Optik 125 (2014) 4845–4849

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A face recognition algorithm based on collaborative representation

Zhengming Li [a,b,*], Tong Zhan [a], Binglei Xie [c], Jian Cao [b], Jianxiong Zhang [a]

[a] Guangdong Industry Training Center, Guangdong Polytechnic Normal University, Guangzhou, China
[b] Bio-Computing Research Center, Harbin Institute of Technology, Shenzhen Graduate School, Shenzhen, China
[c] Shenzhen Key Laboratory of Urban Planning and Decision-Making Simulation, Shenzhen, China

ARTICLE INFO

Article history:
Received 3 September 2013
Accepted 5 April 2014

Keywords:
Face recognition
Sparse coding
Neighbor matrix
Collaborative representation

ABSTRACT

In this paper, we propose a face recognition algorithm by incorporating a neighbor matrix into the objective function of sparse coding. We first calculate the neighbor matrix between the test sample and each training sample by using the revised reconstruction error of each class. Specifically, the revised reconstruction error (*RRE*) of each class is the division of the l_2-norm of reconstruction error to the l_2-norm of reconstruction coefficients, which can be used to increase the discrimination information for classification. Then we use the neighbor matrix and all the training samples to linearly represent the test sample. Thus, our algorithm can preserve locality and similarity information of sparse coding. The experimental results show that our algorithm achieves better performance than four previous algorithms on three face databases.

1. Introduction

In the last decade, a number of sparse representation based classification (SRC) algorithms have been proposed for face recognition. The basic idea is that the test sample can be represented as a linear combination of all the training samples with sparsity constraint, and then can be classified by exploiting the reconstruction error. Huang [1] presents a theoretical framework for signal classification with sparse representation, which sparsely codes a signal over a set of redundant bases and classifies the signal based on its coding vector. Because to minimize the l_0-norm is an NP hard problem, we usually formulate the sparse coding problem as the minimization of the l_1-norm of the reconstruction coefficients or minimization of the l_2-norm of the reconstruction coefficients.

In the past several years, many face recognition algorithms based on the minimization of the l_1-norm of the reconstruction coefficients have been proposed. For example, Wright [2] uses sparse representation for robust face recognition. A test image is first sparsely coded over the template images, and then the classification is performed by checking which class yields the least coding error. Moreover, there are many variations of the SRC. Hui [3] exploits a k-nearest neighbor (KNN) method to classify a test sample using sparse representation, and which can reduce the computational complexity. Kang [4] presents a kernel sparse representation classification framework and utilizes the local binary pattern descriptor in the framework for robust face recognition. Mairal [5] proposes a joint dictionary learning and classifier construction framework. Deng [6] proposes an extended sparse representation-based classifier (ESRC) algorithm, and applies an auxiliary intra class variant dictionary to represent the possible variation between the training and testing images. Gabor features [7] and Markov random fields [8] are also used to further improve the accuracy of SRC. In addition, Ji [9] proposes an improved sparse representation classification algorithm based on non-negative constraint of sparse coefficient. Other nonnegative sparse representation algorithms can be found in Refs. [10–12]. Although SRC and its variations significantly improve the robustness of face recognition, they still need to solve the l_1-minimization problem on the whole dataset, which makes the computation expensive for large-scale datasets. Yang [13] presents a review of iterative shrinkage-threshold based sparse representation methods for robust face recognition. More sparse representation for computer vision and pattern recognition applications can be found in Ref. [14].

Recently, the collaboration representation used in SRC has shown very powerful classification capability, which is based on minimization of the l_2-norm of the reconstruction coefficients. Zhang [15] analyzes the working mechanism of sparse representation based classification (CRC), and indicates that it is the

* Corresponding author at: Guangdong Industry Training Center, Guangdong Polytechnic Normal University, Guangzhou 510665, China. Tel.: +86 20 38256428; fax: +86 20 38256428.
E-mail address: lizhengming2004@126.com (Z. Li).

http://dx.doi.org/10.1016/j.ijleo.2014.04.044
0030-4026/

collaborative representation but not the l_1-norm sparsity that makes SRC powerful for face classification. Therefore, many face recognition algorithms based on collaborative representation have been proposed. For example, Lee [16] presents an efficient sparse coding algorithm that is based on iteratively solving the l_1-regularized least squares problem and the l_2-constrained least squares problem, which can significant enhance the speed of sparse coding. Huang [17] proposes a face recognition algorithm based on collaborative image similarity assessment. Moreover, several variants of collaborative representation have been proposed in recent years by adding some additional regularization and/or constraints. Jadoon [18] proposes a collaborative neighbor representation algorithm for multi-class classification based on the l_2-minimization approach with the assumption of locally linear embedding. Naseem [19] presents a linear regression classification (LRC) algorithm by formulating the pattern recognition problem as a linear regression problem. Yang [20] proposes a regularized robust coding (RRC) model, which could robust regress a given signal with regularized regression coefficients. Moreover, the recently proposed two-phase test sample representation algorithm uses a novel representation based classification algorithm to perform face recognition (TPTSR) [21,22]. He [23] proposes a two-stage sparse representation (TSR) for robust face recognition on a large-scale database. Based on the divide and conquer strategy, TSR decomposes the procedure of robust face recognition into outlier detection stage and recognition stage.

