Neurocomputing 148 (2015) 420–429

Contents lists available at ScienceDirect

Neurocomputing

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

Noise modeling and representation based classification methods for face recognition

Zheng Zhang [a], Long Wang [a], Qi Zhu [b], Zhonghua Liu [c], Yan Chen [a,d,*]

[a] Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, 518055 Guangdong, PR China
[b] Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016 Jiangsu, PR China
[c] Information Engineering College, Henan University of Science and Technology, Luoyang, 471000 Henan, PR China
[d] Shenzhen Sunwin Intelligent Corporation, Shenzhen, 518055 Guangdong, PR China

ARTICLE INFO

Article history:
Received 14 April 2014
Received in revised form
12 June 2014
Accepted 2 July 2014
Communicated by Su-Jing Wang
Available online 11 August 2014

Keywords:
Pattern recognition
Face recognition
Noise modeling
Representation based method

ABSTRACT

In this paper, we propose a novel noise modeling framework to improve a representation based classification (NMFIRC) method for robust face recognition. The representation based classification method has evoked large repercussions in the field of face recognition. Generally, the representation based classification method (RBCM) always first represents the test sample as a linear combination of the training samples, and then classifies the test sample by judging which class leads to a minimum reconstruction residual. However, RBCMs still cannot ideally resolve the face recognition problem owing to the varying facial expressions, poses and different illumination conditions. Furthermore, these variations can immensely influence the representation accuracy when using RBCMs to perform classification. Thus, it is a crucial problem to explore an effective way to better represent the test sample in RBCMs. In order to obtain a highly precise representation metric, the proposed framework first iteratively diminishes the representation noise and achieves better representation solution of the linear combination until it converges, and then exploits the determined 'optimal' representation solution and a fusion method to perform classification. Extensive experiments demonstrated that the proposed framework can simultaneously notably improve the representation capability by decreasing the representation noise and improve the classification accuracy of RCBM.

1. Introduction

Face recognition has achieved a great development in the past few decades and has also caught the great attention from large numbers of research groups [1–3]. However, face recognition suffers from some difficulties, such as varying illumination conditions, different poses, disguise and facial expressions [4–6]. A plenty of face recognition algorithms have been designed to alleviate these difficulties [7–9]. A novel method called the representation based classification method (RBCM) is proposed for face classification. RBCM exploits the training samples to represent the test sample, and then utilizes the representation solution to classify the test sample. The test sample can be classified as a member of the class based on the minimum reconstruction residual [10,11]. The representation based method has been widely applied to many fields such as face recognition [12–15], palmprint recognition [16], signal classification [17], image restoration [18], visual tracking [19].

The representation based classification method (RBCM) has become a powerful and effective method in pattern recognition. With the advancements of the l_0-norm, l_1-norm and l_2-norm minimization techniques, more and more representation based methods for face recognition have been devised. Especially, the sparse representation based classification methods are widely applied to improve the robustness of face recognition [20,21]. The general idea of sparse RBCM is to represent the test sample by exploiting a over-complete dictionary with a sparse constraint, and then to use the sparse representation (or coding) solution to perform final face classification. It is considered that the sparse constraint plays an important role in accurately classifying the test sample. Hereafter "sparsity" specifically refers to an expression of the input signal as a linear combination of training samples and many of the coefficients are zero or very close to zero.

The sparsest solution of RBCM can be obtained by exploiting the sparse representation method with l_0-norm minimization. It has been demonstrated that solving the linear optimization problem with l_0-norm minimization is NP hard problem [1,11,21,22]. However, literatures [21–24] have certified that the solution obtained by l_0-norm minimization can be equivalent to the solution using l_1-norm minimization with enough sparsity. It should be pointed

* Corresponding author at: Bio-Computing Center, Harbin Institute of Technology, Shenzhen, PR China.
E-mail address: jadechenyan@gmail.com (Y. Chen).

http://dx.doi.org/10.1016/j.neucom.2014.07.058
0925-2312/

out that a RBCM named robust face classification via the sparse representation classification (SRC) was proposed by Wright et al. [1]. SRC first represents the test sample using all the training samples and obtains the sparse solution based on l_1-norm minimization, and then classifies the test sample as a member of the class which leads to the minimum reconstruction residual. SRC performs very well in face recognition and its robustness and discriminative capability have also been validated [11]. Although SRC has greatly influenced face recognition and has been widely studied, it still has some unsolved problems. For example, it cannot perfectly address the negative effects from the variations of poses, illumination conditions and facial expressions. Furthermore, the algorithm of SRC usually is to iteratively solve the minimization problem and thus it has very high time consuming. Many methods have been proposed to obtain better performance by improving SRC. For example, Deng et al. [25] proposed an extended SRC scheme which applied an intraclass variant dictionary to represent possible variation between the training samples and the test sample. He et al. [26] proposed a two-phase test sample sparse representation method to enhance the robustness and efficiency of SRC. They first detect the outliers and noise in the face images to learn a metric, and then use the metric to obtain a small subset and finally perform the non-negative SRC. On the contrary, these proposed methods also reflect the imperfection of SRC and these variations from external environments still need to pay much attention for SRC. Thus, overcoming these variations is still a problem for researchers.

