ISPRS Journal of Photogrammetry and Remote Sensing 121 (2016) 11–27

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Robust multitask learning with three-dimensional empirical mode decomposition-based features for hyperspectral classification

Zhi He [a,*], Lin Liu [a,b]

[a] 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
[b] Department of Geography, University of Cincinnati (UC), Cincinnati 45221, USA

ARTICLE INFO

Article history:
Received 14 May 2016
Received in revised form 15 August 2016
Accepted 24 August 2016

Keywords:
Hyperspectral image (HSI)
Classification
Three-dimensional empirical mode decomposition (3D-EMD)
Multitask learning
Sparse
Low-rank

ABSTRACT

Empirical mode decomposition (EMD) and its variants have recently been applied for hyperspectral image (HSI) classification due to their ability to extract useful features from the original HSI. However, it remains a challenging task to effectively exploit the spectral-spatial information by the traditional vector or image-based methods. In this paper, a three-dimensional (3D) extension of EMD (3D-EMD) is proposed to naturally treat the HSI as a cube and decompose the HSI into varying oscillations (i.e. 3D intrinsic mode functions (3D-IMFs)). To achieve fast 3D-EMD implementation, 3D Delaunay triangulation (3D-DT) is utilized to determine the distances of extrema, while separable filters are adopted to generate the envelopes. Taking the extracted 3D-IMFs as features of different tasks, robust multitask learning (RMTL) is further proposed for HSI classification. In RMTL, pairs of low-rank and sparse structures are formulated by trace-norm and $l_{1,2}$-norm to capture task relatedness and specificity, respectively. Moreover, the optimization problems of RMTL can be efficiently solved by the inexact augmented Lagrangian method (IALM). Compared with several state-of-the-art feature extraction and classification methods, the experimental results conducted on three benchmark data sets demonstrate the superiority of the proposed methods.

1. Introduction

Hyperspectral imaging sensors acquire the radiance of materials in hundreds of contiguous bands from the visible to the infrared spectrum, providing high spectral resolution for each pixel to distinguish various materials. Due to the rich information captured by sensors, hyperspectral image (HSI) has opened up new opportunities and challenges in many remote sensing applications, such as classification (Jia et al., 2015; Tan et al., 2015), unmixing (Bioucas-Dias et al., 2012), data fusion (Wang and Glennie, 2015), target detection (Dong et al., 2015), and so on. Supervised classification, which aims at labeling each pixel by one of the land-cover classes based on training samples given for different classes, is an important application that has attracted extensive research efforts over the past few years. The framework of HSI classification contains two main aspects: feature acquisition and classifier construction.

To acquire discriminative features for HSI classification, many feature selection/extraction methods have been developed in the last decades. Feature selection aims to select a subset of the original features which make the classes to be distinguished accurately. Some representative feature selection methods involve the suboptimal search strategy (Serpico and Bruzzone, 2001), clustering (Martínez-Usó et al., 2007; Yuan et al., 2016), genetic algorithm (Ghamisi and Benediktsson, 2015), partial least squares regression (Li et al., 2015; Neumann et al., 2016) and game theory (Gurram et al., 2016). One of the focus of this paper is feature extraction, which transforms the original HSI into new features and preserves the vast majority information. Widespread feature extraction methods include the principle component analysis (PCA) (Zabalza et al., 2014), gray level co-occurrence matrix (GLCM) (Huang et al., 2014), wavelet transform (WT) (Bruce et al., 2002; Hsu, 2007), one-dimensional singular spectrum analysis (1D-SSA) (Zabalza et al., 2014), one-dimensional empirical mode decomposition (1D-EMD) (He et al., 2014) and manifold learning-based methods (Huang et al., 2015; Ma et al., 2016). Besides spectral information, the spatial relationship of neighboring pixels is also crucial to yield promising results. In this regard, spatial information is extracted by extended morphological profiles (EMPs) (Fauvel et al., 2013; Gu et al., 2016), extended multiattribute

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

http://dx.doi.org/10.1016/j.isprsjprs.2016.08.007
0924-2716/

profiles (EMAPs) (Song et al., 2014), image segmentation (He et al., 2016; Li et al., 2015), two-dimensional EMD (2D-EMD) (Demir and Ertürk, 2010; Gormus et al., 2012; Ertürk et al., 2013; He et al., 2013) and two-dimensional SSA (2D-SSA) (Zabalza et al., 2015). The aforementioned feature extraction methods only deal with the HSI by vector or image-based strategies. However, a HSI data set is naturally modeled as three-dimensional (3D) cube which contains a spectral dimension and two spatial dimensions. As such, much work has been carried out in the literature to treat the HSI as volumetric data and detect the spectral-spatial features simultaneously. For instance, 3D discrete WT (3D-DWT) is adopted by Qian et al. (2013) to capture geometrical and statistical spectral-spatial structures of the HSI cube. The traditional GLCM is extended into 3D scenario (Tsai and Lai, 2013) to extract discriminant texture features. Moreover, some researchers (Zhang et al., 2013; Guo et al., 2013; Zhong et al., 2015; Veganzones et al., 2016) extract spectral-spatial features by considering the HSI as a whole tensor rather than a vector or matrix.

