The Visual Computer
https://doi.org/10.1007/s00371-020-01827-3

ORIGINAL ARTICLE

Dissimilarity-based nearest neighbor classifier for single-sample face recognition

Zhengqi Zhang[1] · Li Zhang[1,2] · Meng Zhang[1]

Abstract
In single-sample face recognition (SSFR) tasks, the nearest neighbor classifier (NNC) is the most popular method for its simplicity in implementation. However, in complex situations with light, posture, expression, and obscuration, NNC cannot achieve good recognition performance when applying common distance measurements, such as the Euclidean distance. Thus, this paper proposes a novel distance measurement scheme for NNC and applies it to SSFR. The proposed method, called dissimilarity-based nearest neighbor classifier (DNNC), first segments each (training or test) image into non-overlapping patches with a given size and then generates an ordered image patch set. The dissimilarities between the given test image patch set and the training image patch sets are computed and taken as the distance measurement of NNC. The smaller the dissimilarity of image patch sets is, the closer is the distance from the test image to the training image. Therefore, the category of the test image can be determined according to the smallest dissimilarity. Extensive experiments on the AR face database demonstrate the effectiveness of DNNC, especially for the case of obscuration.

Keywords Face recognition · Single-sample per person · Dissimilarity · Image partitioning

1 Introduction

Face recognition has become one of the hot topics in computer vision and pattern recognition because of its wide applications in law enforcement, information security, and access control. Researchers have proposed many effective face recognition methods [1–4]. In general, these traditional face recognition methods assume that each person has multiple training samples which can be used to extract the discriminative features during the training phase. However, in actual face recognition application scenarios such as verification of electronic passports and ID card photograph, the above assumption could not hold true because there is only a single-sample per person (SSPP) registered or recorded in these systems. Since there are not sufficient training samples for discriminant learning, many mainstream face recognition algorithms could not perform well in single-sample face recognition (SSFR) cases. In order to implement SSFR, a lot of effective methods have been proposed.

In traditional face recognition tasks, linear discriminant analysis (LDA) [5], support vector machine (SVM) [6], and the nearest neighbor classifier (NNC) [7] are common used classifiers and have achieved good performance. However, these classifiers cannot be directly applied to SSFR for one training sample in each class. Therefore, researchers have proposed many schemes to make these classifiers be efficient in SSFR. For example, Liu et al. [8] proposed a novel method called local similarity-based linear discriminant analysis (LS-LDA), which divides a face image into local blocks, classifies each local block, and then integrates all the classification results to make a final decision. Li et al. [9] divided a face image into sub-blocks and then used a multi-manifold judgment analysis algorithm to select features. Finally, SVM is adopted to classify these features of sub-blocks.

✉ Li Zhang
zhangliml@suda.edu.cn

Zhengqi Zhang
20174227025@stu.suda.edu.cn

Meng Zhang
20184227025@stu.suda.edu.cn

[1] School of Computer Science and Technology, Joint International Research Laboratory of Machine Learning and Neuromorphic Computing, Soochow University, Suzhou 215006, Jiangsu, China

[2] Provincial Key Laboratory for Computer Information Processing Technology, Soochow University, Suzhou 215006, Jiangsu, China

Published online: 06 March 2020

Because NNC has the advantages of simple and easy implementation, high accuracy, and high tolerance for outliers and noise, it has become the most widely used classifier in SSFR. In NNC, the commonly used distance measurements are the Euclidean distance, the Manhattan distance, and the Chi-square distance. Therefore, many methods directly use these NNCs based on different distances in the final step of classification. Pan et al. [10] used the locality preserving projection feature transfer, which handles conflicts effectively between small samples and high dimensions and then adopted NNC to perform classification. Wei et al. [11] proposed an NNC algorithm based on automatic weights to the K nearest patches. Pei et al. [12] developed a decision pyramid classifier (DPC) for solving the SSFP problem with large appearance variations (e.g., illumination, expression, and partly occlusions), which is a variant of NNC. Ding et al. [13] addressed a face recognition algorithm based on a kernel principal component analysis network (KPCANet) and then proposed a weighted voting method. Chu et al. [14] proposed a multiple feature subspace analysis (MFSA) approach, which takes advantage of facial symmetry. MFSA formulates SSFP as a MFSA problem by learning a feature subspace for each group, so that the confusion between interclass and intra-class variations in face patches is removed and more discriminative features can be extracted from each subspace. An adaptive generic learning (AGL) method was proposed, which uses multiple training sample sets to predict the intra-class divergence matrix and inter-class divergence matrix [15].

