Neurocomputing 198 (2016) 114–124

Contents lists available at ScienceDirect

Neurocomputing

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

Improved sparse representation with low-rank representation for robust face recognition

Chun-Hou Zheng, Yi-Fu Hou, Jun Zhang *

College of Electrical Engineering and Automation, Anhui University, Hefei, China

ARTICLE INFO

Article history:
Received 15 March 2015
Received in revised form
25 June 2015
Accepted 8 July 2015
Available online 24 March 2016

Keywords:
Low rank representation
HOG feature
Eigenface
Sparse representation
Face recognition

ABSTRACT

In this paper, an approach to learn a robust sparse representation dictionary for face recognition is proposed. As well known, sparse representation algorithm can effectively tackle slight occlusion problems for face recognition. However, if images are corrupted by heavy noise, performance will be not guaranteed. In this paper, to enhance the robustness of sparse representation to serious noise in face images, we integrate low rank representation into dictionary learning to alleviate the influence of unfavorable factors such as large scale noise and occlusion. Among which we extract eigenfaces by singular value decomposition (SVD) from the low rank pictures to reduce dictionary atoms and, thereby, optimize the efficiency of improved algorithm. Otherwise, we characterize each image using the histogram of orientated gradient (HOG) feature which has been proven to be an effective descriptor for face recognition in particular. The performance of the proposed Low-rank and HOG feature based ESRC (LH_ESRC) algorithm on several popular face databases such as the Extended Yale B database and CMU_PIE face database shows the effectiveness of our method. In addition, we evaluate the robustness of our method by adding different proportions of randomly noise and block occlusion and real disgusts. Experimental results illustrate the benefits of our approach.

1. Introduction

Face recognition has attracted significant attention owing to the accessibility of inexpensive digital products and its applications in various domains such as public surveillance, multimedia. Recent years, many literatures about how to extract features and learn a classifier have been published [1,2,4-6,10,14,16,19,20], which are most important to face recognition. Extracting discriminative features and learning a suitable classifier directly associate with the classification accuracy and time consuming.

Extracting discriminative feature is the first step for face recognition. Until now, a large kind of feature extracting methods have been proposed in literatures [1,4,7,8,11,12] and display good performance, e.g. Eigenfaces [1], Local Binary Patterns (LBP) [3], Gabor wavelets [6] etc. Recently, for more robustness to variations in illumination, rotation and small displacements, Albiol et al. [7] combined the HOG descriptors with Elastic Bunch Graph Matching (EBGM) for face recognition. Higher accuracy was obtained on account of the property of HOG descriptors which is more robust to sparse error. To further explore the performance capability of HOG descriptors, Deniz et al. [8] proposed a simple but powerful approach to build a robust HOG descriptor, which leads to a great forward step to the application of HOG feature in face recognition. In addition, Hou et al. [11] jointed eigenface extraction and HOG feature to construct a compact dictionary for sparse representation, the comparison between several kinds of features certified the effectiveness of HOG feature in face recognition.

Moreover, learning a suitable classifier is also a key point for face recognition. In the past several years, sparse representation has made sharply progress in computer vision and has been successfully applied in compressed sensing [14] and image restoration [13]. Recently, sparse representation has also been applied to image classification such as face recognition [15,17] and texture recognition [18], with good performance. With the recent progress of l_1-norm minimization technique, sparse representation based classification has became a hot topic owing to the fact that an image vector can be well presented by the same class samples with a discriminative coefficient vector. At start, researchers connected pixels as off-the-shelf bases for dictionary learning. E.g. Wright et al. [16] cascaded pixels into vectors for constructing a dictionary and achieved good performance when the face images corrupted by slight noise. With the deep research in sparse representation, many researchers found that learning a dictionary is an effective way to improve the performance of signal reconstruction dramatically. Mairal et al. [19] jointly optimized an energy formulation with both sparse reconstruction and class discriminative components during

* Corresponding author.

http://dx.doi.org/10.1016/j.neucom.2015.07.146
0925-2312/

dictionary learning (DL). Yang et al. [20] presented a novel dictionary learning method based on the fisher discrimination criterion to improve the pattern classification performance. These algorithms work well when images are pure or corrupted by slight noise. If serious noise exists both in training images and test images, more robust dictionary optimization algorithms should be adopted. To alleviate the influence of noise, we can separate or suppress the corrupted pixels and extract better features for classification.

Recently, low-rank approximation [9,21,25] has been introduced and applied to subspace clustering and video segmentation, which motivates us to think whether it can be used to alleviate noise on face images? Candes et al. [22] asserted that we can recover the principal components of a data matrix even though a positive fraction of its entries are arbitrarily corrupted, and applied the proposed Robust PCA algorithm in the area of video surveillance where shadows can be effective removed. Liu et al. [23] generalized the Robust PCA and proposed low rank representation (LRR) to deal with the original atoms and utilized the low-rank coefficients to classify test samples. In this paper, to enhance the robustness of sparse representation to serious pollution exiting in face images, we integrate low rank representation into dictionary learning to alleviate the influence of unfavorable factors such as large scale noise and occlusion and even real mask disgusts.