However, in sparse coding, local features are dealt separately. The mutual dependence among local features is ignored, which results in the sparse codes may vary a lot even for similar features. To overcome this drawback, a number of face recognition algorithms based on local features have been proposed. Chen [24] proposes a nonnegative local coordinate factorization (NLCF) for feature extraction. NLCF adds a local coordinate constraint into the standard NMF objective function. In this way, each data point can be represented by a linear combination of only a few nearby basis vectors, which leads to sparse representation. Yu [25] assumes that each data point can be locally approximated by a linear combination of its nearby anchor points, and the linear weights become its local coordinate coding. Wang [26] utilizes the locality constraints to project each descriptor into its local-coordinate system, and the projected coordinates are integrated by max pooling to generate the final representation. Arpit [27] imposes a locality constraint to choose the training samples that are in the vicinity of the test sample. Chao [28] uses both group sparsity and data locality to formulate a unified optimization framework, which produces a locality and group sensitive sparse representation (LGSR) for improved recognition. Gao [29] incorporates Laplacian matrix into the objective function of sparse coding to preserve the consistence in sparse representation of similar local features. Those algorithms demonstrate that locally information can improve the performance of face recognition.

In this paper, we propose a face recognition algorithm based on the neighbor matrix, which can preserve locality and similarity information of sparse coding. The main idea is to calculate the neighbor matrix between the test sample and each training sample by using the revised reconstruction error of each class. Then we use the neighbor matrix and all the training samples to linearly represent the test sample. Furthermore, we propose a new optimization function than can preserve locality and similarity information of sparse coding. The experimental results show that our algorithm is very competitive on the ORL, YALEB and PIE face databases.

The remaining of this paper is organized as follows. Section 2 describes our proposed face recognition algorithm. Section 3 describes the neighbor coefficients assessment between images. The experiments and performance evaluation are reported in Section 4. Finally, the conclusion is presented in Section 5.

2. The proposed algorithm

In this section we describe the proposed algorithm, we assume that there are C classes of training samples $D = [x_1, \ldots, x_n]^T$, n is the number of training samples. If a test sample y belongs to one of the labeled classes in the training samples set D, we can use all the training samples to represent the test sample y. Then, the linear representation of a test sample y can be written as:

$$y = \alpha_1 x_1 + \ldots \alpha_n x_n \tag{1}$$

where $A = [\alpha_1, \ldots, \alpha_n]$ is a coefficient vector. In general, the sparse representation based algorithms reconstruct a test sample using a sparse linear combination of training samples. But they do not consider the underlying neighbor relation between the test sample and each training sample. Therefore, we assume that the following equation is approximately satisfied:

$$y = m_1 x_1 \alpha_1 + \ldots + m_n x_n \alpha_n \tag{2}$$

where $P = [m_1, \ldots, m_n]$ is the neighbor coefficient vector, $m_i (i = 1, \cdots n)$ represents that the neighbor relation between the ith training sample and the test sample. A bigger m_i means that the ith training sample is more close to the test sample. Eq. (2) also shows that if a training sample is far away from the test sample, it has smaller contribution to represent the test sample. Thus, the test sample can be represented better by using Eq. (2) than Eq. (1). Then, we can rewrite Eq. (2) as:

$$y = AMD \tag{3}$$

where $M = diag(m_1, \ldots, m_n)$ is the neighbor matrix whose only nonzero diagonal entries represents the neighbor coefficients between the test sample and each training samples. In general, in collaborative representation based classifier algorithm (CRC), we usually obtain the sparse solution of Eq. (1) by solving the following optimization function.

$$A = \underset{A}{\arg\min} ||y - AD||_2^2 + \lambda ||A||_2^2 \tag{4}$$

Therefore, the sparse solution of Eq. (2) can be

$$A = \underset{A}{\arg\min} ||y - AMD||_2^2 + \lambda ||A||_2^2 \tag{5}$$

where λ is a regularization parameter. In order to better preserve locality and similarity information of sparse coding, we incorporate the neighbor information into the objective function of Eq. (5). Therefore, we propose the following optimization function:

$$A = \underset{A}{\arg\min} \left\{ \frac{1}{2} \left(||y - AMD||_2^2 + \delta \sum_{i=1}^{n} \alpha_i^2 \beta_i^2 + \lambda ||A||_2^2 \right) \right\} \tag{6}$$

where $\beta_i = 1 - m_i (i = 1, \ldots, n)$, a smaller β_i means that the ith training sample is more close to the test sample, δ and λ are regularization parameters having small positive values. The rationale of Eq. (6) is as follows: we first calculate the distances between the test sample and all the training samples. If the test sample is the same class as the training sample, then the distance will be small. It is reasonable to assume that the class close to the test sample has small distance, and has its own contribution to minimize the object function of sparse code. So, we incorporate the distance information into the objective function of sparse code, which can preserve locality and similarity information of sparse code. Furthermore, it can eliminate the side-effect on the classification decision of the