Optik 126 (2015) 3942–3948

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Hyperspectral image classification using FPCA-based kernel extreme learning machine

Yantao Wei [a], Guangrun Xiao [b], He Deng [a,c,*], Hong Chen [d], Mingwen Tong [a], Gang Zhao [a], Qingtang Liu [a]

[a] School of Educational Information Technology, Central China Normal University, Wuhan 430079, China
[b] School of Mechanical and Automotive Engineering, Hubei University of Arts and Science, Xiangyang 441053, China
[c] National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, China
[d] Department of Mathematics, Huazhong Agricultural University, Wuhan 430070, China

ARTICLE INFO

Article history:
Received 9 July 2014
Accepted 27 July 2015

Keywords:
Hyperspectral image classification
Extreme learning machine
Functional principle component analysis

ABSTRACT

In this paper, the capabilities of functional data feature extraction technique are combined with the advantages of kernel extreme learning machine (KELM), to develop an effective hyperspectral image (HSI) classification method. In the proposed method, the hyperspectral pixels are firstly represented by functions. Each pixel in the HSI is processed from the perspective of function rather than high-dimensional vector. These functional representations are transformed to a lower dimensionality feature space using functional principal components analysis (FPCA). And then the obtained lower dimensional representations are processed by a multiclass KELM classifier. Experimental results on two HSI datasets show that the proposed method provides a relatively promising performance compared with other methods.

1. Introduction

Hyperspectral image (HSI) has a relatively narrow spectral bandwidth with fixed sample intervals, which can reach a nearly continuous spectral record for the object. The high spectral resolution of the HSI pixels facilitates superior discrimination of object types in a captured scene [1,2]. Consequently, the classification of HSI is an important task for many practical applications including both commercial and military domains [3–6]. However, the high dimensionality of HSI causes several challenges to classification [7].

In order to improve the classification performance, many methods have been proposed in the past few years. Neural networks (NN) has been applied to HSI classification. Despite its success in this area, several challenges still remain open in order to incorporate such model to real applications. An important limitation of this model is the fact that their computational complexity is quite high. In recent years, support vector machine (SVM)-based approaches have been extensively used for HSI classification [8]. A major reason for the SVM's popularity in this area is its capability to produce higher classification accuracy than the NN model. However, the choice of suitable kernel function, kernel specific parameters, regularization parameter and strategies for multiclass classification are some of the major concerns in the design of an SVM [9]. Demir and Ertürk [10] proposed to use relevance vector machine (RVM) for HSI classification. RVM results in fewer relevance vectors compared with the number of support vectors obtained in the SVM. So RVM-based classification methods are faster. Recently, Huang et al. [11] proposed extreme learning machine (ELM) for classification [12]. Comparing with SVM, ELM requires fewer optimization constraints and results in simpler implementation, faster learning, and better generalization performance. It has been found to provide improved classification accuracy in comparison with the back-propagation NN for the land cover classification. However, the randomly assigned weights produce a large variation in the classification accuracy in different trials with the same number of hidden nodes. To solve this problem, Huang et al. [11] proposed the kernel-based ELM (KELM) which replaces the hidden layer of the ELM with a kernel function. The kernel function used in KELM does not need to satisfy Mercer's theorem. KELM has been applied to HSI classification [13], and the results show that it is similar to, or more accurate than, SVM in terms of classification accuracy and has notable lower computational cost.

Apart from designing effective classifier directly, feature extraction (or feature learning) is also an effective way to improve the HSI classification accuracy [14–16]. Consequently many feature extraction methods, such as, manifold learning algorithms have been

* Corresponding author at: School of Educational Information Technology, Central China Normal University, Wuhan 430079, China. Tel.: +86 18971606945; fax: +86 27 67867597.
E-mail address: denghe@mail.ccnu.edu.cn (H. Deng).

http://dx.doi.org/10.1016/j.ijleo.2015.07.184
0030-4026/

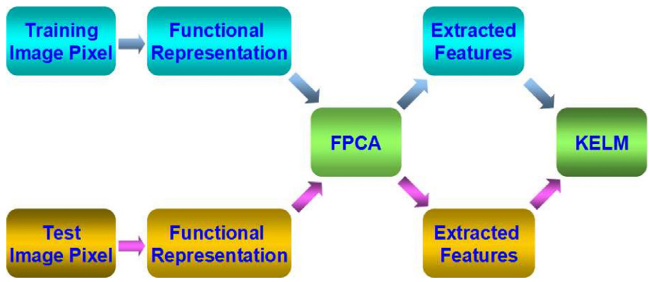

Fig. 1. The flow diagram of algorithm.