For stand or fall of the dictionary, the compactness of constructed dictionary is also attractive [26,27,29]. Zheng et al. [28] decomposed genes expression data matrix into meta-samples matrix and expression pattern matrix and constructed a compact dictionary with meta-samples. Each test sample can be presented as the linear mixture of meta-samples. If we can extract these elements capturing the intrinsic structural information of each face space, an arbitrary face belonging to the same class can be perfectly presented by these elements known as eigenfaces. This will directly reduce the number of atoms in dictionary and, thereby, decrease computational complexity.

Based on the aforementioned analysis, to enhance the robustness of sparse representation, we integrate low rank representation into dictionary learning to alleviate the influence of unfavorable factors such as large scale noise and occlusion. Among which we extract eigenfaces by singular value decomposition (SVD) from the low rank pictures to reduce dictionary atoms and, thereby, optimize the efficiency of improved algorithm. Otherwise, we characterize each image using the histogram of orientated gradient (HOG) feature which has been proven to be an effective descriptor for face recognition in particular. The performance of the proposed Low-rank and HOG feature based ESRC (LH_ESRC) algorithm on several popular face databases shows the effectiveness of our method. In addition, we evaluate the robustness of our method by adding different proportions of randomly noise and block occlusion and real disgusts.

The rest of the paper is organized as follows: Section 2 introduces some related works about low-rank representation and HOG descriptor. Section 3 presents the proposed LH_ESRC algorithm in detail. Section 4 shows the experimental results and analysis. Section 5 concludes the paper and outlines the future work.

2. Related work

2.1. Low-rank representation (LRR)

In literature [22], Candes et al. points out that the principal components of a data matrix can be recovered even though the entries of data matrix are arbitrarily corrupted. The proposed Robust PCA algorithm decomposes a dense but corrupted matrix (i.e.,A) into a low-rank part (i.e.,L) and a large but sparse error part (i.e.,E). Mathematically, it can be expressed as follows:

$$A = L + E, \tag{1}$$

where L and E are unknown but L is a low rank matrix and E is sparse. We reformulated (1) as the following optimization function:

$$\min_{L,E}\ rank(L) + \lambda\|E\|_0 \quad s.t. \quad A = L + E, \tag{2}$$

where parameter λ indicates the percentage of sparse error.

Unfortunately, formula (2) implicitly assumes that matrix A contains a single low-rank structure, which does not match with the face recognition system which usually grouped by multiple subspaces. A more general version of (2) can be presented as follows:

$$\min_{Z^*,E^*} rank(Z) + \lambda\|E\|_0 \quad s.t. \quad A = DZ + E, \tag{3}$$

where D is a dictionary. We usually set $D = A$. Z^* is the low-rank representation of data A (consisting of multi-class data) with respect to D. If we set $A = I$, the formula (3) falls back to (2). DZ^* is the low rank recovery of original data A. The Augmented Lagrange Multiplier (ALM) method can be used to solve optimization problem (3) (Fig. 1).

2.2. ALM method for LRR

The ALM method is adopted to solve the above optimization problem (3). Owing to highly non-convex of formula (3), we relax it by replacing the rank function and the l_0-norm with the nuclear norm and l_1-norms, respectively. That is:

$$\min_{Z^*,E^*} \|Z\|_* + \lambda\|E\|_1 \quad s.t. \quad A = DZ + E, \tag{4}$$

Here, the l_1-norm characterizes the sparse error term E. To obtain optimal solution Z^* and E^*, we convert formula (4) to following equivalent problem:

$$\min_{P^*,E^*} \|P\|_* + \lambda\|E\|_1 \quad s.t. \quad A = DZ + E, \quad Z = P, \tag{5}$$

ALM method is adopted to solve formula (5) by introducing the augmented Lagrange function:

$$\rho(Z,P,E,Y_1,Y_2,u) = \|P\|_* + \lambda\|E\|_1 + tr\left(Y_1^T(A - DZ - E)\right) + tr\left(Y_2^T(Z - P)\right)$$
$$+ \frac{u}{2}\left(\|A - DZ - E\|_F^2 + \|Z - P\|_F^2\right), \tag{6}$$

where u is a positive parameter and has an upper bound. Here, we set $D = A$ and solve each variable P^*,Z^*,E^*, separately:

$$\begin{aligned} P^* &= \arg\ \min \|P\|_* + tr\left(Y_2^T(Z - P)\right) + \frac{u}{2}\|Z - P\|_F^2 \\ &= \arg\ \min \frac{1}{u}\|P\|_* + \frac{u}{2}\left\|P - (Z + Y_2/u)\right\|_F^2 \end{aligned} \tag{7}$$

$$Z^* = \left(I + A^T A\right)^{-1}\left(A^T(A - E) + P + \left(A^T Y_1 + Y_2\right)/u\right) \tag{8}$$

$$\begin{aligned} E^* &= \arg\ \min\ \lambda\|E\|_1 + tr\left(Y_1^T(A - AZ - E)\right) + \frac{u}{2}\|A - AZ - E\|_F^2 \\ &= \arg\ \min \frac{\lambda}{u}\|E\|_1 + \frac{u}{2}\left\|E - (A - AZ + Y_1/u)\right\|_F^2 \end{aligned} \tag{9}$$

Lagrange multipliers Y_1 and Y_2 are updated:

$$Y_1 = Y_1 + u(A - AZ - E) \tag{10}$$

$$Y_2 = Y_2 + u(Z - P) \tag{11}$$