Measurement 80 (2016) 270–280

Contents lists available at ScienceDirect

Measurement

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

A framework of quasiconformal mapping-based kernel machine with its application to hyperspectral remote sensing

Xiaodan Xie [a,*], Bohu Li [a], Xudong Chai [b]

[a] School of Automation Science and Electrical Engineering, Beihang University, Beijing, China
[b] Beijing Simulation Center, Beijing, China

ARTICLE INFO

Article history:
Received 24 April 2015
Received in revised form 20 November 2015
Accepted 21 November 2015
Available online 30 November 2015

Keywords:
Hyperspectral image classification
Kernel learning
Quasiconformal multi-kernels learning
Machine learning

ABSTRACT

Kernel machine is a feasible and effective nonlinear feature extraction method on data analysis, for example, hyperspectral sensing data. Kernel trick improves largely the performance of learning system including recognition, clustering, prediction through the nonlinear kernel mapping from the input data space to output data space. The performance of kernel-based system is largely influenced by the function and parameter of kernel. Optimizing only the parameters is not effective to promote the kernel-based learning system, because the data distribution is not changed only changing the kernel parameter. Moreover, no any universal single kernel is very adaptive to all applications. We present a framework of quasiconformal mapping-based kernel learning machine under single kernel and multiple kernels for hyperspectral image data classification. The performance of learning system is improved largely owing to the two facts: quasiconformal kernel structure changes the data structure in the kernel empirical space; quasiconformal multikernels characterize precisely the data for improving performance on solving complex visual learning tasks. The learning framework is applied to the hyperspectral image classification, and some experiments are implemented on two hyperspectral image databases.

1. Introduction

Hyperspectral Imagery Sensing (HIS) is an important sensor technique with the measurement, analysis and interpretation of spectral data collected by an airborne or satellite based platform sensor, with the prospective applications in military monitoring, energy exploration, geographic information. Hyperspectral instruments with hundreds of contiguous spectral channels bring the developing of collecting remote imagery data. The increasing spectral and space resolution bring two problems in the practical applications: (1) the bandwidth of the communication channel limits the transmission of the full hyperspectral image data for the further processing and analysis on the ground; (2) the demand of the real-time processing for some applications. Data compression is a solution to the transmission problem but little ability for the real-time analysis. So, machine learning-based data analysis technology is a feasible and effective method to produce one image through classifying the spectrum curve of each pixel for the spectrum data of each object. The hyperspectral data machine learning system can be implemented on the satellite platform. Each pixel is classified and denoted to the different objects based on the spectrum database. The relationship of between spectral curves is the classical nonlinear relationship. So the classification is the nonlinear and complex classification problem. Traditional classification methods are not effective to hyperspectral sensing data, kernel learning is a feasible and

* Corresponding author.
E-mail addresses: xiexiaodanbeijing@126.com (X. Xie), bohulibeijing@126.com (B. Li), xudongchaibeijing@126.com (X. Chai).

http://dx.doi.org/10.1016/j.measurement.2015.11.034
0263-2241/

effective nonlinear classifier methods on hyperspectral sensing data.

Kernel-based machine learning is one important pre-processing step of hyperspectral sensing data. Kernel-based learning is to solve the problem of linear learning using kernel trick. Recently many linear methods are kernelized, for example, Linear Discriminant Analysis (LDA) [3], Principal Component Analysis (PCA) [1], are kernelized to Kernel PCA (KPCA), Kernel Discriminant Analysis (KDA) [16]. The performance of the linear methods is improved because kernel method is used to improve the performance of these linear methods, and it represents the complicated nonlinear relationships. Recently researchers are developing the kernel learning methods, Baudat and Anouar [2], Liang and Shi [11], Wang [14], Chen [4] and Lu [12] developed a series of improved KDA methods, and other researchers present the alternative framework of KLPP to develop a framework of KPCA + LPP [9,6] for image recognition and radar target recognition, and other researchers improved LPP with kernels [18,5,17,7,23]. Researchers optimized the parameters of kernel function to improve kernel-based learning [7,13,4]. And these methods is to select the optimal kernel parameter from a set of discrete values, but the geometry structure of data distribution in the kernel-based mapping space is not be changed only adjusting the parameters. For it, Xiong proposed a data-depend kernel for kernel optimization-based machine learning [15], and Amari presents the support vector machine classifier through modifying the kernel function [1]. In the previous works [13,8], authors presented data-dependent kernel for face recognition. In order to solve the kernel function, the more complicated kernel model, multiple kernel learning methods are developed to solve the kernel model selection problems [19–22].