The classification step in those SSFP methods mentioned above uses the idea of nearest neighbor, or minimum distance. Moreover, some other methods adopt the idea of minimum reconstructed distance, which could be also considered as a kind of NNC. Sparse representation (SR) reconstructs the input data by a linear combination of a few items from a dictionary [16–18], which has been applied to SSFR. Lu et al. designed a discriminant multi-manifold analysis (DMMA) method [19], which partitions each enrolled image into several non-overlapping patches and then learns multiple feature spaces to maximize the manifold margins of different persons. NNC-based minimum reconstructed distance is used in the step of classification. Similar improvement method called supervised locality preserving multi-manifold (SLPMM) was proposed by Mehrasa et al. [20]. Zhuang et al. [21] proposed a method based on an SR-based classification framework for SSFR. Ji et al. [22] presented a collaborative probabilistic label (CPL) method based on a robust collaborative representation and a probabilistic graph model for SSFR, which solves the problem of lacking information to predict the variations in the query sample. CPL assumes that the generic training set is full enough with sufficient variations. However, this assumption may not always hold true for the uncontrolled noise or other defects in data. Yu et al. [23] proposed a discriminative multi-scale sparse coding (DMSC) model that works well when face images have obstructions. Wang et al. [24] proposed a new feature extraction method called HOG-LDB (histogram of oriented gradients–local difference binary) and combined it with SVDL (sparse variation dictionary learning) to recognize the probe images with different facial variations (e.g., illuminations, poses, expressions, and disguises).

In NNC, the minimum distance is the basic criterion for classification. Thus, the distance measurement plays a key role in the performance of NNC. It is well known that the Euclidean distance and the Manhattan distance are commonly used measurements. However, NNC has a lower average recognition accuracy without extracting features under the influence of complex conditions, such as expression, light, and occlusion. To address this issue, this paper introduces the method of dissimilarity measurement in [25] and modifies it to measure dissimilarities between ordered image patch sets, which is taken as the distance measurement of NNC. The proposed method, called dissimilarity-based nearest neighbor classifier (DNNC), first segments each image into non-overlapping image patches and then generates the corresponding set of ordered image patches. Next, we can calculate the dissimilarities between the given test image patch set and the training image patch sets. Therefore, the nearest neighbor of the test image is the training image which has the smallest dissimilarity with it.

This paper is organized as follows: In Sect. 2, work related to NNC and dissimilarity measurement is reviewed. Section 3 proposes the dissimilarity-based nearest neighbor classifier and describes it in detail. In Sect. 4, extensive experiments are conducted to demonstrate the superiority of DNNC. Section 5 concludes this paper.

2 Related work

2.1 Nearest neighbor classifier

The nearest neighbor classifier based on distance metric is one of the most popular classification methods in machine learning [26,27]. The effectiveness and simplicity of NNC make it have good applications in general application scenarios.

For a given training set $\{(\mathbf{u}_i, v_i) | i = 1, 2, \ldots, l\}$, where $\mathbf{u}_i \in \mathbb{R}^D$ represents the feature vector of the ith sample, v_i represents the label of this sample and D and l are the number of features and samples, respectively. For a test sample $\mathbf{u}' \in \mathbb{R}^D$ that is to be classified, NNC is to find a training sample $\mathbf{u}^*$ closest to $\mathbf{u}'$ based on the distance measure and then assign the label of $\mathbf{u}^*$ to $\mathbf{u}'$.

There are many distance measurements for NNC. The commonly used distance metrics are the Euclidean distance

and the Manhattan distance. The Euclidean distance between $\mathbf{u}'$ and $\mathbf{u}_i$ can be expressed as follows:

$$d(\mathbf{u}', \mathbf{u}_i) = \sqrt{\sum_{j=1}^{D} \left(u'_j - u_{ij}\right)^2} \tag{1}$$

where u_{ij} is the jth feature of $\mathbf{u}_i$. The Manhattan distance between $\mathbf{u}'$ and $\mathbf{u}_i$ can be calculated by

$$d(\mathbf{u}', \mathbf{u}_i) = \sum_{j=1}^{D} |u'_j - u_{ij}| \tag{2}$$

The label of $\mathbf{u}'$ can be estimated by

$$\hat{v} = v_{i^*} \tag{3}$$

where

$$i^* = \arg \min_{i=1,\ldots,l} d(\mathbf{u}', \mathbf{u}_i)$$

In SSFR, the distance metric is equivalent to measuring the difference (or dissimilarity) between the test sample and the training samples.

2.2 Dissimilarity measurement

Kano et al. [25] proposed a statistical trailer monitoring method based on the dissimilarity of process data. The method of dissimilarity measurement uses the original data sets to calculate dissimilarity between two data sets.

In what follows, we describe in detail how to measure the dissimilarity between two data sets using the method mentioned in [25]. Without loss of generality, assume we have two arbitrary data sets A_1 and A_2. We represent these sets in matrices, or $\mathbf{a}_1 \in \mathbb{R}^{l \times D}$ and $\mathbf{a}_2 \in \mathbb{R}^{l \times D}$, respectively, where l is the number of samples in both data sets and D is the dimension of samples.

