Neurocomputing xx (xxxx) xxxx–xxxx

Contents lists available at ScienceDirect

Neurocomputing

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

Integration of 3-dimensional discrete wavelet transform and Markov random field for hyperspectral image classification

Xiangyong Cao, Lin Xu, Deyu Meng*, Qian Zhao, Zongben Xu

School of Mathematics and Statistics and Ministry of Education Key Lab of Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, PR China

ARTICLE INFO

Communicated by Yue Gao

MSC:
00-01
99-00

Keywords:
Hyperspectral image classification
3-dimensional discrete wavelet transform
Support vector machine
Segmentation

ABSTRACT

Hyperspectral image (HSI) classification is one of the fundamental tasks in HSI analysis. Recently, many approaches have been extensively studied to improve the classification performance, among which integrating the spatial information underlying HSIs is a simple yet effective way. However, most of the current approaches haven't fully exploited the spatial information prior. They usually consider this prior either in the step of extracting spatial feature before classification or in the step of post-processing label map after classification, while don't integratively employ the prior in both steps, which thus leaves a room for further enhancing their performance. In this paper, we propose a novel spectral-spatial HSI classification method, which fully utilizes the spatial information in both steps. Firstly, the spatial feature is extracted by applying the 3-dimensional discrete wavelet transform (3D-DWT). Secondly, the local spatial correlation of neighboring pixels is modeled using Markov random field (MRF) based on the probabilistic classification map obtained by applying probabilistic support vector machine (SVM) to the extracted 3D-DWT feature in the first step, and then a maximum a posterior (MAP) classification problem can be formulated in a Bayesian perspective. Finally, α-Expansion min-cut-based optimization algorithm is adopted to solve this MAP problem efficiently. Experimental results on two benchmark HSIs show that the proposed method achieves a significant performance gain beyond state-of-the-art methods.

1. Introduction

Hyperspectral imaging has opened up new opportunities for analyzing a variety of materials due to the rich information on spectral and spatial distributions of the distinct materials in hyperspectral imagery, such as land-use or land-cover mapping, forest inventory, and urban-area monitoring [1]. Many hyperspectral applications can be essentially converted to a classification task, which aims to classify the image pixels of a hyperspectral image into multiple categories. Multiple state-of-the-art classification techniques have been attempted for this task and achieved good performance in certain applications [2–4].

However, these methods still tend to encounter some problems in practical scenarios. First, the available labeled training samples in HSI classification are typically limited because of the expensive image labeling cost [5], which leads to the high-dimension while low-sample-size classification issue. Second, despite the high spectral resolution of HSI, identical material may have quite different spectral signatures, whereas different materials may share similar spectral signatures [6]. The aforementioned problems, coupled with other difficulties, such as embedded noises from the sensors and environment, incline to further decrease the classification accuracy.

Various classification approaches have been investigated to address these problems. A main approach is to discover the essential discriminant features that are beneficial to classification while reducing the noise embedded in HSIs that impairs the classification performance. As to exploiting discriminant features, it has been pointed out in [7] that spatial information is more crucial than the spectral signatures in HSI classification. Therefore, a pixel-wise classification method following a spatial-filtering preprocessing step becomes a simple yet effective way to implement this technique [4,8]. As to spatial-filtering strategy, square patch is a representative method [9–11], which groups the neighboring pixels by square windows firstly and then extracts features based on the local window using other subspace learning techniques, such as low-rank matrix factorization [12–17], dictionary learning [4,18] and subspace clustering [19]. Compared with the original spectral signatures, the filtered features extracted by square patch method have less intra-class variability and higher spatial smoothness, with reduced noise in some sense.

* Corresponding author.
E-mail addresses: caoxiangyong45@gmail.com (X. Cao), xulinshadow@gmail.com (L. Xu), dymeng@mail.xjtu.edu.cn (D. Meng), timmy.zhaoqian@gmail.com (Q. Zhao), zbxu@mail.xjtu.edu.cn (Z. Xu).

http://dx.doi.org/10.1016/j.neucom.2016.11.034
Received 1 July 2016; Received in revised form 1 November 2016; Accepted 18 November 2016
Available online xxxx
0925-2312/

Except spatial-filtering method, another popular approach to exploit the spatial information is spectral-spatial method, which combines the spectral and spatial information into a classifier. Unlike the pixel-wise classification approach that does not take spatial structure information into consideration, this approach incorporates the local consistency of the neighboring pixel labels. Such a methodology has been proposed to unify the pixel-wise classification with a segmentation map [20]. Besides, Markov random field (MRF) is also a popular technique for this spectral-spatial method, which can be applied to incorporate the local correlation of neighboring pixels into the classification step. Then, the classification task can be formulated into a maximum a posterior (MAP) problem [21,22]. Finally, the graph-based segmentation algorithm [23] can be used to solve this problem efficiently. Compared with the methods that don't utilize the spatial contextual information of labels, the MRF-based HSI classification methods have been validated to achieve higher classification accuracy [21,22].

As to the above two approaches of improving the classification performance, most of the current methods usually utilize only one of them, namely, either using the spatial-filtering approach in feature extraction or spectral-spatial approach in embedding the MRF prior into the classification step. These methods still may not help us obtain the best performance due to the lack of fully utilizing the spatial information. In order to more comprehensively exploit the spatial information, a natural idea is to incorporate both of the two methodologies into a unique framework to further prompt the capability of state-of-the-arts for this HSI classification task.

