Pattern Recognition 47 (2014) 885–898

Contents lists available at ScienceDirect

Pattern Recognition

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

A classification-oriented dictionary learning model: Explicitly learning the particularity and commonality across categories

Donghui Wang*, Shu Kong [1]

Department of Computer Science and Technology, Zhejiang University, Hangzhou 310027, PR China

ARTICLE INFO

Article history:
Received 14 July 2012
Received in revised form
8 January 2013
Accepted 2 August 2013
Available online 12 August 2013

Keywords:
Dictionary learning
Sparse coding
Image classification
Particularity
Commonality

ABSTRACT

Empirically, we find that despite the most exclusively discriminative features owned by one specific object category, the various classes of objects usually share some common patterns, which do not contribute to the discrimination of them. Concentrating on this observation and motivated by the success of dictionary learning (DL) framework, in this paper, we propose to explicitly learn a class-specific dictionary (called particularity) for each category that captures the most discriminative features of this category, and simultaneously learn a common pattern pool (called commonality), whose atoms are shared by all the categories and only contribute to representation of the data rather than discrimination. In this way, the particularity differentiates the categories while the commonality provides the essential reconstruction for the objects. Thus, we can simply adopt a reconstruction-based scheme for classification. By reviewing the existing DL-based classification methods, we can see that our approach simultaneously learns a classification-oriented dictionary and drives the sparse coefficients as discriminative as possible. In this way, the proposed method will achieve better classification performance. To evaluate our method, we extensively conduct experiments both on synthetic data and real-world benchmarks in comparison with the existing DL-based classification algorithms, and the experimental results demonstrate the effectiveness of our method.

1. Introduction

Dictionary learning (DL), as a particular sparse signal model, aims to learn a set of atoms, or called visual words in the computer vision community, in which a few atoms can be linearly combined to well approximate a given signal. From the view of compressive sensing, it is originally designed to learn an adaptive codebook to faithfully represent the signals with sparsity constraint. In recent years, researchers have applied DL framework to other applications and achieved state-of-the-art performances, such as image denoising [1] and inpainting [2], clustering [3–5], classification [6–8].

In this paper, we focus on the classification task based on dictionary learning framework, and propose a novel model to learn a compact and discriminative dictionary for classification. It is well-known that the conventional DL framework is not adapted to classification, as the learned dictionary is used for signal reconstruction. Therefore, to borrow the powerfulness in representation of signals and to circumvent this problem, researchers have developed several approaches to learn a classification-oriented dictionary in a supervised learning fashion. By exploring the label information, the DL-based classification methods learn the classification-oriented dictionary mainly in two ways/tracks as summarized in [6]: either directly forcing the dictionary discriminative, as illustrated by Fig. 1(a), or making the sparse coefficients discriminative (usually through simultaneously learning a classifier) to promote the discrimination power of the dictionary, as demonstrated by Fig. 1(b). We summarize some representative methods of different tracks in Table 1. Note that our method, as well as Fisher discrimination DL (FDDL) method [9], inherits the advantages of both scenarios as explained by Fig. 1(c), thus we classify FDDL and our method to Track-III.

Track-I originates from sparse representation based classification (SRC) method [14], which uses the original training images as a predefined dictionary and achieves very encouraging performances. However, this predefined dictionary can be very large when the number of training data increases, and is not effective enough to represent the query images due to the existing uncertain and noisy information in the original images. For this reason, Yang et al. propose a method called Metaface learning [10] to learn a class-specific dictionary for each class, thus the dictionary becomes more compact and more discriminative. Based on this, Ramirez et al. go further by proposing a structural incoherent dictionary learning method (DLSI), which advocates learning class-specific sub-dictionaries for each class with a structural incoherence penalty term to make the sub-dictionaries as independent as possible [11]. Despite

* Corresponding author. Tel./fax: +86 571 87951916.
E-mail addresses: dhwang@zju.edu.cn, wdh.zju@gmail.com (D. Wang), aimerykong@zju.edu.cn (S. Kong).

[1] Joint first authors, *i.e.* the authors contributed equally to this work.