First, we need to calculate a joint matrix $\mathbf{R}$ for the two data sets. Namely,

$$\mathbf{R} = \frac{l-1}{2l-1}\mathbf{R}_1 + \frac{l-1}{2l-1}\mathbf{R}_2 \tag{4}$$

where

$$\mathbf{R}_t = \frac{1}{l-1}\mathbf{a}_t^T \mathbf{a}_t, \quad t = 1, 2 \tag{5}$$

Next, we perform the eigenvalue decomposition of $\mathbf{R}$ and have

$$\mathbf{R}\mathbf{P}_0 = \mathbf{P}_0 \mathbf{\Lambda} \tag{6}$$

where $\mathbf{P}_0$ is the orthogonal eigenvectors and $\mathbf{\Lambda}$ is the diagonal eigenvalue matrix. The whitening transformation matrix $\mathbf{P}$ can be defined as follows:

$$\mathbf{P} = \mathbf{P}_0 \mathbf{\Lambda}^{-1/2} \tag{7}$$

Thus, we have $\mathbf{P}^T \mathbf{R} \mathbf{P} = \mathbf{I}$, where $\mathbf{I}$ is the identify matrix.

Then, we use the whitening transformation matrix $\mathbf{P}$ to transform the data matrix $\mathbf{a}_t$, or $\bar{\mathbf{a}}_t = \mathbf{a}_t \mathbf{P}$, $t = 1, 2$. On the basis of them, we construct two $\bar{\mathbf{a}}_t$-related matrices $\mathbf{S}_t$ and have

$$\mathbf{S}_t = \frac{1}{l-1}\bar{\mathbf{a}}_t^T \bar{\mathbf{a}}_t, \quad t = 1, 2 \tag{8}$$

By performing the eigenvalue decomposition of $\mathbf{S}_t$, we can get the eigenvalues $\lambda_j^{(t)}$, $j = 1, \ldots, D$ of $\mathbf{S}_t$, $t = 1, 2$. Because $\mathbf{S}_1 + \mathbf{S}_2 = \mathbf{I}$, we have the following equalities

$$\lambda_j^{(1)} = 1 - \lambda_j^{(2)}, \quad j = 1, \ldots, D \tag{9}$$

Finally, the following formula can be defined to evaluate the difference between two matrices $\mathbf{a}_1$ and $\mathbf{a}_2$:

$$d(\mathbf{a}_1, \mathbf{a}_2) = \frac{4}{D}\sum_{j=1}^{D}(\lambda_j - 0.5)^2 \tag{10}$$

where $\lambda_j = \lambda_j^{(1)}$ or $\lambda_j = \lambda_j^{(2)}$. The smaller $d(\mathbf{a}_1, \mathbf{a}_2)$ is, the more similar the data matrices $\mathbf{a}_1$ and $\mathbf{a}_2$ are.

3 Dissimilarity-based nearest neighbor classifier

This section details the proposed method, dissimilarity-based nearest neighbor classifier, whose framework is shown in Fig. 1. DNNC is a kind of image segmentation-based methods and consists of three steps: image segmentation, dissimilarity measurement, and classification. In the following, we describe these three steps in detail.

3.1 Image segmentation

In general, the first step in DNNC is image segmentation, which is to partition original images into non-overlapping patches. An arbitrary face image could be divided evenly into the vertical and horizontal directions. For example, we can obtain 4, 6, or 16 patches by segmenting the image in the grid of 2×2, 3×2, or 4×4, as shown in Fig. 2. Each individual patch has its own structure information such as eyes, nose, or eyebrows. On the whole, these patches contain location relationship information such as eye to eye, nose

