Signal Processing 145 (2018) 12–25

Contents lists available at ScienceDirect

Signal Processing

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

Low-rank tensor learning for classification of hyperspectral image with limited labeled samples

Zhi He*, Jie Hu, Yiwen Wang

School of Geography and Planning, Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, Center of Integrated Geographic Information Analysis, Sun Yat-sen University (SYSU), Guangzhou 510275, China

A R T I C L E I N F O

Article history:
Received 18 July 2017
Revised 27 September 2017
Accepted 15 November 2017
Available online 15 November 2017

Keywords:
Hyperspectral image (HSI)
Classification
Low-rank
Tensor learning
Limited labeled samples

A B S T R A C T

Previous studies have demonstrated that integrating spatial information can potentially provide significant improvements for classification of hyperspectral image (HSI). However, it remains a challenging task to classify the high-dimensional HSI with limited number of training samples. In this paper, we propose a spectral-spatial classification framework based on low-rank tensor learning (lrTL). Unlike the traditional vector/matrix-based methods, the proposed lrTL method aims at improving the classification performance by naturally treating the HSI as a third-order tensor under the umbrella of multilinear algebra. First, small local patches containing the training (or test) samples are extracted from the original HSI cube by superpixel segmentation to preserve the structural information. Second, the lrTL algorithm is proposed to present the local patch of each test sample as a linear combination of all of the training patches. Low-rank constraint is enforced on the parameter tensor to capture the global structure of the HSI. Finally, the class label of the test sample can be determined by the minimal residual between the local patch containing the test sample and its approximation from different class subdictionaries. Experimental results on three benchmark HSI datasets demonstrate the effectiveness of the lrTL in improving the classification performance especially with limited training samples.

1. Introduction

Over the past few years, hyperspectral imaging sensors have emerged as an important breakthrough in remote sensing for capturing hundreds of contiguous and narrow bands from the visible to near infrared wavelength ranges, thus providing abundant information of the land-covers and drawing much attention in many applications, such as agricultural monitoring [1], mineralogy [2], forestry [3], and military affairs [4]. In those applications, one of the fundamental problem is classification [5], where each pixel in the hyperspectral image (HSI) is assigned to one of the classes based on the training samples available for each class. Sufficient training samples are usually required to gain satisfactory results due to various factors like the Hughes phenomenon, etc. Unfortunately, it is extremely hard and expensive to make samples with labels in reality. Therefore, classifying the high-dimensional HSI with limited number of training samples remains an open research issue.

Intensive work has been carried out to design reliable classifiers for classification of the HSI, e.g. decision trees (DT) [6], AdaBoost [7], artificial neural networks (ANN) [8], support vector machine (SVM) [9–11] and sparse representation-based classifiers (SRC) [12–17]. Among these approaches, the SVM [9–11] classifier, which aims at finding an optimal separating hyperplane between two classes, has provided very successful results for HSI classification. Moreover, motivated by the rapid development of compressed sensing, the SRC [12–17] has attracted much interest and become mainstream for HSI classification in the last few years. In addition, the low-rank representation (LRR) [18], which seeks the lowest-rank representation from the candidates to represent all vectors as the linear combination of the bases in a dictionary, has been of growing interest in image processing, as well as in HSI analysis [19–24]. Recently, many researches are focus on developing other promising classifiers based on ensemble learning (EL) [25,26], active learning (AL) [27,28] and deep learning (DL) [29,30]. Those classifiers can improve the classification performance from different perspectives.

