Neurocomputing 204 (2016) 198–210

Contents lists available at ScienceDirect

Neurocomputing

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

Face recognition using class specific dictionary learning for sparse representation and collaborative representation

Bao-Di Liu [a,*], Bin Shen [b], Liangke Gui [c], Yu-Xiong Wang [c], Xue Li [e], Fei Yan [d], Yan-Jiang Wang [a]

[a] College of Information and Control Engineering, China University of Petroleum, Qingdao 266580, China
[b] Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA
[c] School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[d] Lijin County Party Committee Office, Dongying 257499, China
[e] Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 2 March 2015
Received in revised form
7 July 2015
Accepted 20 August 2015
Available online 8 April 2016

Keywords:
Class specific dictionary learning
Sparse representation
Collaborative representation
Face recognition

ABSTRACT

Recently, sparse representation based classification (SRC) and collaborative representation based classification (CRC) have been successfully used for visual recognition and have demonstrated impressive performance. Given a test sample, SRC or CRC formulates its linear representation with respect to the training samples and then computes the residual error for each class. SRC or CRC assumes that the training samples from each class contribute equally to the dictionary in the corresponding class, i.e., the dictionary consists of the training samples in that class. This, however, leads to high residual error and instability. To overcome this limitation, we propose a class specific dictionary learning algorithm. To be specific, by introducing the dual form of dictionary learning, an explicit relationship between the basis vectors and the original image features is represented, which also enhances the interpretability. SRC or CRC can be thus considered as a special case of the proposed algorithm. Blockwise coordinate descent algorithm and Lagrange multipliers are then adopted to optimize the corresponding objective function. Extensive experimental results on five benchmark face recognition datasets show that the proposed algorithm achieves superior performance compared with conventional classification algorithms.

1. Introduction

Face recognition is a classical yet challenging research topic in computer vision and pattern recognition [33]. Two stages are usually considered for effective face recognition: (1) feature extraction, (2) classifier construction and label prediction. For the first stage, Turk and Pentland [25] proposed eigenfaces by performing principal component analysis (PCA). He *et al.* [8] proposed laplacianfaces to preserve local information. Belhumeur *et al.* [2] suggested fisherfaces to maximize the ratio of between-class scatter to within-class scatter. Yan *et al.* [28] proposed a multi-subregion based correlation filter bank algorithm to extract both the global-based and local-based face features. For the latter stage, Richard *et al.* [19] proposed a nearest neighbor method to predict the label of a test image using its nearest neighbors in the training samples. Ho *et al.* [9] and Tao *et al.* [23] proposed a nearest subspace method to assign the label of a test image by comparing its reconstruction error for each category.

Under the nearest subspace [44, 45] framework, Wright *et al.* [27] proposed a sparse representation based classification (SRC) system and achieved impressive performance. Given a test sample, sparse representation techniques represent it as a sparse linear combination of the training samples. The predicted label is determined by the residual error from each class. Different from traditional decomposition frameworks like PCA, non-negative matrix factorization [39], and low-rank factorization [40], SRC allows coding under over-complete bases, and thus makes the attained sparse codes capable of representing the data more adaptively and flexibly. To analyze SRC, Zhang *et al.* [31] proposed collaborative representation based classification (CRC) as an alternative approach. CRC represents a test sample as the linear combination of almost all the training samples. They found that it is the collaborative representation rather than the sparse representation that makes the nearest subspace method powerful for classification. SRC, CRC, and their variants have been also used in other visual data sensing and analysis tasks, such as image

* Corresponding author. Tel.: +86 15764217948.
E-mail addresses: thu.liubaodi@gmail.com (B.-D. Liu), stanshenbin@gmail.com (B. Shen), guiliangke@gmail.com (L. Gui), yuxiongw@cs.cmu.edu (Y.-X. Wang), lixue421@gmail.com (X. Li), 232842281@qq.com (F. Yan), yjwang@upc.edu.cn (Y.-J. Wang).

http://dx.doi.org/10.1016/j.neucom.2015.08.128
0925-2312/

classification [12], image inpainting [41], object detection [42], image annotation [1], and transfer learning [43].

Despite their promise, both SRC and CRC algorithms directly use the training samples as the dictionary for each class. By contrast, a well learned dictionary, especially by enforcing some discriminative criteria, can reduce the residual error greatly and achieve superior performance for classification tasks. Existing discriminative dictionary learning approaches are mainly categorized into three types: shared dictionary learning, class specific dictionary learning, and hybrid dictionary learning. In shared dictionary learning, each basis is associated to all the training samples. Mairal *et al.* [16] proposed to learn a discriminative dictionary with a linear classifier of coding coefficients. Liu *et al.* [15] learned a Fisher discriminative dictionary. Liu *et al.* [35, 36, 37] and Yu *et al.* [38] presented a graph embedded dictionary learning method. Zhang and Li [32] proposed a joint dictionary learning algorithm for face recognition. In class specific dictionary learning, each basis only corresponds to a single class so that the class specific reconstruction error could be used for classification. Yang *et al.* [30] learned a dictionary for each class with sparse coefficients and applied it for face recognition. Sprechmann and Sapiro [22] also learned a dictionary for each class with sparse representation and used it in signal clustering. Castrodad and Sapiro [4] learned a set of action specific dictionaries with non-negative penalty on both dictionary atoms and representation coefficients. Wang *et al.* [26] introduced mutual incoherence information to promote class specific dictionary learning in action recognition. Yang *et al.* [29] embedded the Fisher discriminative information into class specific dictionary learning.

