Optik 126 (2015) 2799–2803

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A nearest neighbor classifier based on virtual test samples for face recognition

Ningbo Zhu, Ting Xu*, Lei Wei, Ting Tang

College of Information Science and Engineering, Hunan University, Changsha 410082, China

ARTICLE INFO

Article history:
Received 17 May 2014
Accepted 2 July 2015

Keywords:
Face recognition
Nearest neighbor classifier
Virtual sample
Linear combination
Sparse representation

ABSTRACT

In this paper we propose a nearest neighbor classifier which aims at improving the classification accuracy of face recognition. The idea of the proposed method is as follows. Firstly, the symmetry of the original test samples is used to generate new test samples. Then, training samples are used to represent original test samples and the virtual test samples respectively. It takes the advantage of the weighted sum to construct a nearest neighbor classifier to improve the accuracy of face recognition. Meanwhile, the proposed method codes a test sample as a linear combination of all of the training samples, and the deviation between the training samples and the test samples is exploited to classify the test sample. The proposed method can perform better in the case with a small number of training samples than the improvement nearest neighbor classifier. In this paper, the proposed method is compared with a simple and fast representation-based face recognition method, an improvement to the nearest neighbor classifier, a novel sparse representation method based on virtual samples for face recognition (SRMVS) and a two-phase test sample sparse representation method (TPTSR). The experimental results show that our method has better classification results than the others.

1. Introduction

Face recognition (FR) has been a research hotspot in computer vision and pattern recognition in recent years. Although FR has been greatly improved in recently years, it is still a challenging task due to its varying lighting, facial expression, poses and environments. And to further improve the correct rate of face recognition, various methods have been proposed. For example, the principal component analysis (PCA) proposed in [1–3], which can deal with dimensionality reduction and feature extraction, is a classical and typical method. Another well-known approach is Linear Discriminant Analysis (LDA) and it is an effective method to feature extraction and dimensionality [4,5]. PCA and LDA are widely used in face recognition and they are two typical examples of linear transform methods. In addition, the nearest neighbor classifier (NNC) is an important classifier for face recognition [6]. NNC is also one of the oldest and simplest classifiers performed by a large of researchers to achieve a better rate of classification in recent years. The key point of the NNC is to determine the "closeness". That is, the first step of NNC is to determine the samples which are close to the test sample.

Recently, a novel and remarkable face recognition method, which is named sparse representation (SR), has been proposed in [7–14]. The key idea of this method is that, one can use all of the training samples to represent the test samples as a sparse linear combination. The "sparse" is regarded as follows: when a method uses all of the training samples to represent a test sample as a linear combination, the coefficients of some training samples are close or equal to zero. By computing the potential contributions of different training samples to express the test sample, the sparse representation method finally classifies the test sample to the class that makes the maximum contribution. These methods usually exhibit a good recognition rate from [15–24], which have a better rate of classification in face recognition.

Though sparse representation can perform very well in face recognition, it is still faced with a limited number of available training samples. A face recognition system usually has a small number of samples so that it cannot obtain a better recognition rate. In order to solve the problem, a series of methods have been proposed to exploit virtual training samples to represent test samples. For example, Tang et al. [22] proposed a novel sparse representation method which is based on virtual samples (SRMVS). The symmetry of the face was used by Xu and Zhu [23] to generate new training samples and to perform more accurate recognition of faces. The methods in [22,23] use the virtual samples of training samples and get a good recognition rate with a small number of samples.

* Corresponding author. Tel.: +86 15084710554.
E-mail address: yoyolan111@163.com (T. Xu).

http://dx.doi.org/10.1016/j.ijleo.2015.07.014
0030-4026/

Other than using virtual training samples, in this paper, we propose a sparse representation method based on virtual test samples and devise a collaborative representation to perform nearest neighbor classification. The sparse representation and the improvement to nearest neighbor classifier are combined in the proposed method. Firstly, the original test samples have been used to generate virtual samples, moreover, the virtual samples have been divided into the symmetry of left side of original face and the symmetry of right side in our scheme. Then, the original test samples and the two of virtual test samples are represented as a linear combination of all the training samples respectively. The nearest neighbors of test samples are determined by the representation results, and the nearest sample is exploited to classify the test sample. The proposed method takes advantages of the score level fusion, which is usually better than the decision level and feature level fusion.

The rest of this paper is organized as follows: In Section 2 we describe the main idea of our scheme. The analysis of our method is provided in Section 3. The experimental results are showed in Section 4 and we conclude our work in Section 5.

2. Description of the proposed method

In this section we mainly describe the proposed method. We assume that there are L classes in the face database and each class includes S test samples, $y_1 \ldots y_s$ and T training samples, $x_1 \ldots x_t$, namely, each class includes $S+T$ samples in total. We take $j(j=1, 2\ldots L)$ as the class label of the training sample which is from the jth class. We will present the main steps of the proposed method in the first subsection. And then describe the steps in details in the next subsection.

