Virtual Training Samples and CRC based Test Sample Reconstruction and Face Recognition Experiments

Wei HUANG and Li-ming MIAO
School of Computer and Information Engineering, Hanshan Normal University, 521041, Chaozhu, China

Keywords: Face recognition, Sparse representation, CRC, MIRROR.

Abstract. Sparse Representation (SR) has the merit to associate each test sample properly with the training samples. Collaborative representation classification (CRC) is a well-known generalized SR method and has achieved outstanding performance in Face Recognition (FR). In this paper, we propose an improvement to CRC, which combines the original training sample and mirror virtual face to form a new training set, uses this new training set to rebuild the test sample and then performs a two-step classification. The face recognition experiments show that the proposed method outperforms CRC and has certain robustness.

Introduction

In the past twenty years, face recognition (FR) has received increasing attention owing to its wide potential applications, such as identity authentication, society security, surveillance, human-computer interface. However, researchers confront some difficulties in face recognition due to great variations in appearance of samples caused by illuminations, expressions, age, postures and so on. Though numerous methods have been proposed to solve these problems as reviewed in, face recognition is still a great challenge [1].

Sparse representation (SR) has gained much attention in face recognition [2,3]. SR is indeed one of the state-of-the-art image classification methods. SR has the merit to properly associate each test sample with the training samples, because it determines relation for every test sample and the set of training samples. In particular, SR achieves this by producing a solution for each test sample. Differing from the statistical face recognition methods such as Eigenface and Fisherface, the face recognition methods based on SR are able to take the test sample into account when obtaining their solutions. Up to now, there have been a number of SR methods with different aims. These SR methods have been applied to various tasks such as feature selection [4], image super-resolution [5], image recovery [6,8] and image denoising [7].

In general, SRC requires that the $l_1$ norm of the coefficient vector of the linear combination be as small as possible. On the other hand, many researchers also proposed $l_2$ norm based representation method such as collaborative representation based classification (CRC) [9] and the two-phase test sample sparse representation (TPTSSR) [10]. CRC has the advantages of efficient computation and high accuracy. CRC has achieved good performance in pattern recognition [11,12]. However, it seems that CRC suffers from a drawback. When CRC represents the test sample via a linear combination of all training samples in face recognition, the representation results may have great difference from the test sample. The main reason is that the dimension of the face image is usually much greater than the number of training samples, so the test sample cannot be accurately expressed by a linear combination of all training samples. Similar analysis has been shown for in linear regression classification [3].

It is assumed that each class has enough training samples in many face recognition methods. There are many variations in face images including illuminations, postures, expressions and et al. If the training set cannot including all these variations, the test samples maybe far from the reconstructions by some or all training samples linearly. In the other words, FR is a typical small sample-size problem, and the training sample set is under-complete in general. If we use it to represent the test sample, the representation residual can be great. This of course will affect the
accuracy of face recognition. This problem can be overcome if more samples can be used to
represent the test sample, but the key is how to obtain additional samples. Intuitively, the faces
of all human beings have a similar image structure, which implies different persons, should also
have similar intrapersonal variations. In other words, one class can “borrow” samples from other
classes in order to faithfully represent the query sample, and we also can generate some new virtual
samples such as symmetrical sample [13] and mirror face samples [14]. In FR, for each class we
may consider the samples from similar classes or the virtual samples and use them to better
represent the query sample.

In this paper, we propose an improvement to CRC, which combines the original training sample
and mirror virtual face to form a new training set. Then the test samples will be reconstructed by
the new training set. At last the two-step classification algorithm is performed to improve the
recognition accuracy furthermore. There two aspects to improve the recognition accuracy. On the
one hand, we enlarge the training set by mirror virtual face image. On the other hand, we use the
TPTSSR classification algorithm to reduce the impact of lack training samples. The efficiency
experiments show that the proposed method outperforms CRC and has certain robustness.

The rest of this paper is organized as follows. Section 2 demonstrates the CRC method. Section 3
discusses our proposed method in detail. Section 4 conducts extensive experiments to demonstrate
the performance of our works, and Section 5 concludes the paper.

The CRC Method

Zhang et al. in reference [9] pointed that collaborative representation was the real power of SRC. As
a typical small-sample-size problem, the training set of FR is always not enough to represent the
testing sample. Collaborative representation classification is proposed to represent y collaboratively
by using all the training samples from different classes. Assume \( X_0 \) includes all the
training samples from all the classes. \( \Psi = [\alpha_1, \alpha_2, \cdots, \alpha_k, \cdots, \alpha_c] (\alpha_k \in R^{kn_k}, n_k \) is the number
of training samples from the \( k^{th} \) class) is the coefficient of \( X \). So the coefficient vector \( \Psi \) can be
estimated by

\[
\Psi = (X_0^T X_0 + \lambda I)^{-1} X_0^T y
\]

where \( \lambda \) is a positive small constant and we set \( \lambda = 0.01 \) in this paper. The representation
residual \( d_k \) can be calculated as follows:

\[
d_k = \| y - \alpha_k X_k \|^2 \quad k = 1, 2, \cdots, c
\]

The rule in favor of the class with minimum distance can be calculated by:

\[
\text{identity}(y) = \arg \min_k d_k
\]

