Pattern Recognition ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Class relatedness oriented-discriminative dictionary learning for multiclass image classification

Dongyu Zhang [a,*], Pengju Liu [b], Kai Zhang [b], Hongzhi Zhang [b], Qing Wang [a], Xiaoyuan Jing [c]

[a] Sun Yat-Sen University, Guangzhou 510275, China
[b] Harbin Institute of Technology, Harbin 150001, China
[c] State Key Laboratory of Software Engineering of Wuhan University, WuHan 430079, China

ARTICLE INFO

Article history:
Received 4 September 2015
Received in revised form
10 November 2015
Accepted 7 December 2015

Keywords:
Dictionary learning
Joint sparsity
$\ell_{1,\infty}$-norm
Support vector machine
Class relatedness

ABSTRACT

Dictionary learning (DL) has recently attracted intensive attention due to its representative and discriminative power in various classification tasks. Although much progress has been reported in the existing supervised DL approaches, it is still an open problem that how to build the relationship between dictionary atoms and the class labels in multiclass classification. In this paper, based on the assumption that the relevance of dictionary atoms could be helpful in multiclass classification task, we proposed a class relatedness oriented (CRO) discriminative dictionary learning method for sparse coding. Utilizing the $\ell_{1,\infty}$-norm regularization on the coding coefficient matrix, the proposed method can adaptively learn the class relatedness between dictionary atoms and the multiclass labels. Experimental results of face recognition, object classification, and action recognition demonstrate that our proposed method is comparable to many state-of-the-art DDL methods.

1. Introduction

As a hot topic in computer vision community, image classification and recognition have inspired many interesting works [1–6]. In recent years, sparse representations based images classification have received considerable interest [7–9]. By using over complete dictionary, sparse representation represents a signal or image as the sparse linear combination of the dictionary atoms. Unlike the principle component analysis based decomposition, sparse representation do not impose the dictionary elements to be orthogonal, which allows more flexibility to adapt the representation to the input signal or image. Recent publications show that sparse representation has been successfully applied to different kinds of images classification tasks, such as face recognition [9,36], handwriting digits recognition [18,37], natural images classification [10], and so on [11–13,40,41].

In sparsity based classification, a query sample $\mathbf{x} \in \mathcal{R}$ is first represented over a dictionary $\mathbf{D} \in \mathbb{R}^{m \times K}$ with a sparse coefficients vector α as $\mathbf{x} \approx \mathbf{D}\alpha$. Then the classification is performed on the coefficients vector α and dictionary $\mathbf{D}$. Thus, the effectiveness of the sparse representation model highly rely on the design of the overcomplete dictionary. One possible route to design the dictionary is to use a prespecified transform matrix, such as FFT bases or wavelet bases, which often leads to simple and fast algorithms for the evaluation of the sparse representation. Although taking these kinds of analytically designed off-the-shelf dictionary is universal to all types of images, it might be not effective enough for specific classification task, such as fine-categories flowers classification [37] and face recognition [9,36]. Another possible way to design the dictionary is to learn the dictionary elements/atoms from the input training data with sparsity regularization. Wright et al. [9] employed all of training samples as the dictionary for sparse representation and achieved impressive performances on face recognition. In [12], by generalizing k-means clustering, Aharon et al. proposed K-SVD algorithm to efficiently learn an overcomplete dictionary from a set of training samples. Recent works show that the dictionary learning methods have been received considerable interest and led to state-of-the-art results in image reconstruction [7,8,12,14] and image classification [9,11,15,16,36,37].

The current dictionary learning methods can be grouped into two categories, unsupervised dictionary learning [12,13] and supervised dictionary learning [15–17,37]. In unsupervised dictionary learning method, the dictionary is designed to minimize the residual error of reconstructing the training samples without using the classification labels. The dictionaries produced in such way can faithfully represent the training samples, which are useful for image reconstruction. However, they are not advantageous for image classification tasks. Although some of unsupervised dictionary learning were applied for classification tasks, recent

* Corresponding author.
E-mail address: cszhangdy@163.com (D. Zhang).

http://dx.doi.org/10.1016/j.patcog.2015.12.005
0031-3203/© 2015 Elsevier Ltd. All rights reserved.

research indicated that the supervised dictionary learning can yield better classification performance by exploiting the class discrimination information [15–17].

The current supervised dictionary learning methods can be roughly grouped into two categories. In the first category, a shared dictionary by all classes is learned with discriminative representation coefficients simultaneously [16,17,21]. In the second category, multiple dictionaries or class-specific dictionaries are learned [18,20,23,27,38]. Ramirez et al. [18] proposed a structured dictionary learning scheme by promoting the discriminative ability of different class-specific sub-dictionaries. Zhou et al. [27] proposed to learn multiple dictionaries for visually correlated object classification. However, there are potential problems in both of the two categories of supervised dictionary learning methods. For the first categories, each of the dictionary atom is associated to all the classes, while the possible mixed information of different classes may reduce the discrimination of the learned dictionary. For the second categories, each dictionary atom is assigned to a single class, but without exploiting the possible correlation of different class dictionary atoms which may help in promoting the discrimination performance. Both of the two cases have ignored the fact that the relationship between the dictionary atoms and class labels needs to be updated during the dictionary learning process. Although much progress have been made in dictionary learning, it is still an open problem to adaptively build the relationship between dictionary atoms and class labels.