Notably that the pixels in a small neighborhood are usually have similar spectral signatures and belong to the same class, many researchers have dedicated to investigating spectral-spatial classification which can incorporate the spatial contextual information into the spectral classifiers. For instance, the spatial dependence can be exploited by various spatial filters, including mor-

* Corresponding author.
E-mail address: hezh8@mail.sysu.edu.cn (Z. He).

https://doi.org/10.1016/j.sigpro.2017.11.007
0165-1684/

phological/attribute profiles [31–34], textural features (e.g. entropy [35]), Gabor filter [36–38], two-dimensional empirical mode decomposition (2D-EMD) [39,40] and two-dimensional singular spectrum analysis (2D-SSA) [41], etc. The above-mentioned methods can extract spatial features from the HSI data before pixel-wise classification is performed. One can also combine the spectral and spatial information [42] in the classifiers. For example, composite spectral-spatial kernel is proposed in [43] to flexibly balance both spectral and spatial information in the SVM classification process. Subsequently, the optimal kernel combination of multiple features are exploited by the multiple kernel learning (MKL) methods [44–46], in which an equivalent single kernel is produced from a linear combination of various base kernels. A patch alignment framework is proposed in [47] to linearly combine multiple features in an optimal way. The spatial information can also be incorporated through a regularization process, such as Markov random field (MRF) [48] and additional structured priors, including the joint sparsity [49–53], Laplacian sparsity [54,55], low-rank prior [56–60] and group-based methods [52,61]. Moreover, some other approaches take the spatial information into consideration by dividing the pixels into several patches with fixed patch sizes or producing adaptive spatial neighborhoods (e.g. watershed segmentation [62,63], graph cut [64] and superpixel segmentation [65–68], etc.).

The aforementioned approaches are almost vector or matrix-based methods, however, the intuitive representation of a HSI cube is a three-dimensional (3D) volumetric array, including one spectral dimension and two spatial dimensions. Therefore, it is more natural to treat the HSI as a 3D cube or tensor [69] to preserve the high-order data structure. A series of 3D or tensor-based methods have been successfully applied on HSI to conjunctively fuse the spatial features with spectral information. For instance, 3D discrete wavelet transform (3D-DWT) [70] is applied to extract the texture features at different scales, frequencies and orientations. The gray level co-occurrence is extended to its 3D version [71] to explore the complicated volumetric data and extract discriminant features for improved classification results. Tensor discriminative locality alignment (TDLA) algorithm is proposed in [72] for feature extraction by optimizing the discriminative locality information. Local tensor discriminant analysis (LTDA) technique is adopted in [73] for spectral-spatial feature extraction. Superpixel tensor sparse coding is proposed in [74] to utilize the high-order structure of HSI along all dimensions to better understand the data. A 3D convolutional neural network (3D-CNN) framework is proposed in [75] to extract the deep spectral-spatial features without any preprocessing or post-processing. Moreover, the 3D extension of the traditional 2D-EMD is proposed in our previous work [76,77] to treat the HSI cube as a whole entity. It is notable that the 3D or tensor-based methods have been shown to be valuable for better performance since the joint spectral-spatial structure is effectively represented.

In this paper, we propose a low-rank tensor learning (lrTL) method[1] [78] for spectral-spatial HSI classification with limited labeled training samples. Different from the vector or matrix-based methods, the lrTL method treats the HSI cube as a third-order tensor by making full use of the multilinear algebra tools. Moreover, the low-rank tensor constraint is more flexible and robust to capture the global structure of the HSI cube than the sparsity prior which strictly enforces the row sparsity and is sensitive to noise. The major steps of the proposed method are threefold. First, the original HSI data cube is divided into a number of small local patches containing the training (or test) samples by superpixel segmentation [79]. The patches can reflect the joint spectral and spatial correlations of the HSI. Then, the lrTL algorithm is proposed to represent each patch containing the test sample as a linear combination of training patches. In this step, the parameter tensor is enforced to have low-rank property. A two-step approach is adopted to solve the optimization problem. As soon as the low-rank parameter tensor is obtained, the error residuals of each class are used to determine the class label of the test sample. To sum up, the main contributions for this work lie in the following four aspects:

- We obtain the small local patches by superpixel segmentation, which flexibly identifies the spatial neighbors of the training (or test) samples;
- We obey the 3D natural of the HSI cube by third-order tensor representation, which helps to preserve the joint spectral and spatial information of the pixels;
- We capture the global structure of the HSI by low-rank constraint, which provides an effective tool for bringing discrimination information in HSI classification;
- We classify the unlabeled test samples by minimal error residuals in the proposed lrTL algorithm, which doesn't require any additional classifiers (e.g. SVM).