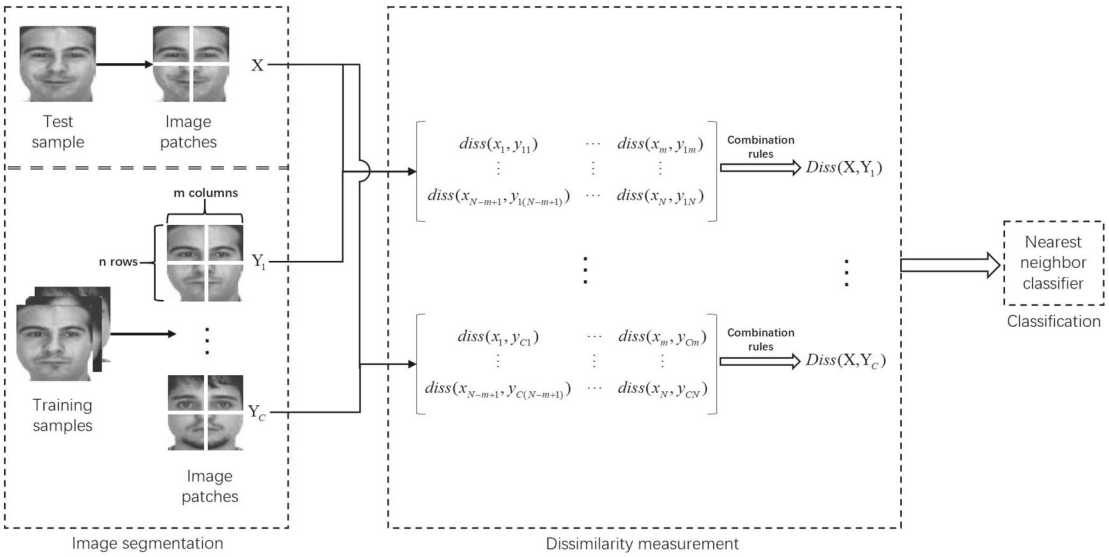

Fig. 1 Framework of DNNC

Fig. 2 Segmentation of a face image from the AR dataset [28]

to mouth, and so on. Thus, the patches and the information among them are very important for SSFR tasks.

Assume that there is a set of training images $\{\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_C\}$. Without loss of generality, we try to partition the training image $\mathbf{Y}_i$ into $n \times m$ patches. $Y_i = \{\mathbf{y}_{i1}, \mathbf{y}_{i2}, \ldots, \mathbf{y}_{iN}\}$ is the ordered set of image patches for $\mathbf{Y}_i$, where $\mathbf{y}_{ij} \in \mathbb{R}^{p\times q}$ is the jth patch of $\mathbf{Y}_i$, the number of patches $N = n \times m$, and $i = 1, 2, \ldots, C$. Note that $j = 1$ to N is in the order from top left to bottom right, which represents each patch of the original training image $\mathbf{Y}_i$ has totally N blocks.

For the given test image $\mathbf{X}$, we use the same segmentation way as the training images. In doing so, we have an ordered set of image patches $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ for $\mathbf{X}$.

3.2 Distance measurement

It is well known that the performance of NNC greatly depends on distance measurements without extracting features. Under complex situations with light, posture, expression, and obscuration, NNC cannot achieve good recognition performance when applying common distance measurements, such as the Euclidean distance and the Manhattan distance [12]. A dissimilarity measurement between two data sets has been described in [25]. We introduce this method and modify it to measure the distance between two ordered image patch sets.

In SSFR, the goal of NNC is to measure the distance (or dissimilarity) between the test image and the training images. In the following, we first describe the way of calculating the dissimilarity between image patches and then extend it to obtain the dissimilarity between the test image and the training image by fusing image patch sets.

3.2.1 Dissimilarity between image patches

Inspired by the measurement way on matrixes in [25], we introduce it to measure the dissimilarity between two image patches that can be taken as matrices.

In Sect. 2.2, we describe in detail how to measure the dissimilarity between two matrices. We can follow the way to directly calculate the dissimilarity between two image patches. Without loss of generality, assume we have two arbitrary image patches $\mathbf{z}_1 \in \mathbb{R}^{p\times q}$ and $\mathbf{z}_2 \in \mathbb{R}^{p\times q}$, where p and q are the number of rows and columns in an image patch, respectively. Namely,

$$\mathbf{z}_t = \begin{pmatrix} z_{11}^{(t)} & z_{12}^{(t)} & \cdots & z_{1q}^{(t)} \\ z_{21}^{(t)} & z_{22}^{(t)} & \cdots & z_{2q}^{(t)} \\ \vdots & \vdots & \ddots & \vdots \\ z_{p1}^{(t)} & z_{p2}^{(t)} & \cdots & z_{pq}^{(t)} \end{pmatrix}, \quad t = 1, 2 \tag{11}$$

In general, if a data set constitutes a sample matrix with $p \times q$, then p (or q) could be either the number of samples or the sample dimension. However, for an image patch with the size of $p \times q$, it is hard to say that p is the dimension or the number of samples. So does q. In view of this, we consider

two cases separately, namely the image patch itself and its transposed image, and take their average as the final result. Therefore, we need to construct two joint matrices based on the two patches and have

$$\mathbf{R} = \frac{1}{2p-1}\mathbf{z}_1^T\mathbf{z}_1 + \frac{1}{2p-1}\mathbf{z}_2^T\mathbf{z}_2 \tag{12}$$

and

$$\mathbf{R}' = \frac{1}{2q-1}\mathbf{z}_1\mathbf{z}_1^T + \frac{1}{2q-1}\mathbf{z}_2\mathbf{z}_2^T \tag{13}$$

On the basis of $\mathbf{R}$, the difference between matrices $\mathbf{z}_1$ and $\mathbf{z}_2$ can be calculated by

$$d(\mathbf{z}_1, \mathbf{z}_2) = \frac{4}{q}\sum_{j=1}^{q}(\lambda_j - 0.5)^2 \tag{14}$$

On the basis of $\mathbf{R}'$, the difference between matrices $\mathbf{z}_1^T$ and $\mathbf{z}_2^T$ can be described by

$$d(\mathbf{z}_1^T, \mathbf{z}_2^T) = \frac{4}{p}\sum_{j=1}^{p}(\lambda'_j - 0.5)^2 \tag{15}$$

To measure the dissimilarity between two image patches, we give a definition based on (14) and (15).

