Neurocomputing 147 (2015) 307–314

Contents lists available at ScienceDirect

Neurocomputing

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

Noise-free representation based classification and face recognition experiments

Yong Xu [a,b], Xiaozhao Fang [a], Jane You [c], Yan Chen [a,d,*], Hong Liu [e]

[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] Department of Computing, The Hong Kong Polytechnic University, Hong Kong
[d] Shenzhen Sunwin Intelligent Corporation, Shenzhen, China
[e] Engineering Lab on Intelligent Perception for Internet of Things, Shenzhen Graduate School, Peking University, Shenzhen, China

ARTICLE INFO

Article history:
Received 1 July 2013
Received in revised form
16 June 2014
Accepted 23 June 2014
Communicated by D. Tao
Available online 30 June 2014

Keywords:
Noise elimination
Representation based classification

ABSTRACT

The representation based classification has achieved promising performance in high-dimensional pattern classification problems. As we know, in real-world applications the samples are usually corrupted by noise. However, representation based classification can take only noise in the test sample into account and is not able to deal with noise in the training sample, which causes side-effect on the classification result. In order to make the representation based classification more suitable for real-world applications such as face recognition, we propose a new representation based classification method in this paper. This method can effectively and simultaneously reduce noise in the test and training samples. Moreover, the proposed method can reduce noise in both the original and virtual training samples and then exploits them to determine the label of the test sample. The virtual training sample is generated from the original face image and shows possible variation of the face in scale, facial pose and expression. The experimental results show that the proposed method performs very well in face recognition.

1. Introduction

As we know, noise extensively exists in the data [1]. In order to better model, predict and classify the data, we should well deal with noise. For pattern classification problems, it is significant to devise a way to resist noise and to improve the robustness of the classifier [2–4].

As shown in literature [5], noise has great influence on the face recognition accuracy. Moreover, the image of a face usually varies with the illumination, pose and facial expression [6–11]. This is indeed a great challenging problem in face recognition. We can treat the difference between the images of the same face as generalized noise. For pattern classification problems, if we can identify and reduce noise, a better result can be obtained. A number of methods have been proposed for noise-free face recognition. For example, it is conceived that not-necessarily orthogonal basis which may reconstruct the data is better than principal component analysis (PCA) in the presence of noise and independent component analysis (ICA) is proposed for face recognition in the presence of noise [12]. A unified framework of subspace is proposed for robust face recognition [13]. The enhanced fisher linear discriminant (EFLD) model is also proposed for overcoming noise in face images [14]. In recent years, the low-rank decomposition is also applied to eliminate noise [15–17].

The recently proposed representation based classification has performed very well in high-dimensional pattern classification problems. Especially, the representation based classification proposed by Wright et al. [18,19], i.e. sparse representation classification (SRC) is viewed as a breakthrough of face recognition. Besides SRC, there have also been a number of other representation based classification methods. For example, representation based classification with the l_2 norm minimization constraint on the solution vector is not only able to obtain a high accuracy but also computationally very efficient. For example, collaborative representation [20], the two-phase test sample representation (TPTSR) method [21], the feature space representation method [22–24] have achieved satisfactory results in face recognition. The recently proposed linear regression classification (LRC) is also a representation based classification with the l_2 norm minimization constraint [25,26]. LRC is closely related with the previously proposed nearest intra-class space (NICS) method [27]. The pattern recognition community has paid much attention to the theoretical foundation of representation based classification and to design new representation based classification algorithms [28–30].

Rationales of representation based classification have been demonstrated from different aspects. Wright et al. considers that the

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

http://dx.doi.org/10.1016/j.neucom.2014.06.058
0925-2312/

"sparsity" of the representation is very helpful for achieving the high classification accuracy [18,19]. However, Zhang et al. claimed that for representation based classification it is not the "sparsity" but the way to represent and classify the test sample, i.e. collaborative representation that contributes the most to the face recognition performance [20]. Moreover, Yang et al. considered that for SRC "locality" is more significant than "sparsity" because in representation based classification "locality" always leads to "sparsity" but not vice visa [28]. Our studies on representation based classification with the l_2 norm minimization constraint show that the "sparsity" can be achieved by using a simple scheme which is very helpful for identifying the class that the test sample is truly from [21,31,32]. The idea of representation based classification has been applied to improve various methods such as tensor discriminant analysis and eigen-subspace methods [30,33–35]. A number of studies on manifold learning also made notable contributions in designing methods that preserve localities of samples. For example, Laplacian faces [36], semisupervised multiview distance metric learning [37], elastic manifold embedding [38] and adaptive hypergraph learning [39] all address the problem of preserving locality structures of samples from different viewpoints. These methods can be applied to various issues such as object correspondence construction in animation [40] and Cartoon character retrieval [41]. Sparse representation was also integrated with other methods such as the wavelet decomposition for face recognition [42].

It seems that it is crucial to properly model noise in data [43]. Though noise exists in both the training and test samples, the algorithm of conventional representation based classification is established on the basis of the conventional least squares algorithm and it cannot take noise in the training samples into account. This will cause side-effect to the classification accuracy. For face recognition, besides noise from the acquisition stage, the variation of the facial pose and expression [44,45] of the same face can also be viewed as generalized noise.

