as the objective function for the metric learning, to learn a subspace which can maximally separate target samples from background ones, especially when the target sample number is very small or the targets are difficult to detect.
(2) The proposed MMML method with pairwise constraints. By using the constrained optimization via the cutting plane method, our algorithm can improve the calculation efficiency with a strong generalization ability, and can perform well without certain assumptions.
(3) Which can be optimized in a constant number of iterations, has a strong generalization ability.
(4) We creatively combine metric learning with the ACE algorithm by transforming the original space to the Mahalanobis space with a low dimensionality, which is suitable for handling high-dimensional problems and can greatly enhance the target detection performance.

The rest of the paper is organized as follows. In Section 2, we describe the proposed MMML algorithm, including the conventional distance metric learning problem, the proposed MMML method for HSI target detection, and its solution algorithm. Section 3 details the experiments undertaken to conduct a comparison between the different algorithms with a simulated hyperspectral dataset and two challenging real-world hyperspectral datasets. Finally, the conclusions are summarized in Section 4. A flowchart of the MMML algorithm for HSI target detection is shown in Fig. 1. The input items of the MMML algorithm include both positive samples (red[1] points) and negative samples (green points).

2. The maximum margin metric learning algorithm

In this section, we first introduce metric learning and the maximum margin approach, which we use as the objective function for the metric learning, and we then present how to transfer the metric learning into such a maximum margin approach. Finally, we use the cutting plane algorithm to optimize this problem for obtaining the metric matrix **A**, and we accomplish the optimization via a primal sub-gradient method.

2.1. Metric learning

Given a set of training data, $\{\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n\} \in R^{L\times n}$ where n represents the number of samples and L is the number of features, we denote z_{ij} as the relationship between the constrained points $\mathbf{t}_i$ and $\mathbf{t}_j$. Thereafter, we assume that the number of target samples, which are known as prior positive samples, can be denoted as n^+, and the number of background samples, which are known as prior negative samples, can be denoted as n^-, in which $N = n^+ + n^- + m$ and $n = n^+ + n^-$, where m denotes the unlabeled samples and N denotes the total number of the data. We then have a set of *must-link* constraints **M** and a set of *cannot-link* constraints **C** as:

$$\begin{aligned}&\mathbf{M}:\ \forall(\mathbf{t}_i,\mathbf{t}_j)\in\mathbf{M},\quad \mathbf{t}_i,\mathbf{t}_j\in \text{same class and } z_{ij}=1\\&\mathbf{C}:\ \forall(\mathbf{t}_i,\mathbf{t}_j)\in\mathbf{C},\quad \mathbf{t}_i,\mathbf{t}_j\in \text{different class and } z_{ij}=-1.\end{aligned} \tag{1}$$

Our goal is to learn the positive semi-definite (PSD) matrix **A**, which specifies a Mahalanobis distance metric $d(\mathbf{t}_i,\mathbf{t}_j)$. Furthermore, the distance defined by matrix **A** between $\mathbf{t}_i$ and $\mathbf{t}_j$ is:

[1] For interpretation of color in Figs. 1, 7, 9 and 12, the reader is referred to the web version of this article.