Definition 1 For two image patches $\mathbf{z}_1$ and $\mathbf{z}_2$, given the difference $d(\mathbf{z}_1, \mathbf{z}_2)$ between image patches and the difference $d(\mathbf{z}_1^T, \mathbf{z}_2^T)$ between the transposed image patches, the dissimilarity between $\mathbf{z}_1$ and $\mathbf{z}_2$ is defined as:

$$diss(\mathbf{z}_1, \mathbf{z}_2) = \frac{d(\mathbf{z}_1, \mathbf{z}_2) + d(\mathbf{z}_1^T, \mathbf{z}_2^T)}{2} \tag{16}$$

where $diss(\mathbf{z}_1, \mathbf{z}_2) \in [0, 1]$.

According to Definition 1, the value of $diss(\mathbf{z}_1, \mathbf{z}_2)$ must be close to 0 when two image patches are similar to each other. Contrarily, when the two image patches are completely different from each other, $diss(\mathbf{z}_1, \mathbf{z}_2)$ must be close to 1.

3.2.2 Dissimilarity between image patch sets

The dissimilarity between two image patches has been described above. Now, we discuss how to measure the dissimilarity between the given test image $\mathbf{X}$ and the training image $\mathbf{Y}_i$, or between the ordered image patch sets $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ and $Y_i = \{\mathbf{y}_{i1}, \ldots, \mathbf{y}_{iN}\}$.

We first calculate the dissimilarity between the image patches $\mathbf{x}_j$ and $\mathbf{y}_{ij}$ in terms of Definition 1 and then obtain $diss(\mathbf{x}_j, \mathbf{y}_{ij})$, $j = 1, \ldots, N$. To obtain the dissimilarity between $\mathbf{X}$ and $\mathbf{Y}_i$, we fuse the dissimilarities between image patches and design three combination rules, including the average rule, the maximum rule, and the minimum rule.

The average rule is to make an average on the generated dissimilarities between image patches, which is described in the following definition:

Definition 2 Given the dissimilarities $diss(\mathbf{x}_j, \mathbf{y}_{ij})$, $j = 1, \ldots, N$ between image patches of the test image $\mathbf{X}$ and the training image $\mathbf{Y}_i$, the dissimilarity between $\mathbf{X}$ and $\mathbf{Y}_i$ is defined as:

$$Diss(\mathbf{X}, \mathbf{Y}_i) = \frac{1}{N}\sum_{j=1}^{N} diss(\mathbf{x}_j, \mathbf{y}_{ij}) \tag{17}$$

where $Diss(\mathbf{X}, \mathbf{Y}_i) \in [0, 1]$.

The maximum rule is to find the maximum dissimilarity between image patches, which is described in Definition 3:

Definition 3 Given the dissimilarities $diss(\mathbf{x}_j, \mathbf{y}_{ij})$, $j = 1, \ldots, N$ between image patches of the test image $\mathbf{X}$ and the training image $\mathbf{Y}_i$, the dissimilarity between $\mathbf{X}$ and $\mathbf{Y}_i$ is defined as:

$$Diss(\mathbf{X}, \mathbf{Y}_i) = \max_{j=1,2,\ldots,N} diss(\mathbf{x}_j, \mathbf{y}_{ij}) \tag{18}$$

where $Diss(\mathbf{X}, \mathbf{Y}_i) \in [0, 1]$.

The minimum rule is to find the minimum dissimilarity between image patches, which is defined in Definition 4:

Definition 4 Given the dissimilarities $diss(\mathbf{x}_j, \mathbf{y}_{ij})$, $j = 1, \ldots, N$ between image patches of the test image $\mathbf{X}$ and the training image $\mathbf{Y}_i$, the dissimilarity between $\mathbf{X}$ and $\mathbf{Y}_i$ is defined as:

$$Diss(\mathbf{X}, \mathbf{Y}_i) = \min_{j=1,2,\ldots,N} diss(\mathbf{x}_j, \mathbf{y}_{ij}) \tag{19}$$

where $Diss(\mathbf{X}, \mathbf{Y}_i) \in [0, 1]$.

