Optik 125 (2014) 3908–3912

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A novel virtual samples-based sparse representation method for face recognition

Yuyao Wang[a,b,*], Min Wang[a,b], Yan Chen[b,c], Qi Zhu[a,b]

[a] Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
[b] Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, China
[c] Shenzhen SUNWIN Intelligent Co., Ltd., Shenzhen, China

ARTICLE INFO

Article history:
Received 1 August 2013
Accepted 28 January 2014

Keywords:
Virtual training samples
Sparse representation
Face recognition

ABSTRACT

A limited training set usually limits the performance of face recognition in practice. Even sparse representation-based methods which outperform in face recognition cannot avoid such situation. In order to effectively improve recognition accuracy of sparse representation-based methods on a limited training set, a novel virtual samples-based sparse representation (VSSR) method for face recognition is proposed in this paper. In the proposed method, virtual training samples are constructed to enrich the size and diversity of a training set and a sparse representation-based method is used to classify test samples. Extensive experiments on different face databases confirm that VSSR is robust to illumination variations and works better than many representative representation-based face recognition methods.

1. Introduction

The sparse representation method has been popular for face recognition in recent years, and is commented as a breakthrough in face recognition [1]. Many sparse representation-based face recognition methods have been proposed [2–10]. Xu et al. [11] proposed a two-phase sparse representation method (TPTSR) which picked out some “neighbor” training samples in the first phase and classified a test sample with the representation of “neighbor” samples. Yang and Zhang [12] applied Gabor-feature to sparse representation-based classification (SRC), which improved SRC accuracy greatly on occluded face images. Wang et al. [13] proposed a method that made use of sparse representation to discard some training samples, and classified test samples based on Scale Invariant Feature Transaction feature (SIFT feature). Xu and Zhu [14] firstly picked out a “nearest” training sample from each training class for a test sample base on Euclidean metric. Then a test sample was expressed as a sparse representation of all selected training samples to accomplish the classification. Gao and Yang [15] applied sparse representation to face image super-resolution method. In their method, a sparse representation was obtained for each low-resolution image, and representation coefficients were used to generate corresponding high-resolution output. Gao et al. [16] proposed a Kernel Sparse Representation method (KSR) that combined kernel trick with sparse representation. KSR was applied to image classification and face recognition. Naseem et al. [17] addressed the video-based face recognition in SRC for the first time. They fused SRC with SIFT-based methods using weighted sum rules. Elhamifar [18] applied a sparse representation technique to classification in multi-sub space setting. They aimed to find a representation of a test sample that used minimum number of training samples.

Although sparse representation-based methods have an impressive performance on face recognition, they also have to face the limited training set problem that limits the performance of many face recognition methods in practice. Few methods have been proposed to solve the limited training set problem for sparse representation. Chang et al. [19] proposed to generate new images by shifting original images or reconstructing original images via PCA. New images were regarded as training samples and SRC was used for face recognition. Wang and Zhu [20] generated 13 virtual samples for each original sample using geometric transformation and SVD decomposition. Tang et al. [21] proposed a method that constructed virtual training samples via adding random noise to original images. Random noise was added to the whole image of an original sample, and a new training set consisted of original samples and virtual samples.

The method proposed in [21] was simple and effective on several face databases, but its performance was far from satisfactory.

* Corresponding author at: Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology Shenzhen, China.
E-mail addresses: shinewyy@gmail.com (Y. Wang), leafcano@gmail.com (M. Wang), jadechenyan@gmail.com (Y. Chen), ksqiqi@sina.com (Q. Zhu).

http://dx.doi.org/10.1016/j.ijleo.2014.01.161
0030-4026/

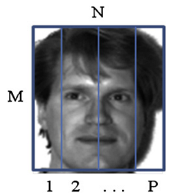

Fig. 1. A divided sample. This sample is divided transversely into P parts, and the size of each part is $M \times (N/P)$.

Constructing virtual samples via adding random noise to the whole image of a sample was just to simulate simple illumination change, leading to the whole image darken or lighten. Inspired by the method proposed in [21], we propose a virtual samples-based sparse representation method (VSSR) for face recognition in this paper. The proposed method contains two steps. The first step is to enlarge an original training set, which divides an original image into several parts and generates a corresponding virtual training sample via adding different random noise to pixels of different parts. This step simulates various illumination changes. Therefore, original samples and virtual samples compose a new training set. The second step is to apply SRC for face recognition. VSSR not only generates virtual samples to simulate various illumination changes of original samples, but also enlarges the original training set by combing virtual samples and original samples. In a word, VSSR enriches a training set in size and diversity. Extensive experiments on face databases demonstrate that the proposed method is robust to illumination and outperforms many state-of-the-art sparse representation-based face recognition methods.

