



# Learning low-rank and discriminative dictionary for image classification<sup>☆</sup>

Liangyue Li<sup>1</sup>, Sheng Li<sup>1</sup>, Yun Fu<sup>\*</sup>

Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA



## ARTICLE INFO

### Article history:

Received 1 June 2013

Received in revised form 25 October 2013

Accepted 21 February 2014

Available online 5 March 2014

### Keywords:

Sparse representation

Dictionary learning

Low-rank regularization

Image classification

## ABSTRACT

Dictionary learning plays a crucial role in sparse representation based image classification. In this paper, we propose a novel approach to learn a discriminative dictionary with low-rank regularization on the dictionary. Specifically, we apply Fisher discriminant function to the coding coefficients to make the dictionary more discerning, that is, a small ratio of the within-class scatter to between-class scatter. In practice, noisy information in the training samples will undermine the discriminative ability of the dictionary. Inspired by the recent advances in low-rank matrix recovery theory, we apply low-rank regularization on the dictionary to tackle this problem. The iterative projection method (IPM) and inexact augmented Lagrange multiplier (ALM) algorithm are adopted to solve our objective function. The proposed discriminative dictionary learning with low-rank regularization ( $D^2L^2R^2$ ) approach is evaluated on four face and digit image datasets in comparison with existing representative dictionary learning and classification algorithms. The experimental results demonstrate the superiority of our approach.

Published by Elsevier B.V.

## 1. Introduction

Sparse representation has been extensively studied in recent years due to its promising performance [26,55,36]. Given a test signal and an over-complete dictionary with prototype signals as atoms, sparse representation seeks a sparsest representation of the test signal among all the linear combinations of the dictionary atoms. As a result, sparse representation can reveal the underlying structure of high dimensional signals. This is especially significant in the big data age when massive high dimensional data (image, video, web, bioinformatic data, etc.) are emerging which require fast processing. Sparse representation has been applied to many problems, ranging from speech denoising [19] to super resolution [54], and from blind source separation [31] to bioinformatics [56]. In this paper, we focus on image classification based on sparse representation which demands the representation to be discriminative as well. Wright et al. [51] propose sparse representation based classifier (SRC) for face recognition based on sparse signal representation theory. This sparsity has been supported by research in human visual system which found that nerve cells in the connecting pathway only react to a certain amount of stimuli [43].

The essence of sparse representation is to recover a signal from a small number of linear measurements. Given an over-complete dictionary  $D$  and a query sample  $y$ , the problem can be formulated as the following objective function:

$$\min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad y = D\alpha, \quad (1)$$

where  $\alpha$  is the coding coefficient whose non-zero elements are those corresponding to the category  $y$  belongs to. Here the dictionary  $D$  can either be pre-specified or gradually adapted to fit the training samples given. Wright et al. [51] pre-specified the dictionary as the original training samples. A problem with this strategy is that the original images in the training set may not faithfully represent the test samples due to the noise and uncertainty in it. Besides, the distinctive message resided in the training set might be ignored in this way and the dictionary cannot guarantee the sparse property. As a result, we need to learn the dictionary adaptively from the specific training sample set.

Research progress on dictionary learning has been made on learning a well adapted dictionary for discriminative representation of test samples. Generalizing K-means clustering process, Aharon et al. [1] presented K-SVD algorithm to learn an over-complete dictionary by updating dictionary atoms and sparse representations iteratively. Recently, a discriminative K-SVD method that considers classification error when learning the dictionary was proposed [58]. Jiang et al. [21] associated label information with each dictionary

<sup>☆</sup> This paper has been recommended for acceptance by Lijun Yin, Ph.D.

<sup>\*</sup> Corresponding author. Tel.: +1 617 373 7328.

E-mail addresses: [liangyue@ece.neu.edu](mailto:liangyue@ece.neu.edu) (L. Li), [shengli@ece.neu.edu](mailto:shengli@ece.neu.edu) (S. Li),

[yunfu@ece.neu.edu](mailto:yunfu@ece.neu.edu) (Y. Fu).

<sup>1</sup> These authors contributed equally to this work.

atom to enforce discriminability. To reduce computational complexity, Lee et al. and Wang et al. [26,50] emphasized specific discriminative criteria to learn an over-complete dictionary. In Ref. [55], the authors introduced the Fisher criterion to learn a structured dictionary. Studer and Baraniuk [48] investigated dictionary learning from sparsely corrupted signals. However, the methods above can only work well on clean training samples or with small noise and sparse corruption. Imagine the training samples are corrupted with large noise, then in order for representing the training samples, the dictionary atoms will also get corrupted.

Recent advances in low-rank learning for the purpose of visual representation have shown excellent performance for handling large noise. Given a matrix  $M$  of low rank, matrix completion aims at recovering it from noisy observations of a random small portion of its elements. It has been proved that under certain assumptions, the problem can be exactly solved and several methods have been proposed [23,5,6]. In our case of dictionary learning for image classification, training samples in the same class are linearly correlated and lie in a low dimensional manifold. Therefore, a sub-dictionary for representing samples from one class should reasonably of low rank. Ma et al. (2012) integrated rank minimization into sparse representation and achieved impressive face recognition results especially when corruption existed.