2.1. The main steps of the proposed method

In this subsection, the steps of the algorithm are described as follows:

Step1. Use the original test samples to generate two 'symmetrical face' test samples. We denote them by y_n, y_n^l, y_n^r $(n=1, 2\ldots s)$, respectively. y_n^l is the virtual test samples which copy the left side of the original test samples to the right side, namely, the two sides of the 'symmetrical face' are same as the left side of the original test samples. However, y_n^r is the virtual test samples which copy the right side of the original test samples to the left side, that is, the two sides of the 'symmetrical face' are same as the right side of the original test samples.

Step2. Use the original training samples to represent the original test samples and the two of virtual samples, respectively.

Step3. Exploit the nearest neighbors of the original test samples and the virtual test samples to perform the nearest neighbor classifier. The algorithm presents the solution in detail.

Step4. When the deviation of rth training samples from each test sample is reached. If the deviation of rth training samples from the original test samples is D_r^1, D_r^l, D_r^r is referred to the deviation of rth class from the virtual test samples. Then, conduct weighted score level fusion. For example, $D_r = w_1 D_r^1 + w_2 D_r^l + w_3 D_r^r$ is used to denote the contribution to the rth training samples in total. w_1, w_2, w_3 are the weights and the sum of them equals to 1, the w_2, w_3 should be equal and smaller than w_1. In the last, the test samples will be classified into the class which is same as the training sample that has the minimum D_r.

2.2. The description of the proposed method

Lei Zhang et al. showed that the sparse representation coded by the l_2 norm can lead a better recognition rate with few samples [23]. In our opinion, though the l_2 norm is much weaker than l_1 norm, it has a similar classification results but with significantly lower complexity. In order to collaboratively represent the test samples using training samples with low computational burden and effective classification, we also propose to use the l_2 norm.

The first phase of the proposed method works as follows: it first uses every original test sample to generate two 'symmetrical face' test samples. We use y_n, y_n^l, y_n^r $(n=1, 2\ldots s)$ to denote the original test samples, the symmetrical of left side of the original test samples and the symmetrical of right side of the original test samples, respectively.

The algorithm of the proposed method is similar with sparse representation. The test sample can be represented as a linear combination of all of the training samples, with the following formula:

$$y = \sum_{i=1}^{l\times n} a_i x_i \tag{1}$$

where y is the test sample, x_i are the training samples and a_i are the corresponding coefficients.

For simplicity of presentation, only the algorithm on the original test samples is described in this section and the classification of virtual test samples is similar to it. We suppose each class has t training samples, respectively, there are $l\times t$ training samples denoted by $l\times t$ column vectors $a_1 \ldots a_{t\times l}$. And $a_i\,(i=1, 2\ldots l\times t)$ can be describe as a matrix A, $[a_1\ldots a_{t\times n}]^T$, likewise, x_i can be described as a matrix X, $[x_1\ldots x_{t\times n}]^T$. So, the Eq. (1) can be rewritten into the following equation:

$$y_n = A \times X \tag{2}$$

If X is a nonsingular square matrix, A would be solved by Eq. (3):

$$A = X^{-1} y_n \tag{3}$$

Otherwise, it also can be solved by Eq. (4):

$$A = \left(X^T X + \rho I\right)^{-1} X^T y_n \tag{4}$$

where ρ is a small positive constant and I denotes the identity matrix.

Since we get the A, the test samples can be represented by Eq. (1), and the different training samples have corresponding contributions. That is, the contribution of the ith training sample is $a_i x_i$ where a_i is the ith component of A. Then, the distance between the test sample y and the ith class would be measured as follows:

$$e_i = ||y_n - a_i x_i||_2^2 \tag{5}$$

e_i is also referred to as the deviation of the ith training sample from nth test sample y_n. The smaller deviation has greater contribution in representing the test sample and the a_i denotes corresponding coefficients of the training samples.

e_i is exploited to determine the smallest deviation and classifies it into the smallest deviation. As a result, it identifies the class label of the training samples which has the minimum deviation and classifies the test sample into the same class.

3. Analysis of the proposed method

This section will describe the rationales of the proposed method and analyze it in detail.

The rationale of the proposed method is to use a linear combination of a subset of training samples to represent the test samples and virtual test samples and exploit the weighted fusion of the deviation of each class from test samples and virtual test samples to perform classification. In this paper, the key point is to generate virtual samples of test samples and it will make better classification results than the method which generates and exploits virtual training samples to perform classification. Fig. 1 shows some original test samples from the ORL face database and the virtual test