Non-negativity and dependence constrained sparse coding for image classification

Hong Han, Sanjun Liu, Lu Gan

PII: S1047-3203(14)00199-0
DOI: http://dx.doi.org/10.1016/j.jvcir.2014.12.002
Reference: YJVCI 1455

To appear in: *J. Vis. Commun. Image R.*

Received Date: 13 November 2013

Please cite this article as: H. Han, S. Liu, L. Gan, Non-negativity and dependence constrained sparse coding for image classification, *J. Vis. Commun. Image R.* (2014), doi: http://dx.doi.org/10.1016/j.jvcir.2014.12.002

Non-negativity and Dependence Constrained Sparse Coding for Image Classification

Hong Han*, Sanjun Liu, Lu Gan
School of Electronic Engineering, Xidian University
Xi'an, 710071, China
*Email: hanh@mail.xidian.edu.cn

Abstract: Sparse coding method is a powerful tool in many computer vision applications. However, due to the combinatorial optimization of sparse coding involving both additive and subtractive interactions, the features can cancel each other out with subtraction. And also, in the process of independent coding, the locality and the similarity among the instances which be encoded may be lost. To solve these problems, an image classification framework by leveraging the Non-negative Matrix Factorization and graph Laplacian techniques is presented. Firstly, the Non-negative Matrix Factorization is used to constrain both of the codebook and the corresponding coding coefficients non-negativity. To preserve the dependence properties of the locality and the similarity among the instances, the graph Laplacian regularization is utilized. Then, along with max pooling and spatial pyramid matching, we extend our method to Bag-of-Words image representation. Finally, the linear SVM is leveraged for image classification. Experimental results show that the proposed method achieves or outperforms the state-of-the-art results on several benchmarks.

Keywords：Non-negative Matrix Factorization; Graph Laplacian; Sparse Coding; Image Classification

1. Introduction

In recently years, lots of researchers have turned to the study of image classification, which is one of the fundamental problems in computer vision. To address this problem, many image representation models have been proposed, among of which the bag of visual words (BoW) model [1] has shown