Inspired by the previous work, we aim at learning a discriminative dictionary for image classification that can handle training samples corrupted with large noise. We propose a discriminative dictionary learning with low-rank regularization ( $D^2L^2R^2$ ) approach and illustrate it in Fig. 1. In the figure, the training sample set can be approximately recovered by the multiplication of the dictionary and the coding coefficient matrix. Each sub-dictionary is of low rank (can be seen as multiplication of two matrices of smaller size) to reduce the negative effect of noise contained in training samples. The coding coefficients conform to Fisher discrimination criterion. Benefiting from the above design, our approach has the following advantages. First, the Fisher discriminant function can help us achieve a small ratio of the within-class scatter to between-class scatter on the coefficients, making the dictionary learned has strong discerning power. Second, low-rank regularization will output a compact and pure dictionary that can reconstruct the denoised images even when the training samples are contaminated.

Unlike FDDL proposed in Ref. [55], our  $D^2L^2R^2$  approach can well cope with training samples with large noise and can still achieve impressive performance due to the low-rank regularization on the sub-dictionaries. Our approach also differs from the recently

proposed DLRD\_SR algorithm proposed in Ref. [36]. Though DLRD\_SR was claimed to be able to handle noisy samples as well, it may suffer from certain information loss because of the low-rank regularization, our  $D^2L^2R^2$  approach compensates this by enforcing the Fisher criterion on the coding coefficients of the training sets.

This paper is a substantial extension of our previous conference paper [30]. Compared with Ref. [30], we give more details of our approach in this paper, and extensive experimental results are reported. In addition, we introduce the background of our work more comprehensively, and add more discussions in methodology and experimental parts. Beyond face recognition applications in Ref. [30], we also evaluate the performance of our approach on digit recognition. Thus, this is a more systematic and comprehensive paper of our work. The rest of this paper is organized as follows. Section 2 gives a brief review of some related work. Section 3 introduces our discriminative dictionary learning with low-rank regularization ( $D^2L^2R^2$ ) approach. Section 4 presents the optimization algorithm for our model. Section 5 describes the classification scheme. Section 6 shows experimental results on several image datasets. Finally, we draw conclusions in Section 7.

## 2. Related work

In this section, we briefly review related work on sparse representation, dictionary learning and low-rank learning.

### 2.1. Sparse representation

We briefly review robust face recognition using sparse representation proposed in Ref. [51]. Define matrix  $Y$  as the entire training set which consists of  $n$  training samples from all  $c$  different classes:  $Y = [Y_1, Y_2, \dots, Y_c]$  where  $Y_i \in \mathbb{R}^{d \times n_i}$  is all the training samples from  $i$ -th class,  $d$  is the dimension of samples, and  $n_i$  is the number of samples from  $i$ -th class. To classify a test sample  $y$ , we need to go through two phases: coding and classification.

(a) Coding phase: we obtain the coding coefficient of  $y$  by solving the following minimization problem:

$$\alpha = \operatorname{argmin} \|\alpha\|_0 \quad \text{subject to} \quad \|y - Y\alpha\|_2 \leq \epsilon. \quad (2)$$

The model seeks the sparsest representation for  $y$  among all the possible linear combinations of the atoms in the dictionary. For general

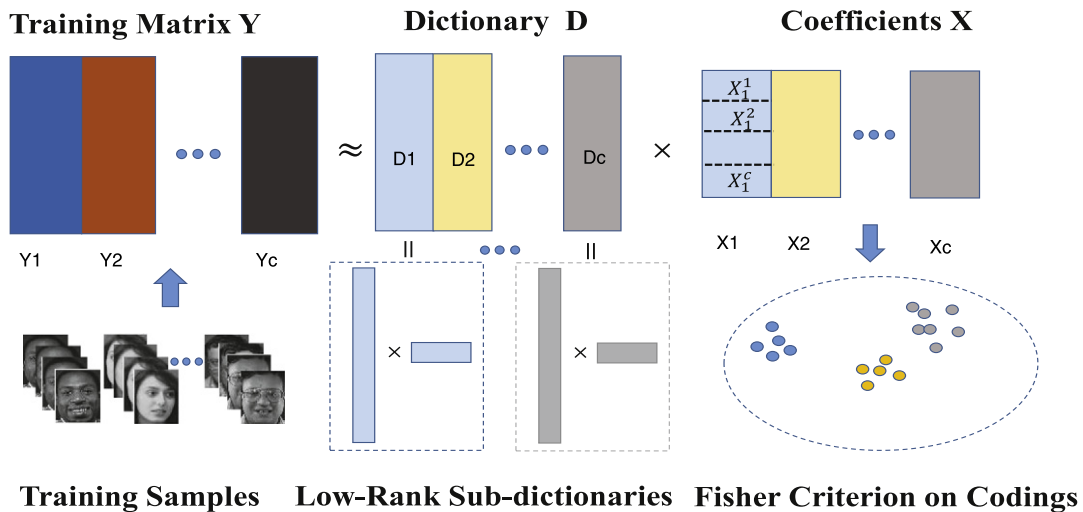


Fig. 1. Illustration of  $D^2L^2R^2$  model. Each sub-dictionary learned is of low rank, which can reduce the negative effect of noise contained in training samples. The coding coefficients conform to Fisher discrimination criterion, making the dictionary discriminative for the training samples.

Download English Version:

<https://daneshyari.com/en/article/528420>

Download Persian Version:

<https://daneshyari.com/article/528420>

[Daneshyari.com](https://daneshyari.com)