Table 1
Ground truth classes for the Indian Pines scene and their respective samples number.

Class	Class name	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93
Total		10,249

Table 2
Recognition accuracy comparisons on Indian Pines image (the best results are highlighted in bold).

Percentage	5%	10%	15%
SVM	75.08	80.61	83.32
KELM	76.02	81.91	84.19
PCA + KELM	75.95	81.71	84.17
KPCA + KELM	76.22	82.01	84.36
LPP + KELM	75.74	81.12	83.52
LDA + KELM	76.34	82.10	84.46
FPCA-based KELM	**80.86**	**85.69**	**87.62**

adopted for HSI classification in the last few years. Han and Goodenough [17] utilized the locally linear embedding (LLE) method to reduce the dimensionality of HSIs. Local preserving projection (LPP) has also been applied to HSI classification [18]. In order to make use of the label information, some supervised methods have been proposed to extract the features [14,19,20]. Recently, semi-supervised methods such as discriminative locally enhanced alignment (SDLEA) [1] have been proposed for HSI dimension reduction. The feature extraction methods mentioned above all take vectors as inputs. However, as we know, the spectrum of a target is a continuous curve. The spectral curve maps the reflectance or radiance of the studied sample as a function of wavelengths of the illuminating light. The observed hyperspectral data can be viewed as independent realizations of a smooth stochastic process [5]. Consequently, it is natural to take each pixel as a continuous curve. In this way, the functional properties of the HSI can be explored. Note that functional principal component analysis (FPCA) is a popular functional data analysis technique [21].

Keeping these points in mind, this paper designs a HSI classification method called FPCA-based KELM. In the proposed method, FPCA is adopted to increase the classification accuracy of KELM. Consequently, FPCA-based KELM is expected to achieve better performance.

The remainder of this paper is structured as follows. The details of the two methods used to compose the proposed image classification method are described in Sections 2 and 3, respectively. The proposed HSI classification algorithm is introduced in Section 4. The effectiveness of the proposed method is demonstrated in Section 5 by experimental results on several real HSIs. Finally, Section 6 summarizes this paper and makes some closing remarks.

2. Functional principal component analysis

Recently, functional data analysis (FDA) is becoming increasingly popular, since it offers an alternative way to deal with high-dimensional data problems [22]. Here "functional data" refers that observations that can be thought as real-valued curves over some domain rather than vectors in a high-dimensional space. In this paper, pixels in the HSI are treated as functions. Note that FPCA is an important dimension reduction method for functional data. Consequently, FPCA is introduced firstly to deal with the functional data.

2.1. Functional data representation

In practice, finite-dimensional representation of the analyzed functions is needed for fast computing in computer. In FDA implementation a truncated functional basis expansion is commonly used. Let us denote a functional basis on the interval J by $\{\theta_1, \theta_2,$

Table 3
Class-specific classification accuracies on Indian Pines image (the best results are highlighted in bold).

Class	SVM	KELM	PCA + KELM	KPCA + KELM	LPP + KELM	LDA + KELM	FPCA-based KELM
1	44.63	42.44	43.17	44.63	44.88	46.59	**66.10**
2	75.53	76.58	75.59	76.56	74.88	76.56	**80.89**
3	70.83	68.02	69.57	69.53	68.66	70.24	**74.00**
4	52.07	50.28	50.52	51.36	50.56	51.41	**63.76**
5	90.99	90.25	90.51	90.25	90.07	90.39	**93.29**
6	94.72	96.12	95.59	95.91	95.34	95.97	**96.71**
7	75.60	68.00	69.60	68.80	68.00	68.00	**79.20**
8	97.81	98.40	98.19	98.23	97.84	97.91	**99.37**
9	31.11	33.89	31.67	34.44	36.67	39.44	**57.78**
10	74.98	74.90	75.86	75.37	74.02	75.84	**81.45**
11	81.46	85.36	84.78	85.04	84.03	84.85	**88.04**
12	69.55	71.35	71.09	71.89	70.77	72.12	**80.58**
13	96.63	97.83	97.34	97.77	97.55	97.55	**98.53**
14	94.24	**96.18**	95.70	95.80	95.40	95.47	96.07
15	54.70	57.52	56.80	58.07	58.18	59.77	**62.77**
16	**89.28**	84.82	83.49	83.37	79.88	82.17	81.08