Proposed Method

In this section we formally present the proposed method. In general FR is a typical small
sample-size problem, and the training sample set is under-complete. So we create the virtual mirror
(using Eq. 4) for each training sample to form the new training set. Also, the virtual mirror sample
can increase the post variation samples. After that the test sample is represented by the new training
set. At last the CRC is used to identify the test sample. The detail of the proposed method is shown
in Table 1.
Table 1. The proposed algorithm.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Normalize the columns of $X_0$ to have unit L2-norm. $X_0$ is the original training sample set.</td>
</tr>
<tr>
<td>2.</td>
<td>Compute a mirror virtual training sample $m$ of original sample $r$ by</td>
</tr>
<tr>
<td></td>
<td>$m(i, j) = r(i, q - j + 1)$ (4) where $i = 1, \ldots, p$, $j = 1, \ldots, q$, $p$ and $q$ stand for the numbers of rows and columns of $r$ respectively. $r(i, j)$ denotes the pixel located in the $i^{th}$ row and $j^{th}$ column of the original sample. The mirror sample set $M$ is made of all mirror samples. Then these two sets are combined to produce a new one $\hat{X} = [X_0 \ M]$.</td>
</tr>
<tr>
<td>3.</td>
<td>Rebuild test sample using combined training sample set by CRC</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha} = (\hat{X}^T \hat{X} + \lambda I)^{-1} \hat{X}^T \hat{y}$ (5)</td>
</tr>
<tr>
<td></td>
<td>$\hat{y} = \hat{\alpha}X$ (6)</td>
</tr>
<tr>
<td>4.</td>
<td>Compute regularized residuals of $\hat{y}$</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha} = (\hat{X}^T \hat{X} + \lambda I)^{-1} \hat{X}^T \hat{y}$ (7)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_k = | \hat{y} - \hat{\alpha}_k \hat{X}_k |^2, \ k = 1, 2, \ldots, c$ (8)</td>
</tr>
<tr>
<td></td>
<td>$\hat{X}_k$ denotes the set of training samples from class $k$ and $\hat{\alpha}_k$ denotes the coefficient from class $k$.</td>
</tr>
<tr>
<td>5.</td>
<td>Output the identity of $y$ as</td>
</tr>
<tr>
<td></td>
<td>$\text{identify}(y) = \arg \min { \hat{\alpha}_k }$ (9)</td>
</tr>
</tbody>
</table>

**Experiments**

In this section, we will implement our method in two standard face databases, FERET face database and AR face database which are usually used to test the performance of face recognition methods. The compared algorithm is CRC which is one of the stat of the art face recognition algorithms.

A subset from the cropped FERET database [15] consisting of 1400 images from 200 subjects is used here. The variations of images include facial expression, illumination and face posture. Now FERET face database has become a standard database for testing and evaluating state-of-the-art face recognition algorithm. The original and virtual mirror samples of one subject are shown in Fig. 1.

![Figure 1. Some original samples and virtual mirror samples on FERET database. The first row is the original samples of one subject the second row is the virtual mirror samples.](image)

From Fig. 1 we can see that the frontal face is similar with its virtual mirror face and the virtual mirror face will be a turning left face (see the second face in the second row) when the original face is a turning right image (see the second face in the first row). That means the virtual mirror faces can enlarge the training set and overcome pose problem in some degree.

In our experiment, the first $p$ samples per subject of non-occluded frontal views with various facial expressions were used for training, while the others were used for testing. The images were resized to 32×32. The experiment results are shown in Table 2.
Table 2. The accuracy rate in Feret database.

<table>
<thead>
<tr>
<th>$p$</th>
<th>CRC</th>
<th>The proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.576</td>
<td>0.6390</td>
</tr>
<tr>
<td>3</td>
<td>0.455</td>
<td>0.6062</td>
</tr>
<tr>
<td>4</td>
<td>0.533</td>
<td>0.5917</td>
</tr>
</tbody>
</table>

From Table 2 we can find that the proposed method has a big improvement compared with CRC in any case. At most, the recognition rate of the proposed method is superior to that of CRC by 33.2% when $p=2$. Also, the proposed method has 11.01% and 10.94% improvement when $p=3$ and $p=4$ respectively.

A subset from the cropped AR database [16] consisting of 3,120 images from 120 subjects, is used here. The AR face image database includes 960 images (about 14 samples per subject) of non-occluded frontal views with various facial expressions, 720 images (6 samples per subject) of sunglasses and 720 images (6 samples per subject) of scarves (as shown in Fig. 2). All the images were resized to 32×32. In this experiment the first $p$ samples per subject were used for training and the others were used for testing. The experiment results are shown in Fig. 2.

![Figure 2. Some original samples and virtual mirror samples on AR database. The first row is the original samples of one subject the second row is the virtual mirror samples.](image)

The compared experiment results are show in Fig. 3.

Fig. 3 shows the experiment results which are compared with CRC in AR face databases.

![Figure 3. The accuracy rate in AR database.](image)

From Fig. 3 we can learn that the accurate recognition accuracies of the proposed method are always higher than CRC in AR face database. Also, in FERET database the proposed method is more excellent than that in AR database. The experiment results show that virtual mirror sample and the two-step classification all contribute to performance improvement of the proposed method.

**Conclusion**

In this paper we proposed a improvement CRC method, which uses a combined training set which is not only include the original training samples but also include their mirror virtual samples to rebuild the test sample by collaborative representation classification (CRC). The TPTSSR classification algorithm is used to improve the performance of the proposed method furthermore. To sum up, there are two aspects in our method to reduce the influence which is brought by the lack of training samples. First we increase the number of the training sample set. The proposed method utilizes the symmetry of face image to generate the new virtual mirror samples. So the new training sample set includes the original training samples and the virtual mirror samples. Secondly, using the new training set to represent the test sample can reduce the noise and the brutal residual. So the TPTSSR classification algorithm is performance to improve the robustness of the proposed method.
furthermore. The experiment results on FERET and AR face database show that the proposed method outperforms CRC and has certain robustness.

Acknowledgement

This work is partially supported by the Natural Science Foundation of Guangdong Province (No. 2016A030307050), the Special Foundation of Public Research of Guangdong Province (No. 2016A020225008, No. 2017A040405062).

Reference