The shared dictionary learning approaches usually lead to a dictionary of small size and the discriminative information (i.e., the label information corresponding to coding coefficients) is embedded into the dictionary learning framework. The class specific dictionary learning approaches usually focus on the classifier construction aspect since each basis vector is fixed to a single class label. The combination of shared basis vectors and class specific basis vectors is then learned in hybrid dictionary learning. Zhou *et al.* [34] learned a hybrid dictionary with Fisher regularization on the coding coefficient. Gao *et al.* [6] learned a shared dictionary to encode common visual patterns and learned a class specific dictionary to encode subtle visual differences among different categories for fine-grained image representation. Liu *et al.* [12] proposed a hierarchical dictionary learning method to produce a shared dictionary and a cluster specific dictionary. In spite of the demonstrated performance of hybrid dictionary learning, it is still a challenge to balance between the shared dictionary and the class specific dictionary.

In this paper, motivated by the superior performance of the SRC and CRC algorithms and the class specific dictionary learning method, we propose class specific dictionary learning (CSDL) for both sparse representation based classifier (CSDL-SRC) and collaborative representation based classifier (CSDL-CRC). Fig. 1 shows the framework of our proposed CSDL. The major distinction between our approach and the existing class specific dictionary learning methods is that the existing methods directly optimize the dictionary basis vectors (the "primal" form), whereas we leverage a "dual" reformulation of dictionary learning and optimize the weights instead. Compared with the "primal" form widely used by the existing methods, our novel "dual" form offers several benefits:

- It provides an explicit relationship between the basis vectors and the original image features, thus enhancing the interpretability of the learned dictionary.
- It is easy to be kernelized due to the separation of original data, which is difficult for the existing methods. The generalization to kernel spaces will be elaborated in Section 5.7.3.
- Most of the existing class specific dictionary learning methods focus on introducing additional regularization terms, which could be easily incorporated into our dual formulation of class specific dictionary learning to further improve the performance.

Our main contributions are threefold:

- We propose a novel class specific dictionary learning scheme that considers the weight of each sample when generating the dictionary (i.e., subspace). The traditional CRC and SRC methods perform face recognition without training procedures (i.e., the training samples are directly used for predicting the labels). By contrast, our proposed method compensates this deficiency by introducing class specific dictionary learning. It is applicable to both CRC and SRC. Furthermore, it is reasonable to assume additionally that different samples contribute unevenly in constructing the corresponding subspace. CRC or SRC can be thus viewed as special cases of our proposed algorithm.
- We propose the dual form of dictionary learning to enhance the interpretability.

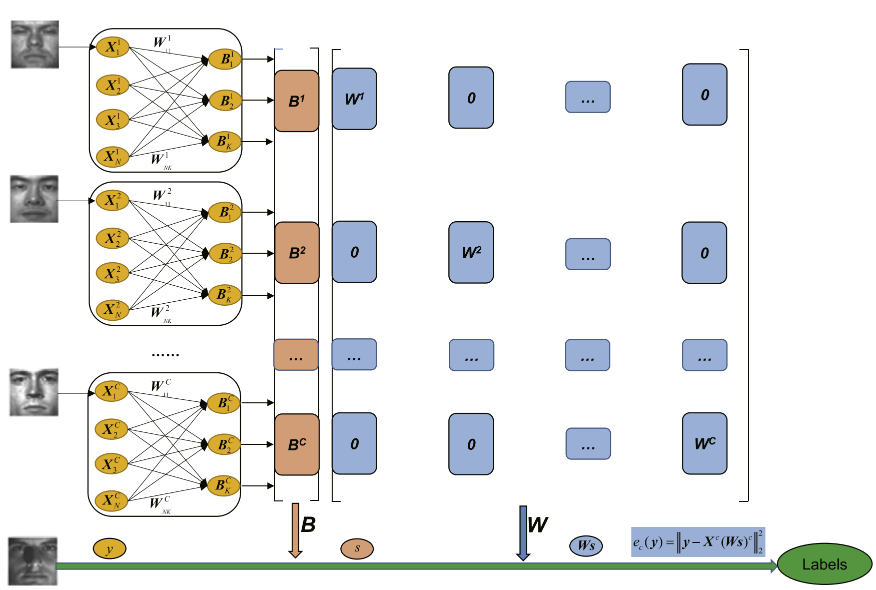

Fig. 1. The framework of our proposed class specific dictionary learning algorithm.