Designing suitable classifiers also plays a vital role to yield significant classification performance. Two types of state-of-the-art classifiers are support vector machine (SVM) (Vapnik, 1995; Cavallaro et al., 2015; Dalponte et al., 2015) and sparse representation-based classification (SRC) (Chen et al., 2011, 2013; Wu et al., 2015). The former is based on structural risk minimization, while the latter is motivated by the rapid development of compressed sensing. Other powerful methods, such as ensemble learning (Samat et al., 2014), active learning (Crawford et al., 2013) and deep learning (Zhao and Du, 2016) also announce impressive results for HSI classification. Recently, many techniques, which include some variations of SVM or SRC-based methods, have been proposed to incorporate both spectral and spatial characteristics of the samples. For instance, composite kernels (Camps-Valls et al., 2006) and multiple kernel learning (He and Li, 2015; Wang et al., 2016) are proposed to balance the spectral and spatial content by optimizing the linear combination of various kernels. Markov random filed (MRF) regularization is adopted by Tarabalka et al. (2010) to take advantage of the spatial contextual information to refine the classification results. Support tensor machines (STM) (Guo et al., 2016) is developed to tackle the classification problem of HSI with tensor-based data structure. Joint sparsity model (JSM) is applied in (Chen et al., 2011; Fu et al., 2016) to incorporate the spatial correlation between neighboring pixels into a joint SRC. Multitask learning (MTL) (He et al., 2014; Li et al., 2015; Yuan et al., 2015; Jia et al., 2016) is proposed to deal with multiple features of the HSI simultaneously by treating each type of feature as a task. Moreover, low-rank representation (LRR) (He et al., 2016; Xu et al., 2015; de Morsier et al., 2016) has also gained much popularity since the high spatial similarity of HSI implies the low-rank characteristic and LRR provides a robust tool for capturing the correlation of data belonging to several subspaces. Versatile as the MTL and LRR are, to the best of our knowledge, existing MTL rarely exploits the low-rank structure of HSI to improve the classification performance.

In this paper, we extend the traditional 1D/2D-EMD into 3D (i.e. 3D-EMD) by naturally treating the HSI as a cube. The HSI can be decomposed into varying oscillations named 3D intrinsic mode functions (3D-IMFs), each of which is a 3D feature of the original HSI. In general, the computational efforts and memory requirements will exponentially increase with the addition of dimension. Therefore, two strategies are adopted to accelerate the 3D-EMD: (1) 3D Delaunay triangulation (3D-DT) (Golias and Dutton, 1997) is used to determine the distances of extrema. Rather than all of the extrema, only part of the extrema are involved in calculating the filter size. (2) Separable filters are adopted to generate the envelopes. That means, instead of performing a complicated 3D filter, we separately execute a 1D filter three times to achieve the same results as 3D filter, thus reducing the computational requirements. Subsequently, robust MTL (RMTL) is proposed to classify all of the 3D-IMFs simultaneously by taking each IMF as a task. The RMTL captures the task relationships by low-rank structure, while the specificities can also be identified in the RMTL by sparse structure. The pairs of low-rank and sparse structures are realized by trace-norm and l_{12}-norm, respectively. Inexact augmented Lagrangian method (IALM) (Lin et al., 2009) is adopted to effectively solve the optimization problem in the RMTL. Based upon the above analysis, the framework of the proposed HSI classification method is outlined in Fig. 1.

To sum up, the main innovative contributions of this work lie in the following two aspects:

(1) We present the first attempt to develop a fast 3D-EMD, which treats the HSI as a data cube and decomposes the HSI into several 3D-IMFs. Instead of extending the 1D/2D-EMD directly into 3D, two strategies (i.e. 3D-DT and separable filters) are adopted to facilitate implementation of the proposed 3D-EMD in filter size determination and envelope generation. Moreover, the 3D nature of 3D-IMFs help to better represent the spectral-spatial features of the samples in HSI. As such, compared to the vector or image-based methods, the proposed 3D-EMD method facilitates the preservation of spectral-spatial information.
(2) We take each 3D-IMF as a task and simultaneously classify the 3D-IMFs by the RMTL, which integrates multiple features with both low-rank and sparse regularizations. It is of great importance to combine multiple features due to their potential advantages in characterizing the HSI over a single feature. Note that the high spatial similarity of the HSI implies the low-rank property and the sparse matrix identifies task specificities, the proposed low-rank and sparse basedRMTL can capture both shared factors and specificities of the tasks.

The layout of this paper is as follows. Section 2 describes the proposed 3D-EMD method. Section 3 presents the RMTL for HSI classification. Experimental results on three benchmark HSI data sets are illustrated in Section 4. Finally, conclusions are drawn in Section 5.

2. 3D extension of EMD for analyzing HSI

EMD, which is first proposed by Huang et al. (1998) has attracted a great deal of attention in various applications (Li et al., 2013; Song et al., 2014) due to its ability to extract local characteristics of the non-linear and/or non-stationary data. Contrary to most of the signal processing methods (e.g. Fourier transform and WT), the EMD can adaptively decompose the non-linear and/or non-stationary data as sum of zero-mean amplitude modulation and frequency modulation (AM-FM) components termed as IMFs. As stated in Rilling et al. (2003), Deléchelle et al. (2005), Patel et al. (2016), and Ren et al. (2016), the sifting process of EMD is realized by detecting the local maxima and minima, generating the upper and lower envelops of the extrema, subtracting the mean envelopes to isolate the high-frequency oscillatory components and repeating the above-mentioned procedures recursively on the rest of data. Both 1D-EMD and 2D-EMD have been receiving considerable attention in the signal/image processing literature. As mentioned earlier, the 1D-EMD/2D-EMD methods have been successfully applied in hyperspectral classification (He et al., 2014; Demir and Ertürk, 2010; Gormus et al., 2012; Ertürk et al., 2013; He et al., 2013). Note that a HSI data is naturally formed as 3D cube, it would be promising to develop the 3D extension of