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.08.004

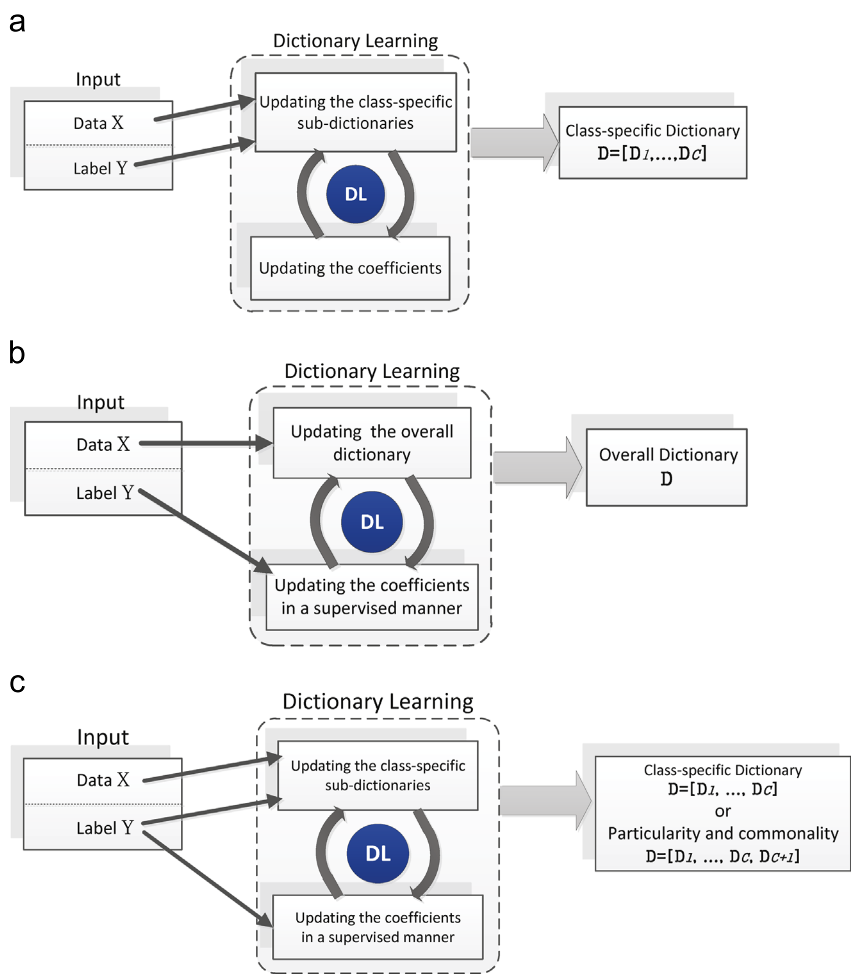

Fig. 1. (a) Methods from Track-I employ the labeling information to learn class-specific sub-dictionaries, by which way, the discrimination power of the overall dictionary concatenated by the sub-dictionaries is improved. (b) Methods from Track-II apply the labeling information to a criterion on the coefficients, such as logistic regression [8] and linear classifier [12], and thus propagate the discrimination power to the dictionary. (c) FDDL [9] and our method use the label information on the updating of both the dictionary and the coefficients, in which way, better discrimination performance in dictionary learning framework can be anticipated.

Table 1
Some representative algorithms from the two tracks of DL-based classification methods. Our method is different from the two scenarios, thus we put it to the third scenario, *i.e.* Track-III. Besides, FDDL [9] also belongs to Track-III.

Scenario	Representative methods
Track-I	Metaface learning [10], DLSI [11]
Track-II	SupervisedDL [8], D-KSVD [12], LC-KSVD [13]
Track-III	FDDL [9], our method

the improvements of meta-face learning and DLSI, both of them ignore the benefit of making the coefficients discriminative (see Fig. 1(a)), as the Track-II does. Thus, further enhancement on the classification performance is prevented. As for classification scheme, the methods of Track-I usually exploit reconstruction-based classifier for the final classification, *i.e.* reconstructing the novel signal by each class-specific dictionary and identifying the signal to the class whose sub-dictionary produces the smallest reconstruction error.

Most methods belong to Track-II, *i.e.* driving the sparse coefficients more discriminative to enhance the discrimination power of the overall dictionary, as illustrated by Fig. 1(b). Mairal et al. propose a supervised DL method by adding a logistic loss function on the sparse coefficients to the dictionary learning framework [8]. Their method simultaneously learns a classifier, and achieves very impressive performances in hand-written digit recognition and texture classification. Zhang and Li propose a discriminative K-SVD method (D-KSVD), which embeds a linear classifier on the sparse coefficients into the dictionary learning framework. D-KSVD produces a desired dictionary which has good representation power, and in the meantime supports optimal discrimination of the classes [12], and brings out fairly good performance in face recognition. Furthermore, Jiang et al. add a label consistence term on D-KSVD (LC-KSVD) to bridge the labeling information and the dictionary atoms, thus driving the sparse coefficients more discriminative for further improvement of the performance. LC-KSVD achieves impressive results on face recognition and object classification. However, these methods purely treat the sparse coefficients as a new feature representation of the original images, and intuitively makes them discriminative to improve the classification performance which is carried out on the coefficients. Obviously, merely driving the coefficients discriminative is not enough for better classification performance. Thus, if we incorporate the process of forcing the dictionary discriminative as described previously, we can obtain a more desirable dictionary learning method.

Out of these two scenarios, Yang et al. propose Fisher discrimination DL (FDDL) method to simultaneously learn class-specific sub-dictionaries and make the coefficients more discriminative based on Fisher criterion [9]. From Fig. 1(c), we can see FDDL is such a method that inherits the advantages of the two scenarios as presented above. Similarly, our approach learns the classification-oriented dictionary by borrowing the benefit of the two scenarios