Most of literatures on the SRC emphasize the sparsity achieved by l_1-norm minimization constraint for face recognition. However, there also exist disagreements. For example, literatures [14,27,28] doubted the role of sparsity in RBCM with l_1-norm minimization constraint and demonstrated that the influence of the sparsity on recognition was not so strong via lots of experiments. Shi et al. [28] pointed out that the sparse approximation could not satisfy the needs of robustness and required performance, and then proposed to use simple l_2-norm minimization to obtain more effective classification. Zhang et al. [14] confirmed that the sparsity based on l_1-norm minimization based could not really make critical differences in classification and then proposed an efficient RBCM named collaborative representation classification (CRC) method for face recognition. The mechanism of CRC employs l_2-norm minimization rather than l_1-norm minimization to obtain representation solution and the computational efficiency of CRC is extremely higher than SRC with similar classification accuracy [14]. Subsequently, many approaches based on l_2-norm minimization have been proposed. For example, Xu et al. [29] exploited the scheme of RBCM to improve the nearest neighbor classifier (INNC). Liu et al. [16] proposed a representation-based palmprint recognition method called the fusion classification method (FCM) based on the representation residual by using a weighted sum of all the training samples to replace the test sample, and then exploited the fused representation result to conduct classification. Xu et al. [12] proposed a semi-supervised two-step sparse representation classification based on the l_2-norm minimization and achieved extraordinarily high classification accuracy with much more efficiency. Thus, it is confirmed that the RBCM implemented based l_2-norm minimization can also obtain impressive classification accuracy with much high efficiency.

However, all these literatures related to RBCM above [8,10–16,29] still ignore a very important fact that the test sample still cannot be precisely represented exploiting the training samples with variations and the representation noise cannot be directly discarded. These methods [1,6,8,10–16,29] all assume that each test sample could be represented by a linear combination of the training samples. Nevertheless, this assumption is very hard to be satisfied and representation noise is unavoidable during the representation procedures. Furthermore, from the viewpoint of linear algebra, if the dimensionality of the sample is larger than the number of the training samples, a linear combination of the training samples cannot precisely represent the test sample. Literatures [25,26] both have a tendency to diminish or remove the negative effects from the external environments and exploit the uncontaminated face images for robust face recognition. To this end, it is beneficial to eliminate or minimize the representation noise in the process of RBCM. The proposed framework is a novel way to lower the side effects from variations and also demonstrates that the significance of more precisely representing the test sample is never more demand.

In this paper, we propose a novel framework of improving RBCM to solve the problem of representation noise. In order to more precisely obtain the linear combination of training samples, $x_1, x_2, \ldots, x_n$, to represent the testing sample y, we first assume that $y = a_1x_1 + a_2x_2 + \cdots + a_nx_n + s_i$ where s_i is viewed as representation noise. In order to obtain a minimum representation noise s_i, we first iteratively minimize the representation noise and achieve better representation solution of the linear combination until the condition of convergence is reached, and then exploit the determined 'optimal' representation solution to perform effective classification.

The remainder of this paper is structured as follows: we first briefly review the existing related RBCMs in Section 2. In Section 3, we present the proposed framework of NMFIRC. In Section 4, we discuss the rationale of the proposed approach. In Section 5, the proposed method is verified by performing extensive experiments on several well-known face databases. Finally, we offer the conclusion in Section 6.

2. Description of typical RBCMs

Suppose that there are n training samples, $X = [x_1, x_2, \ldots, x_n]$ from c classes and the number of the training samples in each class is m. The testing sample is y. All the samples should be firstly normalized to unit column vectors with the length of 1.

2.1. RBCM with l_2-norm minimization

When it comes to the RBCM with l_2-norm minimization (RBCM2NM), collaborative representation classification (CRC) [14] is the most representative method. It first determines a linear combination of all the training samples to represent the testing sample y. Thus, CRC assumes that the following equation is approximately satisfied:

$$y = a_1x_1 + a_2x_2 + \cdots + a_nx_n + s_i \quad (1)$$

where y is the normalized unit testing sample vector and the a_i $(i = 1, 2, \ldots, n)$ is the coefficient of x_i. For convenient description, we can rewrite Eq. (1) into the following equation:

$$y = X\alpha \quad (2)$$

where $\alpha = [a_1 a_2 \ldots a_n]^T$ and $x_1, x_2, \ldots, x_n$ and y are all column vectors. The representation solution, α, can be represented by solving l_2-norm minimization problem:

$$\hat{\alpha} = \operatorname{argmin} \|\alpha\|_2 \quad \text{s.t. } y = X\alpha \quad (3)$$

We exploit the Lagrangian algorithm to obtain the solution of Eq. (3) and the solution of α is to solve the l_2-norm minimization problem below:

$$\operatorname{argmin}(\|y - X\alpha\|_2 + \mu\|\alpha\|_2) \quad (4)$$

where μ is a small positive constant. We can solve α by using $\alpha = (X^TX + \mu I)^{-1}X^Ty$, where I is the identity matrix.