3.3 Classification and algorithm description

The smaller the dissimilarity $Diss(\mathbf{X}, \mathbf{Y}_i)$ is, the more similar the two images are. Thus, $Diss(\mathbf{X}, \mathbf{Y}_i)$ could be taken as the distance between $\mathbf{X}$ and $\mathbf{Y}_i$. After calculating the dissimilarity values of the test image $\mathbf{X}$ and all $\mathbf{Y}_i$ in the training image set, DNNC assigns the test image $\mathbf{X}$ the class label

$$k = \arg\min_{i=1,\ldots,C} Diss(\mathbf{X}, \mathbf{Y}_i) \tag{20}$$

The detailed algorithm of DNNC is described in Algorithm 1.

4 Experimental design and results analysis

In this section, we conduct a large number of comparative experiments on the AR facial database [28] to analyze the superiority of our proposed DNNC compared with other SSFR methods. All numerical experiments are performed on a personal computer with a 3.6 GHz Intel(R) Core(TM) i7-7700 and 8G bytes of memory. This computer runs Windows 10, with MATLAB 8.3.0.532 (R2014a) and VC++ 14.0 installed.

4.1 Data description

The AR facial database contains more than 4000 color facial images of 126 people (70 men and 56 women), and each person has a frontal face view with different facial expressions, lighting conditions, and occlusion (sunglasses and scarves). Each person has 26 different images taken in two shots (two weeks apart) and in the size of 768×576. In the following experiments, face images of 100 different subjects (50 men and 50 women) are signed as ARs, and each object has 26 images. In addition, each image is cropped to 60×60. Figure 3 shows the 26 faces of the first person in this database.

Without loss of generality, we put the 1st, 2nd, ..., 26th images of each person into the corresponding image subsets. Then, we get 26 subsets from A to Z as shown in Fig. 3, and each subset contains 100 face images of different people. The detailed description about these 26 subsets is given in Table 1.

Algorithm 1: DNNC

Input: The set of training images $Y = \{\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_C\}$, $i = 1, 2, \ldots, C$; a test image $\mathbf{X}$; the number of patches N;
Output: The estimated label k for $\mathbf{X}$;
1: Partition training images into N non-overlapping patches and then generate training image patch sets $Y_i = \{\mathbf{y}_{i1}, \mathbf{y}_{i2}, \ldots, \mathbf{y}_{iN}\}$, $i = 1, \ldots, C$.
2: Partition the test image into N non-overlapping patches and then generate the test image patch set $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$;
3: **for** $i = 1$ to C **do**
4: **for** $j = 1$ to N **do**
5: Calculate the difference $d(\mathbf{x}_j, \mathbf{y}_{ij})$ using (14);
6: Calculate the difference $d(\mathbf{x}_j^T, \mathbf{y}_{ij}^T)$ using (15);
7: Calculate the dissimilarity $diss(\mathbf{x}_j, \mathbf{y}_{ij})$ using (16);
8: **end for**
9: Calculate the dissimilarity $Diss(\mathbf{X}, \mathbf{Y}_i)$ according to the average rule (17) or the maximum rule (18) or the minimum rule (19);
10: **end for**
11: Assign the estimated label k to $\mathbf{X}$ using (20);
12: **return** k;

4.2 Experiments on expression change

4.2.1 Analysis of combination rules for DNNC

To observe the effect of combination rules on the performance of DNNC, we perform experiments on eight expression variation subsets, including subsets A (normal), B (smiling), C (angry), D (screaming), N (normal), O (smiling), P (angry), and Q (screaming). In the eight subsets, images are taken from two different time periods and have different expressions. Table 1 provides detailed information for expression variation subsets.

We discuss three combination rules: the average, maximum, and minimum rules for DNNC. The subset A is used for training, and the other seven subsets are used for test. Here, three image segmentation strategies are adopted. Each original image is divided into 3×3, 4×4, or 5×5. In other words, all 800 original images are divided into 9 patches, 16 patches, or 25 patches.

Table 2 lists the comparison of three different sizes for DNNC in different partitions. From Table 2, we can see that among the three different combination rules, the average rule has the highest average recognition accuracy. Meanwhile, the average rule can achieve better accuracies in all eight subsets. For the subset Q with large variations in expression, the recognition accuracy is very low. However, the experimental results show that DNNC can obtain a relatively higher accuracy when using the average combination rule. Therefore, the following experiments use the average combination rule.

4.2.2 More compared methods

In order to prove that DNNC can still achieve better recognition results in the situation of changing expressions, we compare our method with 13 state-of-the-art algorithms which can be used to address the SSPP face recognition problem, including PCA [33], $(PC)^2A$ (projected-combined PCA) [34], $E(PC)^2A$ (enhanced projected-combined PCA) [35], 2DPCA (two-dimensional PCA) [36], $(2D)^2PCA$ (two-directional two-dimensional PCA) [37], SOM (self-organizing map) [38], LPP(locality preserving projection) [39], SVD-LDA (singular value decomposition-based linear discriminant analysis) [40], block-PCA [41], block-LDA [42], UP (uniform pursuit) [43], DMMA [19], SLPMM [20], and CPL [44].

