Neurocomputing 190 (2016) 124–131

Contents lists available at ScienceDirect

Neurocomputing

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

Learning a structure adaptive dictionary for sparse representation based classification

Heyou Chang [a,b], Meng Yang [b], Jian Yang [a]

[a] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, PR China
[b] Shenzhen University, Shenzhen, PR China

A R T I C L E I N F O

Article history:
Received 12 October 2015
Received in revised form
9 December 2015
Accepted 14 January 2016
Communicated by Su-Jing Wang
Available online 4 February 2016

Keywords:
Structure adaptive dictionary learning
Sparse representation
Fisher criterion
Image classification

A B S T R A C T

Dictionary learning (DL), playing a key role in the success of sparse representation, has led to state-of-the-art results in image classification tasks. Among the existing supervised dictionary learning methods, the label of each dictionary atom is predefined and fixed, i.e., each dictionary atom is either associated to all classes or assigned to a single class. In this paper, we propose a structure adaptive dictionary learning (SADL) method to learn the relationship between dictionary atoms and classes, which is indicated by a binary association matrix and jointly optimized with the dictionary. The binary association matrix can not only represent class-specific dictionary atoms, but also hyper-class dictionary atoms shared by multiple classes. Furthermore, discrimination is explored by introducing Fisher criterion on coding coefficient and reducing between-class dictionary coherence. The extensive experimental results have shown that the proposed SADL can achieve better performance than previous supervised dictionary learning methods on various classification databases.

1. Introduction

Sparse representation has been shown to be a powerful and efficient tool for various vision tasks [1–3]. The basic model of sparse representation suggests that natural signals can be compactly expressed or efficiently approximated as a linear combination of dictionary atoms, where the coefficients are sparse. The dictionary plays an important role in the success of sparse representation [4]. Using off-the-shelf bases as dictionary (wavelets) might be universal to all types of images but would not be effective enough for some specific image classification tasks (such as face classification). Learning a dictionary from training samples has led to the state-of-the-art results in many particular applications, such as image processing [5], face recognition [6,7] and image classification [8,9].

Current prevailing supervised DL methods can be divided into three main categories. In the first category, a dictionary shared by all classes and a classifier trained on the representation coefficients computed on the dictionary are learnt jointly [5,6,10–13]. Based on [6], Song et al. [14] adopted sparse error correction (SEC) model to minimize the energy of the errors and proposed SEC-DKSVD for iris recognition. By introducing Fisher criterion, Zheng and Tao [15] proposed Fisher discriminative KSVD (FD-KSVD) to learn an over complete discriminative dictionary and an optimal linear classifier for image classification and Dong et al. [16] proposed a supervised dictionary learning algorithm for action recognition in still images. Methods in the second category try to learn a dictionary whose atoms have correspondence to the class labels. Sprechmann and Sapiro [17] learnt a class-specific dictionary with sparse coefficients and applied it to signal clustering. Ramirez et al. [18] introduced an incoherence term between the dictionaries associated with different classes, which encourages the dictionaries belonged different classes to be as independent as possible. Based on [18], Wang et al. [19] proposed a class-specific DL method for sparse modeling in action recognition. Yang et al. [20] introduced Fisher discrimination both in the sparse coding coefficients and class-specific representations. Castrodad and Sapiro [21] learnt a set of action-specific dictionaries with non-negative penalty on both dictionary atoms and representation coefficients. Instead of adding the sparsity constraint on the coefficients of each input signal, Meng et al. [22] proposed to learn a dictionary under global sparsity constraint to fittingly assign the atoms of the dictionary to represent signals. In the third category, a hybrid dictionary which contains a set of class-specific dictionaries and a shared dictionary were learnt. Kong and Wang [8] learnt a hybrid dictionary by introducing an incoherence penalty term to the class-specific sub-dictionaries. Zhou et al. [23] learnt a hybrid dictionary with a Fisher-like regularization on the coding coefficients.