We present a framework of quasiconformal mapping-based kernel learning under a single kernel and multiple kernels learning, and the practical application framework of hyperspectral image classification is also proposed. The framework is to improve the performance of the learning system, from the following viewpoints: (1) quasiconformal single kernel structure changes the data structure in the kernel empirical space; (2) quasiconformal multiple kernels are combined to more precisely characterize the data for improving performance on solving complex visual learning tasks.

2. Framework

2.1. Motivation and framework

Computing the kernel matrix with the selected kernel function is the first key step of kernel learning for the classification, clustering and other statistical pattern analysis on image processing, pattern recognition, and machine learning. The performance of kernel-based system is largely influenced by the function and parameter of kernel. Optimizing the parameters is not effective to promote the kernel-based learning system owing to the unchanged data structure with the changing of the parameter of kernel function. No a universal single kernel is very effective way to detecting intrinsic information for the complicate sample data in the input data space. For example, in the kernel-based image processing, each kind of kernel function can only be used to describe the characteristics of the data effectively, such as texture, color and edge. The researchers began to have a significant interest in the multi kernel function of the combination of basic nuclear functions. The multiple kernel learning has the different kernel representation for the different feature subspace. Multiple kernel learning is an effective method to solve the above problems. Multiple kernel learning method can choose the kernel function according to the different characteristics of the data, and then combine them. Multiple kernel learning is a feature extraction method of combining many features, and is better to describe the data features than the single feature extraction method. The multiple kernels learning can not only improve the efficiency of feature extraction, but also adjust the weights of the adaptive selection and the basic kernel function. Multiple kernel learning not only preserves the nonlinear mapping characteristics of kernel functions, but also shows the possibility of using different kernel functions and also shows the possibility of a unified framework for the classical method of nonlinear feature extraction. Quasiconformal mapping-based kernel is to change the data structure in the feature space through adjusting the parameter of quasiconformal function. In order to improve the performance of kernel learning, we combine the idea of the quasiconformal kernel and multiple kernel combination.

2.2. Empirical kernel mapping

Kernel-based learning performs on data classification, clustering, and regression in the feature embedding subspace. The embedding map is called by "empirical kernel map" [24], correspondingly the embedding space is called by the "empirical feature space". Since $K = [k_{ij}]_{m\times m}$ is a symmetrical positive semidefinite matrix, then K can be decomposed as

$$K_{m\times m} = P_{m\times r}\Lambda_{r\times r}P^T_{r\times m} \tag{1}$$

where Λ is a diagonal matrix consisted of r positive eigenvalues of K, and P is consisted with the eigenvectors corresponding. The mapping to an r dimensional Euclidean space is the empirical kernel map [24] in the empirical feature space – $\Phi^e_r \in R^r$.

$$\Phi^e_r : \chi \longrightarrow R^r$$

$$x \longrightarrow \Lambda^{-1/2}P^T(k(x,x_1), k(x,x_2), \ldots, k(x,x_m)) \tag{2}$$

Kernel matrix is computed with the selected kernel function. The performance is influenced by kernel function and its parameter, because the sample data distribution is different with the different kernel function mapping. So it is feasible to improve the kernel learning performance through adjust the kernel model. But only changing the kernel parameter for a selected kernel function is not so effective to improve the kernel performance, because the data distribution is determined by kernel model while data structure is not be adjusted by the different kernel