In this paper, we use the following scheme to improve the representation based classification: we first perform matrix decomposition for the matrix consisting of all the training samples and obtain the approximations of all the training samples, referred to as approximation training samples (ATRS). Then we exploit the ATRS to obtain an approximation of the test sample, referred to as approximation test sample (ATES). Finally, conventional representation based classification (CRBC) is applied to ATES and ATRS and the classification result of the test sample is obtained. Moreover, motivated by the fact the face usually has an axis-symmetrical structure, the proposed method also exploits the original face images to generate virtual symmetrical face images, which are helpful for showing possible variation of the face in scale and pose. Simply speaking, an original face image will generate two virtual symmetrical face images. The left half of the first virtual symmetrical face image is the same as the left half of the original face image and the right half of the first virtual symmetrical face image is just the mirror image of its left half. The right half of the second virtual symmetrical face image is the same as the right half of the original face image and the left half of the second virtual symmetrical face image is just the mirror image of its right half. The proposed method uses both the original and virtual training samples to represent and classify the test sample and outperforms the state-of-art face recognition methods. The main contributions of this work are as follows: (1) it proposes a simple and reasonable way to simultaneously reduce noise in the training and test samples. (2) The designed algorithm can lead to very accurate recognition of faces by properly integrating the original and virtual training samples.

2. The proposed method

In this section we describe the proposed method in detail. Suppose that there are C classes and each class has n training samples. Let $x_1, x_2, \ldots, x_N (n = nC)$ be all the training samples from the first, second,…, and C-th classes, respectively. In other words, $x_1, x_2, \ldots, x_n$ are the n training samples from the first class. $x_{n(i-1)+1}, x_{n(i-1)+2}, \ldots, x_{ni}$ are the n training samples from the i-th class. Let y be the test sample. $x_1, x_2, \ldots, x_N$ and y are all M dimensional column vector. The following context describes the main steps of the algorithm of approximation representation (AAR).

Step 1. Let $x_1, x_2, \ldots, x_N$ be all available training samples. Let $X = [x_1 x_2 \ldots x_N]$ and perform singular value decomposition (SVD) to obtain $X = U\Lambda V^T$, $U = [u_1 \ldots u_d]$, $V = [v_1 \ldots v_d]$, Λ is a diagonal matrix and $diag(\Lambda) = [\lambda_1 \ldots \lambda_d]$, $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_d$. Take $\hat{X} = \sum_{i=1,\ldots,K} \lambda_i u_i v_i^T (K \leq d)$ as the reconstruction of X. The j-th column of $\hat{X}$ i.e. $\hat{x}_j$ is the j-th approximation training sample (ATRS) and is also referred to as reconstruction of the j-th training sample x_j.

Step 2. Establish equation $y = \hat{X}A$ and solve it using $\hat{A} = (\hat{X}^T\hat{X} + \mu I)^{-1} y$. μ is a small positive constant and I is the identity matrix. Treat $\hat{y} = \hat{X}\hat{A}$ as the approximation test sample (ATES).

Step 3. Use the ATES and ATRS to establish equation $\hat{y} = \hat{X}B$ and solve it using $\tilde{B} = (\hat{X}^T\hat{X} + \mu I)^{-1}\hat{y}$.

Step 4. Let $\tilde{B} = [\tilde{b}_1 \ldots \tilde{b}_N]$ and $\hat{X} = [\hat{x}_1 \ldots \hat{x}_N]$. Define $S_j = ||\hat{y} - \sum_{k=1,\ldots,n} \tilde{b}_{n(j-1)+k}\hat{x}_{n(j-1)+k}||_2$ as the distance between the test sample and the j-th class.

The proposed method is implemented for test sample y as follows:

(1) Implement the algorithm of approximation representation (AAR) for the original training samples. The score of the test sample with respect to the j-th class is denoted by S_j^1.
(2) Generate virtual training samples and perform AAR for them. The score of the test sample with respect to the j-th class is denoted by S_j^2.

Virtual training samples are generated using the procedure below, which was first proposed in [31]. Let $x_{i0} \in \Re^{p \times q}$ be the i-th training sample in the form of image matrix. Let y_i^1 and y_i^2 respectively stand for the first and second virtual training samples generated from x_{i0}. The left half columns of y_i^1 is set to the same as that of x_{i0} and the right half columns of y_i^1 is the mirror image of the left half columns of y_i^1. However, the right half columns of y_i^2 is set to the same as that of x_{i0} and the left half columns of y_i^2 is the mirror image of the right half columns of y_i^2. y_i^1 and y_i^2 are also referred to as the first and second 'symmetrical face' training samples. The mirror image S of an arbitrary image R is defined as $S(i,j) = R(i, H-j+1)$, $i = 1, \ldots, G$, $j = 1, \ldots, H$. G and H stand for the numbers of the rows and columns of R, respectively. $S(i,j)$ denotes the pixel located at the ith row and jth column of S. For more details, please refer to [31].

AAR is performed for the set consisting of all the first and second virtual training samples. The score of the test sample with respect to the j-th class is denoted by S_j^2.

(3) The scores generated from the original and virtual training samples are integrated using $d_j = uS_j^1 + vS_j^2$, where u, v are the weights and $u + v = 1$. If $q = \arg\min_j d_j$, then test sample y is assigned to the q-th class.

We summarize the algorithms below. First, the algorithm framework of the proposed method is presented by the algorithm table named "Algorithm 1".