We follow the experimental setting in [19] and [20]. The original image is divided into 5×5 patches, and the subset A is taken as training samples. We compare these methods on some expression variation subsets (subsets B, C, D, N, O, and P. The results of SLPMM are from Ref. [20], CPL are from Ref. [44], and the other 12 methods are from Ref. [19].

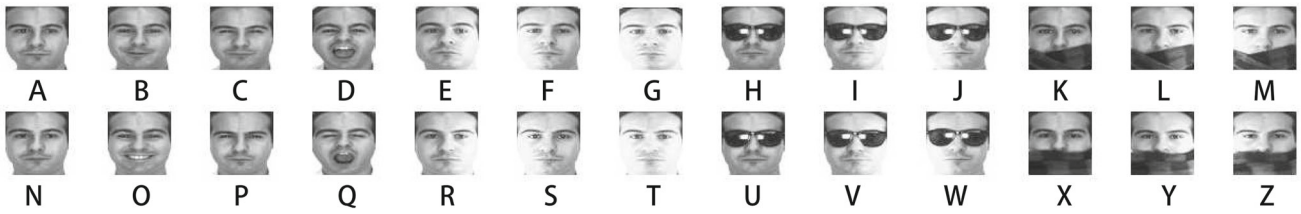

Fig. 3 26 facial images of the first person in the ARs database

Table 1 Description on 26 subsets of the ARs database

Subset	Collection condition
A	The first shot, normal expression
B	The first shot, smiling facial expression
C	The first shot, angry expression
D	The first shot, screaming expression
E	The first shot, normal expression, right light
F	The first shot, normal expression, left light
G	The first shot, normal expression, frontal light
H	The first shot, with sunglasses, without light
I	The first shot, with sunglasses, right light
J	The first shot, with sunglasses, left light
K	The first shot, with scarf , without light
L	The first shot, with scarf, right light
M	The first shot, with scarf, left light
N	The second shot, normal expression
O	The second shot, smiling facial expression
P	The second shot, angry expression
Q	The second shot, screaming expression
R	The second shot, normal expression, right light
S	The second shot, normal expression, left light
T	The second shot, normal expression, frontal light
U	The second shot, with sunglasses, without light
V	The second shot, with sunglasses, right light
W	The second shot, with sunglasses, left light
X	The second shot, with scarf, without light
Y	The second shot, with scarf, right light
Z	The second shot, with scarf, left light

Table 3 gives the experimental results, where "–" denotes there is no experimental result for the corresponding method, and the highest results are in bold type. By comparison, we can find that on subsets B, C, N, and P, the recognition accuracy of the proposed DNNC method is much better than other methods in recent years. Although DNNC does not achieve the highest accuracy on subsets D and O, its average accuracy is the highest among 15 methods. This also shows that the proposed DNNC method can achieve good recognition performance in the situation of facial expression changes.

4.3 Experiments on object occlusion

This subsection discusses the effect of object occlusion on the performance of DNNC and compares DNNC with NNC based on Manhattan distance (NNC-M) and NNC based on Euclidean distance (NNC-E) methods.

There are 12 subsets containing occlusion, including subsets H (sunglasses), I (sunglasses), J (sunglasses), K (scarf), L (scarf), M (scarf), U (sunglasses), V (sunglasses), W (sunglasses), X (scarf), Y (scarf), and Z (scarf) as shown in Fig. 3. Subsets H, I, J, U, V, and W are occluded with sunglasses, and subsets K, L, M, X, Y, and Z are occluded with scarf. In these 12 subsets, each person's images are taken from two different time periods and have different lighting conditions. Each subset consists of 100 different objects, and the total number of face images is 1200. Table 1 also provides detailed information for object occlusion subsets. Similar to Sect. 4.2, we take the subset A as the training set and 12 object occlusion subsets as the test set. Let $n \times m$ be 3×3, 4×4, or 5×5, which means that the number of patches N is 9, 16, or 25.

We first discuss the occlusion with sunglasses. Table 4 gives the comparison of accuracy of three NNC methods in the condition of test images with sunglasses occlusion (subsets H, I, J, U, V, and W). As can be seen from Table 4, among the three NNC methods, our proposed DNNC has the highest average accuracy when test images are occluded with sunglasses.

Next, we discuss the occlusion with scarf. Table 5 shows the accuracy of three NNC methods when test images are occluded with scarf (subsets K, L, M, X, Y, and Z). From Table 5, it can be clearly seen that for the six subsets with scarf occlusion, DNNC is much better than the other two NNC methods.

Figure 4 shows a comparison of average recognition accuracy of three NNC methods when there is an object occlusion in test images including sunglasses and scarf. It is easy to see that DNNC is much better than NNC-E and NNC-M.