In this paper, we proposed a new discriminative dictionary learning (DDL) scheme to adaptively learn the relationship between the dictionary atoms and the class labels by using a joint sparsity constraint on the coding vectors of each class of training samples. Specifically, the joint sparsity is enforced on the representation coefficients with the $\ell_{1,\infty}$-norm regularization, which has been widely applied as penalties in signal processing [24] and machine learning [25,26]. Besides, in order to make the coding vectors more discriminative, we also add a linear classifier to the objective function. Thus, we can learn a class related dictionary and a multiclass linear classifier simultaneously. For a test sample, we could use the learned dictionary to obtain the corresponding coding vector, and then predict its label with the linear classifier. Similar work of exploiting the relation of dictionary atoms and class labels has been studied recently. For example, in [37] Gao et al. learned a shared dictionary and category specific dictionaries for fine-grained flower images categorization, in which the shared dictionary used to describe the correlated relation of different classes. However, in their method the structures of shared dictionary and category specific dictionaries are pre-specified. We argue that the relationship between dictionary atoms and class labels should not be predefined. Instead, they should be learned adaptively during the dictionary learning process. This piece of work of is an extension of our conference paper, in which we further analyze the motivation and the principle of the proposed class relatedness dictionary method. More experiments have also been performed to comprehensively evaluate the proposed method.

The remainder of this paper is organized as follows. In Section 2, we briefly introduce the related work about DDL methods. Then we propose our class relatedness oriented (CRO) DDL model, together with the corresponding optimization method in Section 3. Experimental results of the proposed method are discussed in Section 4, and finally Section 5 concludes this paper.

2. Related work

Discriminative dictionary learning (DDL), has been successfully applied in pattern recognition applications such as image classification [18,19,29] and face recognition [9,20,36]. DDL methods concentrate on the discriminative classification capability of the dictionary as its goal is to assign correct class labels to the test samples. To enrich such capability, how to build the relationship between dictionary atoms and class labels plays a crucial role in the dictionary learning process.

Based on the relationship between dictionary atoms and class labels, prevailing DDL methods can be roughly divided into two main categories: global dictionary learning methods and class-specific dictionary learning methods. In global dictionary learning, the dictionary atoms are shared by all class and the coding vectors are generally explored for classification. Mairal et al. [21] proposed a DDL method by learning a shared dictionary and training a classifier on coding vectors simultaneously for handwriting digit recognition and texture classification. Zhang and Li [17] proposed a joint learning algorithm based on K-SVD for face recognition. Pham and Venkatesh [16] proposed to jointly train the dictionary and classifier for face recognition and object categorization. Cai et al. [22] proposed a support vector guided dictionary method (SVGDL) to jointly optimize the dictionary and classifier. Even though a global dictionary can be powerful to represent training data, all the above methods fail to adaptively learn the correspondence between dictionary atoms and class labels. As each dictionary atom is shared by all classes, the mixed information from different class samples may reduce the discrimination of the learned dictionary.

In class-specific DDL methods, each dictionary atom is assigned to a single class and the dictionary atoms associated with different classes are encouraged to be as independent as possible. Ramirez et al. [18] proposed a structured dictionary learning scheme by promoting the discriminative ability between different class-specific sub-dictionaries. Castrodad and Sapiro [23] learned a set of class-specific sub-dictionaries with non-negative penalty on both dictionary atoms and coding vectors. Yang et al. [20] proposed a DDL framework which employs Fisher discrimination criterion to learn class-specific dictionaries. Since each dictionary atom has a single label, the reconstruction error with respect to each class could be used for classification. However, those methods ignored the cross relatedness of different dictionary atoms and class labels, e.g., sometimes it is helpful in promoting the performance by assigning some dictionary atoms to different class labels in multiclass classification task.

The above DDL approach associates dictionary atoms and class labels in two extreme manners: the dictionary atom is either associated to all classes, or assigned to a single one. In order to adaptively build the relationship between dictionary atoms and class labels, we propose a well-principled DDL scheme by applying joint sparsity constraint on the coding vectors of each class with $\ell_{1,\infty}$-norm regularization, respectively. Since the $\ell_{1,\infty}$-norm is a matrix norm that encourages entire rows of a matrix to be zeros, the resultant coding vector of a certain class should be row sparse. Besides, by incorporating a classifier into the objective function to promote the discriminative ability of coding vectors, our method would adaptively build the relatedness between class labels and the dictionary atoms in the training phase.

3. The CRO-DDL method

In this section, we first briefly describe the general DDL model, and then propose our CRO-DDL method. The process of parameters optimization for the proposed method is also presented, and the classification rule is discussed in the end of this section.

Please cite this article as: D. Zhang, et al., Class relatedness oriented-discriminative dictionary learning for multiclass image classification, Pattern Recognition (2016), http://dx.doi.org/10.1016/j.patcog.2015.12.005