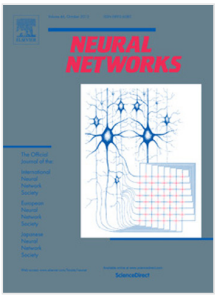
Accepted Manuscript



Orthogonal self-guided similarity preserving projection for classification and clustering

Xiaozhao Fang, Yong Xu, Xuelong Li, Zhihui Lai, Shaohua Teng, Lunke Fei

PII: S0893-6080(17)30001-1
DOI: http://dx.doi.org/10.1016/j.neunet.2017.01.001
Reference: NN 3699

To appear in: *Neural Networks*

Received date: 23 August 2016
Revised date: 20 November 2016
Accepted date: 5 January 2017

Please cite this article as: Fang, X., Xu, Y., Li, X., Lai, Z., Teng, S., & Fei, L. Orthogonal self-guided similarity preserving projection for classification and clustering. *Neural Networks* (2017), http://dx.doi.org/10.1016/j.neunet.2017.01.001

Orthogonal Self-Guided Similarity Preserving Projection for Classification and Clustering

Xiaozhao Fang[a], Yong Xu[b,*], Xuelong Li[c], Zhihui Lai[d], Shaohua Teng[a], Lunke Fei[b]

[a]*School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, 510006, China*
[b]*Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, 518055, Guangdong, China*
[c]*Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Chinese Academy of Sciences, Xian, 710119, Shaanxi, China*
[d]*College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, 518055, Guangdong, China*

Abstract

A suitable feature representation can faithfully preserve the intrinsic structure of data. However, traditional dimensionality reduction (DR) methods commonly use the original input features to define the intrinsic structure, which makes the estimated intrinsic structure unreliable since redundant or noisy features may exist in the original input features. Thus a dilemma is that 1) one needs the most suitable feature representation to define the intrinsic structure of data and 2) one should use the proper intrinsic structure of data to perform feature extraction. To address the problem, in this paper we propose a unified learning framework to simultaneously obtain the optimal feature representation and intrinsic structure of data. The structure is learned from the results of feature learning, and the features are learned to preserve the refined structure of data. By leveraging the interactions between the process of determining the most suitable feature representation and intrinsic structure of data, we can capture accurate structure and obtain the optimal feature representation of data. Experimental results demonstrate that our method outperforms state-of-the-art methods in DR and subspace clustering. The code of the proposed method is available at "http://www.yongxu.org/lunwen.html".

Keywords: Dimensionality reduction, intrinsic structure, subspace clustering, feature representation

1. Introduction

In many computer vision and pattern recognition applications, high-dimensional data often contain some redundant or noisy features. The performance of algorithm may drop exponentially as the dimensionality of data increases [3, 40, 21, 17]. Therefore, it is essential to seek a low-dimensional representation for the original high-dimensional data [30, 9]. Principal component

*Corresponding author
Email address: yongxu@ymail.com (Yong Xu)