A Robust Image Classification Scheme with Sparse Coding and Multiple Kernel Learning

Dongyang Cheng[1], Tanfeng Sun[1,2,3], and Xinghao Jiang[1,2,*]

[1] School of Information Security Engineering Shanghai Jiao Tong University, Shanghai 200240, China
{dycheng,tfsun,xhjiang}@sjtu.edu.cn
[2] National Engineering Lab on Information Content Analysis Techniques, GT036001, Shanghai 200240, China
[3] Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark 07102, USA

Abstract. In recent researches, image classification of objects and scenes has attracted much attention, but the accuracy of some schemes may drop when dealing with complicated datasets. In this paper, we propose an image classification scheme based on image sparse representation and multiple kernel learning (MKL) for the sake of better classification performance. As the fundamental part of our scheme, sparse coding method is adopted to generate precise representation of images. Besides, feature fusion is utilized and a new MKL method is proposed to fit the multi-feature case. Experiments demonstrate that our scheme remarkably improves the classification accuracy, leading to state-of-art performance on several benchmarks, including some rather complicated datasets such as Caltech-101 and Caltech-256.

Keywords: Sparse coding, MKL, Feature fusion.

1 Introduction

Nowadays, image classification has captured a lot of interest in computer vision. The common classification schemes mainly consist of two parts: image representation and classification.

With regard to image representation models, Bag of Words (BoW) model with following three modules has been widely used and shows good performance: (i) Region selection and representation; (ii) Codebook generation and feature quantization; (iii) Frequency histogram based image representation. Specifically, the codebook consisting of entries of visual words is used to reconstruct the input local features. The process to generate the codebook and quantize features governs the quality of image representation. But the frequently used k-means method may lead to severe information loss since it assigns each feature to only one visual word in the codebook.

* Corresponding author.

Y.Q. Shi, H.J. Kim, and F. Pérez-González (Eds.): IWDW 2012, LNCS 7809, pp. 520–529, 2013.

After the image is represented as a histogram of visual words, a classifier will be required to make the decision that which category the histogram belongs to. Kernel based classifiers such as support vector machine (SVM) are now widely used by many researchers for their wonderful performance. For SVM, the input histograms are mapped to a higher dimensional space by kernel function, in which they can be easily classified in a linear way. However, the sensitiveness of kernel function to categories will increase the fluctuation in accuracy, resulting in a relatively unsatisfying overall performance.

Many works have been done to improve the classification performance. Yang *et al.* [1] applied sparse coding instead of k-means since it can learn the optimal codebook and reduce the information loss. Zhang *et al.* [2] proposed a framework by leveraging an improved sparse coding method, low-rank and sparse matrix decomposition techniques. Linear SVM classifier is used for classification. Gao *et al.* [3] proposed a robust Laplacian sparse coding algorithm for feature quantization which generated more discriminative sparse codes. Naveen *et al.* [4] presented a new framework which was built upon a way of feature extraction that generates largely affine-invariant features and an AdaBoost based classifier. From the perspective of classifier, multiple kernel learning (MKL) can increase the stability of overall performance by learning a linear combination of a series of kernel functions. Bosch *et al.* [5] combined different features by using a weighted linear combination of kernels, where the weights were learnt on a validation set. Lampert *et al.* [6] proposed a method to combine the efficiency of single class localization with a subsequent decision process that worked jointly for all given object classes.

In this paper, we devise a novel image classification scheme by adopting sparse coding and multi-feature MKL, which can ameliorate the image representation and classification phase respectively. The improved multi-feature MKL is proposed based on original MKL, in order to adapt to multi-feature case. Specifically, SIFT and SURF descriptors are extracted and then converted into sparse vectors precisely by the trained dictionaries. The images can be represented by these vectors using max-pooling method which is proved to be more robust than others. After that, the two descriptors are combined into a single vector. Finally, multi-feature MKL approach is implemented to train and test those histograms, generating stable results due to the auto adjustment of the linear combination of kernel functions for each feature.

2 Proposed Scheme

As two main parts in image classification scheme, image representation and classification can substantially affect the classification performance. On one hand, a good kernel method for classification is necessary, for it provides an intuitive and principled tool for learning from high-dimensional vectors that represent images. On the other hand, the performance of kernel method strongly depends on the data representation of images, which means an accurate image representation algorithm is indispensable. Our paper is to enhance the image classification accuracy through the amelioration of both parts.