Although these DL methods have achieved good performance, the label and the property of dictionary atom are pre-given, which may not be able to reflect the practical data structure accurately, i.e., particularity and commonality shared by multiple classes. In addition, how to balance the size of each class-specific part in the

http://dx.doi.org/10.1016/j.neucom.2016.01.026
0925-2312/

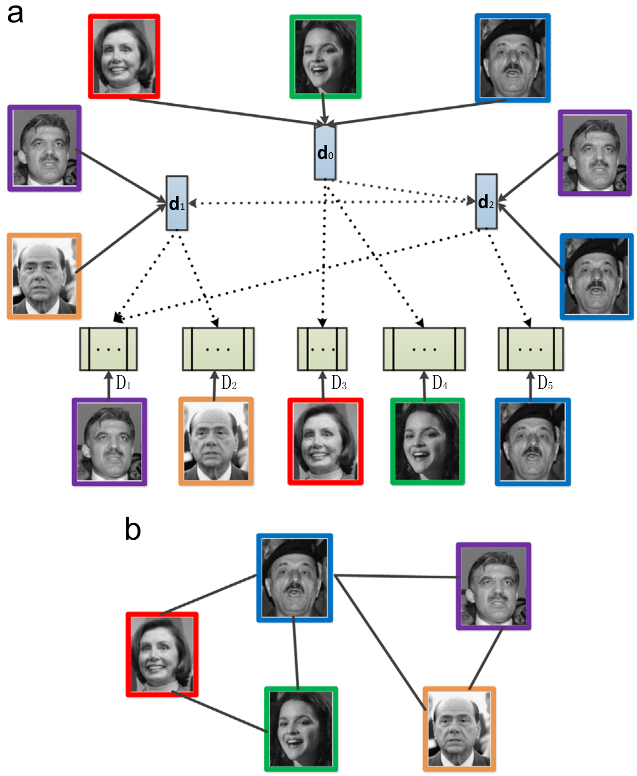

Fig. 1. (a) The relationships between 5 persons and dictionary atoms learnt on LFW database. (b) The graph structure between class labels.

dictionary is not a trivial task. Yang et al. [24] proposed a latent dictionary learning (LDL) method to build the relationship between dictionary atoms and class labels by learning a real-number latent matrix. Although the promising performance has been reported, the real-number latent matrix learnt by LDL could not represent which classes a dictionary atom belongs to. The label and the property of each dictionary atom were still blurry.

In this paper, we propose to learn a dictionary where the label and the property of each dictionary atom can be learnt adaptively. Although high-order group structure among data has been explored by building a probabilistic model in [25,26], how to define the relationship between class labels and dictionary atoms is seldom discussed. In order to avoid predefining the relationships between dictionary atoms and class labels, we proposed to learn a structure adaptive dictionary with a binary matrix to indicate the relationships. As shown in Fig. 1(a), dictionary atom $\mathbf{d}_0$, $\mathbf{d}_1$ and $\mathbf{d}_2$ are shared by more than one class, and $\mathbf{D}_1,\ldots,\mathbf{D}_5$ are class-specific sub-dictionary. In traditional dictionary learning methods, the shared atoms (e.g., $\mathbf{d}_0$, $\mathbf{d}_1$ and $\mathbf{d}_2$) are assumed to be shared by all classes. The size of each sub-dictionary (shared sub-dictionary and class-specific sub-dictionary) is also predefined. In SADL, it is easy to see that the shared dictionary atom $\mathbf{d}_0$ is shared by 3 persons; $\mathbf{d}_1$ and $\mathbf{d}_2$ are shared by 2 persons respectively. At the same time, $\mathbf{d}_1$ and $\mathbf{d}_2$, $\mathbf{d}_0$ and $\mathbf{d}_2$ are shared by one person respectively (violet box and turquoise box respectively). Besides, the size of each class-specific sub-dictionary $\mathbf{D}_1,\ldots,\mathbf{D}_5$ and shared sub-dictionary $[\mathbf{d}_0,\mathbf{d}_1,\mathbf{d}_2]$ can be adaptively learned during dictionary learning. Compared to LDL, the binary relationship is more meaningful because it clearly shows which class or classes a dictionary atom is associated with. On the other hand, to enhance the discrimination of the dictionary, a discrimination term based on Fisher discrimination criterion is applied on the coefficients. A weighted dictionary coherence term is also added to reduce the correlation of dictionary atoms between different classes. Compared with other supervised DL methods, SADL has two advantages. Firstly, SADL can adaptively adjust the label and property of each dictionary atom, which results in a more compact dictionary. Secondly, SADL can construct a graph structure between class labels by exploring the class relation among dictionary atoms, shown in Fig. 1(b). Extensive experiments on various image classification tasks, such as face recognition, gender classification and action recognition, showed the proposed SADL can achieve competitive performance with those state-of-the-art dictionary learning methods.

