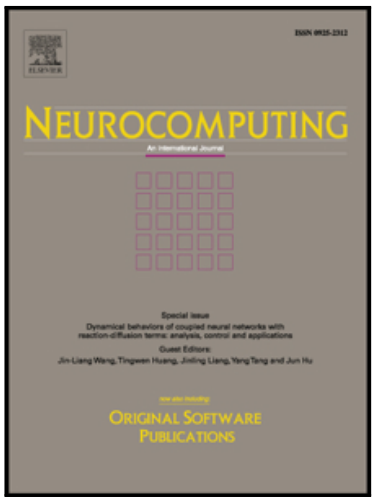

Big Data Analytics Enabled by Feature Extraction Based on Partial Independence

Qiao Ke, Jiangshe Zhang, Houbing Song, Yan Wan

PII: S0925-2312(17)31921-5
DOI: 10.1016/j.neucom.2017.07.072
Reference: NEUCOM 19193

To appear in: *Neurocomputing*

Received date: 14 November 2016
Revised date: 11 February 2017
Accepted date: 2 July 2017

Please cite this article as: Qiao Ke, Jiangshe Zhang, Houbing Song, Yan Wan, Big Data Analytics Enabled by Feature Extraction Based on Partial Independence, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.07.072

Big Data Analytics Enabled by Feature Extraction Based on Partial Independence

Qiao Ke[a], Jiangshe Zhang[a,*], Houbing Song[b], Yan Wan[c]

[a]*School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China*
[b]*Department of Electrical and Computer Engineering, West Virginia University, Montgomery, WV 25136*
[c]*Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019*

Abstract

Complex cells in primary visual cortex (V1) selectively respond to bars and edges at a particular location and orientation. Namely, they are relatively invariant to the phase as well as selective to the frequency and orientation emerging from natural images that are analogous to the characteristics of complex cells in V1 with the energy function of receptive fields (RFs) from tuning curve test with sinusoidal function in our related jobs. In this paper, we propose a feature learning algorithm based on the overcomplete AISA to apply on big data in parallel computing. In order to demonstrate the effectiveness of the overcomplete AISA features in the classification task, two feature representation architectures are evolved into the partial independent signal bases and partial independent factorial representation, respectively. Experiments on four datasets (Coil20, Extended YaleB, USPS, PIE), acquired conjunction with two classification architectures based on the overcomplete AISA features, show that the classification accuracy is mostly higher than those obtained from the other ICA related features and two other sparse representation features with a small number of training samples via nearest neighbor (NN) classification method.

Keywords: Independent Component(IC); Overcomplete Features; Sparse Representation; Big Data;

*Corresponding author
Email address: jszhang@mail.xjtu.edu.cn (Jiangshe Zhang)