The rest of the paper is organized as follows. Section 2 briefly describes the related works, including the multilinear algebra tools and low-rank representation. Section 3 introduces the proposed lrTL method for HSI classification. Section 4 reports the experimental results. Finally, conclusions are drawn in Section 5.

2. Related works

In this section, we briefly describe some background on the multilinear algebra and low-rank representation.

2.1. Multilinear algebra tools

A tensor [80] is an array of numbers that transform linearly under different coordinate transformations, which can be denoted by an underlined boldface capital letter, e.g. $\underline{\boldsymbol{A}} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$, with N refers to the order of $\underline{\boldsymbol{A}}$ and the nth order of the tensor is of size I_n $(n = 1, 2, \ldots, N)$. Specifically, a third-order tensor can be represented as $\underline{\boldsymbol{A}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, while a matrix (i.e. 2D array) can be denoted by boldface uppercase letter, e.g. $\boldsymbol{A} \in \mathbb{R}^{I_1 \times I_2}$, and a vector can be expressed by boldface lowercase letter, e.g. $\boldsymbol{a} \in \mathbb{R}^{I}$. An element $(i_1, i_2, \ldots, i_N)$ of the tensor $\underline{\boldsymbol{A}}$ is denoted by $a_{i_1, i_2, \ldots, i_N}$, where $1 \le i_n \le I_n$ and $1 \le n \le N$. Moreover, the Frobenius norm of a tensor is calculated by $\|\underline{\boldsymbol{A}}\|_F = \sqrt{\sum_{i_1, i_2, \ldots, i_N} |a_{i_1 \ldots i_N}|^2}$.

Given a matrix $\boldsymbol{A}$, the rank-r projection $p(\boldsymbol{A}, r)$ is defined as the projection of $\boldsymbol{A}$ to the top-r spaces. In merit of the singular value decomposition (SVD) of $\boldsymbol{A}$ (i.e. $\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathrm{T}}$), the $p(\boldsymbol{A}, r)$ can be obtained by $p(\boldsymbol{A}, r) = \boldsymbol{U_r \Sigma_r V_r}^{\mathrm{T}}$, which is the top-$r$ truncated SVD of $\boldsymbol{A}$.

A sub-tensor is formed by restricting the indices to certain subsets of values. Particularly, the mode-n fiber of a tensor $\underline{\boldsymbol{A}} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is a vector determined by fixing all indices to single values except i_n. The mode-n unfolding (called also mode-n matricization) of a tensor $\underline{\boldsymbol{A}} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ yields a matrix $\boldsymbol{A}_{(n)} \in \mathbb{R}^{I_n \times \bar{I}_n} (\bar{I}_n = \prod_{m \neq n} I_m)$, whose columns are the corresponding mode-n fibers rearranged in a certain order, i.e. $\boldsymbol{A}_{(1)} \in \mathbb{R}^{I_1 \times I_2 I_3 \ldots I_N}$, $\boldsymbol{A}_{(2)} \in \mathbb{R}^{I_2 \times I_1 I_3 \ldots I_N}$, etc. Moreover, the n-rank of a tensor $\underline{\boldsymbol{A}}$ is determined by the matrix rank of the mode-n unfolding rank$(\boldsymbol{A}_{(n)})$, and the sum-n-rank of a tensor $\underline{\boldsymbol{A}}$ is calculated by the summation of n-rank, that means

$$\text{sum-}n\text{-rank}(\underline{\boldsymbol{A}}) = \sum_{n=1}^{N} \text{rank}(\boldsymbol{A}_{(n)}) \tag{1}$$

[1] To be exact, lrTL is an algorithm to determine the parameter tensor $\underline{\boldsymbol{W}}$ (see Section 3.2). Notably that the lrTL plays the most important role in the proposed classification method, we call our proposed method as lrTL for simplicity.