The reminder of this paper is organized as follows. Related supervised dictionary learning methods are introduced in Section 2. Then in Section 3 we present the proposed structure adaptive dictionary learning method. In Section 4, we evaluate the performance of SADL on various image classification tasks. Finally, we conclude this paper in Section 5.

2. Related work

2.1. All-class-shared dictionary learning

In all-class-shared dictionary learning, a dictionary shared by all classes is learnt. In order to improve the classification ability, it is popular to learn a shared dictionary and a classifier on the coefficients jointly. Bach et al. [10] proposed to learn a discriminative dictionary with a logistic function simultaneously. Venkatesh and Pham [11] proposed to learn a dictionary with linear classifier for both the labeled and unlabeled data. Inspired by the work of [11], Zhang and Li [6] proposed discriminative KSVD (DKSVD) for face recognition. Jiang et al. [13] proposed Label-Consistent KSVD (LCKSVD) by adding a label consistency regularization. The optimization problem can be written as

$$\min_{\mathbf{D},\mathbf{Z}} \sum_{i=1}^{C} (\|\mathbf{X}_i - \mathbf{D}\mathbf{Z}_i\|_F^2) + f_{\mathbf{Z}}(\mathbf{Z}) \quad \text{s.t.} \ \|\mathbf{d}_m\|_2^2 = 1 \quad \forall m \tag{1}$$

where $\mathbf{D}=[\mathbf{d}_1,\mathbf{d}_2,\ldots,\mathbf{d}_K]\in\mathcal{R}^{d\times K}$, $f_{\mathbf{Z}}(\mathbf{Z})$ is a function on $\mathbf{Z}$.

2.2. Class-specific dictionary learning

In class-specific dictionary learning, the atoms in the learnt dictionary $\mathbf{D}=[\mathbf{D}_1,\ldots,\mathbf{D}_C]$ have class labels correspond to the subject classes, where $\mathbf{D}_i\in\mathcal{R}^{d\times K_i}$ is the sub-dictionary corresponding to class i. Sprechmann and Sapiro [17] learnt a dictionary for each class and applied it to signal clustering. Ramirez et al. [18] introduced an incoherence term between the dictionaries associated with different classes, which encourages the dictionaries belonged to different classes to be independent as much as possible. Based on [18], Wang et al. [19] proposed a class-specific dictionary learning method for sparse modeling in action recognition. Yang et al. [20] introduced Fisher discrimination both in the sparse coding coefficients and class-specific representations to make both the representation residual and the representation coefficients be discriminative. The dictionary $\mathbf{D}$ can be learnt class by class:

$$\min_{\mathbf{D},\mathbf{Z}} \sum_{i=1}^{C} (\|\mathbf{X}_i - \mathbf{D}_i\mathbf{Z}_i\|_F^2) + f_{\mathbf{Z}}(\mathbf{Z}) + f_{\mathbf{D}}(\mathbf{D}) \quad \text{s.t.} \ \|\mathbf{d}_m\|_2^2 = 1 \quad \forall m \tag{2}$$

where $\mathbf{D}=[\mathbf{D}_1,\mathbf{D}_2,\ldots,\mathbf{D}_C]\in\mathcal{R}^{d\times K}$, $\mathbf{D}_i=[\mathbf{d}_1^i,\ldots,\mathbf{d}_{K_i}^i]$, $K=\sum_{i=1}^{C}K_i$, $f_{\mathbf{Z}}(\mathbf{Z})$ and $f_{\mathbf{D}}(\mathbf{D})$ are functions on $\mathbf{Z}$ and $\mathbf{D}$, respectively.

2.3. Hybrid dictionary learning

In hybrid dictionary learning, a shared dictionary $\mathbf{D}_{C+1}$ by all classes and a set of class-specific dictionaries $\mathbf{D}_i, i\in\{1,..,C\}$ were learnt. Kong and Wang [8] learnt a hybrid dictionary by introducing an incoherence penalty term to the class-specific sub-