The rest of this paper is organized as follows: In Section 2, procedure of the proposed method is detailed and an analysis on VSSR is described. Experimental results are demonstrated in Section 3. A brief conclusion is given in Section 4.

2. Description and analysis of the proposed method

2.1. Details of the proposed method

The VSSR method proposed in this paper mainly contains two steps. The first step is to enlarge the original training set by constructing virtual samples. The second step is to classify test samples. The procedure of VSSR is detailed in this section.

2.1.1. STEP 1: Construct virtual samples based on original training samples

In the first step, virtual samples are constructed based on original samples. Assume X be a training sample and its size is $M \times N$. X is regarded as a prototype sample to construct virtual samples. Define X' as a virtual sample with the same size as X. X' is constructed in the following way. The prototype sample X is divided into P parts, and the size of each part is, $M \times (N/P)$ as shown in Fig. 1.

Each pixel of X' is defined as follows:

$$X'(i_p, j_p) = (1 + f)X(i_p, j_p) \tag{1}$$

In Eq. (1), (i_p, j_p) denotes a position in pth part of a sample, $X'(i_p, j_p)$ denotes the pixel value in position (i_p, j_p) of X', and $X(i_p, j_p)$ denotes the pixel value in the same position of X as X'. In VSSR method, f is a function and defined as Eq. (2):

$$f = a \cdot (rand() - b) \tag{2}$$

where a and b are positive constants, and $rand()$ is a function that generates a set of random real numbers between 0 and 1.

A virtual sample is constructed for each original training sample by using Eqs. (1) and (2). Virtual samples and original training samples are combined to constitute a new training set $TrainSet$. In $TrainSet$, K denotes the number of training classes, and Q denotes the number of training samples in kth class.

$$TrainSet(k, q) = \begin{cases} Z_k^q, & k \in [1, 2 \ldots K], \quad q \in [1, 3 \ldots Q-1] \\ (Z_k^q)', & k \in [1, 2 \ldots K], \quad q \in [2, 3 \ldots Q] \end{cases} \tag{3}$$

2.1.2. STEP 2: Classify test samples using sparse representation-based method.

In the second step, test samples are classified by using sparse representation classification method (SRC), based on the new training set $TrainSet$ that is obtained in the first step. The main idea of SRC is to represent a test sample as a linear combination of all training samples.

Assume Y be a test sample. Y is represented as a linear combination of all samples in $TrainSet$:

$$Y = \sum_{k=1}^{K} \sum_{q=1}^{Q} a_{(k-1)\times Q+q} \hat{X}_k^q \tag{4}$$

where $\hat{X}_k^q$ denotes qth sample of kth class in $TrainSet$, and $a_{(k-1)\times Q+q}$ denotes the corresponding coefficient of the training sample $\hat{X}_k^q$ for linearly representing a test sample. Eq. (3) can be rewritten to the following equation:

$$Y = \hat{X}A \tag{5}$$

where $\hat{X} = [x_1, \ldots, x_{Q\times K}]$, $A = [a_1, \ldots, a_{Q\times K}]^T$ and Y is a column vector. If $\hat{X}$ is a nonsingular matrix, A can be solved by using:

$$A = \hat{X}^{-1} Y \tag{6}$$

Otherwise, A can be solved by using:

$$A = (\hat{X}^T \hat{X} + \mu I)^{-1} \hat{X}^T Y \tag{7}$$

where μ is a positive constant and usually set to 0.01 and I is the identity matrix.

Once the coefficient matrix A is solved, all training samples can be used to represent a test sample. It is learned from Eq. (4) that each training sample has its own contribution for linearly representing a test sample. Therefore, Eq. (8) is applied to calculate the contribution of each training class for representing a test sample:

$$con_k = \sum_{q=1}^{Q} a_{(k-1)\times Q+q} \hat{X}_k^q \tag{8}$$

Eqs. (4) and (8) show that the contribution con_k is the representation result of the kth class in $TrainSet$. Eq. (9) is used to evaluate the similarity of a test sample Y with the representation result of the kth training class.

$$D_k = \left\| Y - con_k \right\| \tag{9}$$

In this step, the sparse representation-based method classifies the test sample Y into a training class that is the most similar to Y. That is to say, the test sample Y is classified into the kth training class if $D_k = \min\{D_1, D_2, \ldots, D_K\}$.

As described above, a new training set is obtained via using the first step of VSSR, and SRC is used for face recognition on the new training set.

2.2. Analysis of the proposed method

In the first step of VSSR, a training set is enlarged in size via constructing virtual samples. Different from the method proposed in