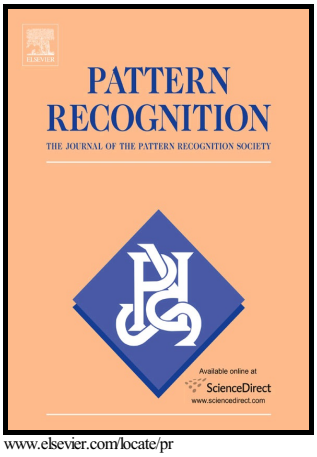
Joint Hypergraph Learning and Sparse Regression for Feature Selection

Zhihong Zhang, Lu Bai, Yuanheng Liang, Edwin Hancock



www.elsevier.com/locate/pr

PII: S0031-3203(16)30127-3
DOI: http://dx.doi.org/10.1016/j.patcog.2016.06.009
Reference: PR5763

To appear in: *Pattern Recognition*

Received date: 8 August 2015
Revised date: 18 May 2016
Accepted date: 14 June 2016

Cite this article as: Zhihong Zhang, Lu Bai, Yuanheng Liang and Edwin Hancock, Joint Hypergraph Learning and Sparse Regression for Feature Selection, *Pattern Recognition,* http://dx.doi.org/10.1016/j.patcog.2016.06.009

Joint Hypergraph Learning and Sparse Regression for Feature Selection

Zhihong Zhang[a], Lu Bai[b,*], Yuanheng Liang[a], Edwin Hancock[c]

[a]*Xiamen University, Xiamen, China*
[b]*Central University of Finance and Economics, Beijing, China.*
[c]*University of York, York, UK.*

Abstract

In this paper, we propose a unified framework for improved structure estimation and feature selection. Most existing graph-based feature selection methods utilise a static representation of the structure of the available data based on the Laplacian matrix of a simple graph. Here on the other hand, we perform data structure learning and feature selection simultaneously. To improve the estimation of the manifold representing the structure of the selected features, we use a higher order description of the neighbourhood structures present in the available data using hypergraph learning. This allows those features which participate in the most significant higher order relations to be selected, and the remainder discarded, through a sparsification process. We formulate a single objective function to capture and regularise the hypergraph weight estimation and feature selection processes. Finally, we present an optimization algorithm to recover the hyper graph weights and a sparse set of feature selection indicators. This process offers a number of advantages. First, by adjusting the hypergraph weights, we preserve high-order neighborhood relations reflected in the original data, which cannot be modeled by a simple graph. Moreover, our objective function captures the global discriminative structure of the features in the data. Comprehensive experiments on 9 benchmark data sets show that our method achieves statistically significant improvement over state-of-art feature selection methods, supporting the effectiveness of the proposed method.

*Corresponding author
Email address: bailucs@cufe.edu.cn (Lu Bai)