Specifically, in this paper, a novel approach is proposed for supervised HSI classification. The key idea is to unify the spatial-filtering technique and the spatial smoothness (MRF) prior of labels into one framework. Firstly, a spatial-filtering method is used to produce spectral-spatial features. In our work, the 3-dimensional discrete wavelet transform (3D-DWT) [24] is adopted to generate spectral-spatial features, which have been validated to be more discriminative than the original spectral signature [7]. Secondly, the local correlation of neighboring pixels should also be introduced in order to fully exploit the spatial information. In our paper, the Markov random field (MRF), which assumes that adjacent pixels are more likely to belong to the same class, is utilized to model this local correlation. To our best knowledge, this is the first work that spatial feature extraction using 3D-DWT and spatial post-processing using MRF are taken into consideration simultaneously. Besides, probability support vector machine (SVM) is firstly utilized on the spatial 3D-DWT feature. Our extensive experimental results substantiate that such amelioration is insightful to this issue and can always guarantee a considerable improvement beyond the state-of-the-art methods in the scenarios both with relatively less labeled samples and with noisy input training samples. The proposed approach is thus expected to further prompt the frontier of this line of study.

The rest of the paper is organized as follows. In Section 2, related work regarding hyperspectral image (HSI) classification is introduced. In Section 3, the 3-dimensional discrete wavelet transform (3D-DWT) is briefly reviewed. In Section 4, the whole classification method is described. In Section 5, experimental results on two benchmark HSIs are reported. Finally, conclusions are drawn in Section 6. Throughout the paper, we denote scalars, vectors, matrices and tensors as the non-bold, bold lower case, bold upper case and curlicue letters, respectively.

2. Related work

The past two decades has witnessed prosperous developments in the field of hyperspectral image (HSI) classification. Numerous extensions along this line of research can be roughly classified into three categories: support vector machine (SVM)-based methods, sparse representation classifier (SRC)-based methods and multinomial logistic regression (MLR)-based methods.

In HSI classification, support vector machine (SVM) is a state-of-the-art approach that has shown impressive performance in high dimensional scenario [2]. The effectiveness of SVM largely depends on the choice of kernel functions, among which radial basis function (RBF) is the most widely used one. However, SVM [2] is only a pixel-wise classification method and ignores the correlations among distinct pixels in the image, and thus always results in unsatisfying classification performance. To improve the performance of SVM, multiple improved versions have been proposed, including SVM with composite kernels (SVM-CK) [25], which combines both spectral and spatial information in kernels, and multiple kernel learning (MKL) [26], which enhances the flexibility of kernels in machine learning. Except the popular SVM classifier, sparse representation classifier (SRC) has also been widely used in HSI classification [4,27]. Specifically, SRC method is based on the observation that hyperspectral pixels belonging to the same class approximately lie in the same low-dimensional subspace, and then an unknown test sample can be sparsely represented by the combination of a few training samples from the entire dictionary, while the corresponding sparse representation vector encodes the class information implicitly. Many improved versions based on SRC method have also been conducted to discover the inherent structure of adjacent pixels, including joint sparsity model (JSM) [28], which assumes small neighborhood pixels share a common sparsity support, Laplacian regularized Lasso [4], which introduces another weighting matrix to characterize the similarity among neighboring pixels based on JSM, and collaborative group Lasso (CGL) [11], which assumes that the representation matrix has a group-wise sparsity pattern and further enforce sparsity within each group. Except the two previous classifiers, multinomial logistic regression (MLR)-based classifier [29] has also been adopted in HSI classification. It aims to maximize the posterior class distributions for each sample and seems more suitable to the multi-classification task of HSI. Many studies on applying MLR to the HSI classification have obtained promising results [7]. Other utilized methods for HSI classification have also been studied, such as convolutional neural network (CNN) [30], extreme learning machine (ELM) [31], boosting [32] and semi-supervised method [33].

In summary, the performance of HSI classification can be significantly improved by integrating spatial information. More concretely, the methods for utilizing the spatial information of the hyperspectral image can be roughly divided into two main categories, namely spatial-filtering method [4,11] and spectral-spatial method [12,22,21]. For the spatial-filtering strategy, square patch is a representative approach [9–11]. This method can be used to group the neighboring pixels and then these neighboring pixels are applied in SRC and low-rank-based methods. For example, spatial correlation between neighboring pixels is utilized in [28], while square patches are taken as contextual groups in [10]. As to spectral-spatial methods, Markov random field (MRF) model is a commonly used strategy, which can be adopted to incorporate spatial information into the classification step by adding a smoothness prior term of labels on a probabilistic discriminative classification function [12,21,22]. Additionally, conditional random field (CRF) [34] is also an alternative to formulate the spectral-spatial classification model.

Although most of these aforementioned approaches achieve good performance in some scenarios, they still lack of fully utilizing the spatial information, which is a main drawback for these methods. Specifically, most of the current methods utilize the spatial knowledge either in the pre-processing feature extraction stage or in the smoothness post-processing the recovered labels, while not integratively take both useful information into a unique framework, such as SOMP [4], CGL [11] and MKL [26], which only utilize the square patch method to extract spatial information, SVM-GC [22,35] and MLRsubMLL [21], which just utilize the MRF to model the smoothness prior of labels. In order to fully exploit the spatial information, a natural idea is to incorporate both of the two techniques into the classification process. By fully discovering the spatial information, the proposed method can