To further prove the effectiveness of our proposed method, we also compare DNNC with methods proposed in recent years, including CRC [29], SVDL [30], DMMA [19], and SeetaFace [31]. For DNNC, let $n \times m$ be 5×5. In the same experimental environment as described in [32], Table 6 shows the accuracy of these methods on subsets H and K (occlu-

Table 2 Comparison of three combination rules for DNNC

$n \times m$	Combination rule	Accuracy (%)							Average accuracy (%)
		B	C	D	N	O	P	Q	
3 × 3	Minimum	82.00	76.00	43.00	59.00	28.00	36.00	9.00	47.57
	Average	98.00	96.00	64.00	89.00	67.00	72.00	27.00	73.29
	Maximum	53.00	63.00	16.00	47.00	32.00	36.00	12.00	37.00
4 × 4	Minimum	77.00	60.00	43.00	55.00	20.00	34.00	11.00	42.86
	Average	98.00	99.00	62.00	94.00	80.00	77.00	30.00	77.14
	Maximum	51.00	60.00	21.00	46.00	29.00	33.00	11.00	35.86
5 × 5	Minimum	74.00	63.00	35.00	57.00	24.00	43.00	5.00	43.00
	Average	99.00	99.00	67.00	95.00	74.00	79.00	31.00	77.71
	Maximum	50.00	51.00	15.00	37.00	20.00	22.00	10.00	29.29

Table 3 Comparison of 15 algorithms on expression change subsets

Method	Accuracy(%)						Average accuracy (%)	Year
	B	C	D	N	O	P		
PCA	97.00	87.00	60.00	77.00	76.00	67.00	77.33	1991
$(PC)^2A$	97.00	87.00	62.00	77.00	74.00	67.00	77.33	2002
$E(PC)^2A$	97.00	87.00	63.00	77.00	75.00	68.00	77.83	2004
2DPCA	97.00	87.00	60.00	76.00	76.00	67.00	77.17	2004
$(2D)^2PCA$	98.00	89.00	60.00	71.00	76.00	66.00	76.70	2005
SOM	98.00	88.00	64.00	73.00	77.00	70.00	78.30	2005
LPP	94.00	87.00	36.00	86.00	74.00	78.00	75.83	2005
SVD-LDA	73.00	75.00	29.00	75.00	56.00	58.00	61.00	2005
Block-PCA	97.00	87.00	60.00	77.00	76.00	67.00	77.33	2004
Block-LDA	85.00	79.00	29.00	73.00	59.00	59.00	64.00	2004
UP	98.00	88.00	59.00	77.00	74.00	66.00	77.00	2010
DMMA	99.00	93.00	69.00	88.00	**85.00**	85.50	79.00	2013
SLPMM	99.00	94.00	65.00	–	–	–	–	2017
CPL	92.22	88.06	**83.61**	83.59	77.95	72.82	83.04	2017
DNNC	**100.00**	**98.00**	69.00	**92.00**	76.00	**85.00**	**86.67**	

Table 4 Comparison of three NNC methods under sunglasses occlusion

$n \times m$	Method	Accuracy (%)						Average accuracy (%)
		H	I	J	U	V	W	
3 × 3	NNC-M	89.00	44.00	41.00	42.00	28.00	22.00	44.33
	NNC-E	70.00	51.00	43.00	29.00	30.00	20.00	40.50
	DNNC	92.00	67.00	53.00	57.00	31.00	26.00	54.33
4 × 4	NNC-M	89.00	44.00	41.00	42.00	28.00	22.00	44.33
	NNC-E	81.00	54.00	42.00	36.00	32.00	21.00	44.33
	DNNC	94.00	77.00	62.00	67.00	43.00	37.00	62.83
5 × 5	NNC-M	89.00	44.00	41.00	42.00	28.00	22.00	44.33
	NNC-E	85.00	54.00	42.00	39.00	35.00	22.00	46.17
	DNNC	97.00	82.00	78.00	66.00	54.00	42.00	69.83

Table 5 Comparison of three NNC methods under scarf occlusion

$n \times m$	Method	Accuracy (%)						Average accuracy (%)
		K	L	M	X	Y	Z	
3×3	NNC-M	17.00	21.00	18.00	6.00	12.00	9.00	13.83
	NNC-E	16.00	28.00	17.00	6.00	15.00	11.00	15.50
	DNNC	78.00	64.00	56.00	54.00	39.00	27.00	53.00
4×4	NNC-M	17.00	21.00	18.00	6.00	12.00	9.00	13.83
	NNC-E	14.00	27.00	18.00	8.00	13.00	12.00	15.33
	DNNC	84.00	67.00	63.00	59.00	37.00	29.00	56.50
5×5	NNC-M	17.00	21.00	18.00	6.00	12.00	9.00	13.83
	NNC-E	19.00	26.00	20.00	8.00	13.00	12.00	16.30
	DNNC	88.00	77.00	70.00	52.00	47.00	30.00	60.67

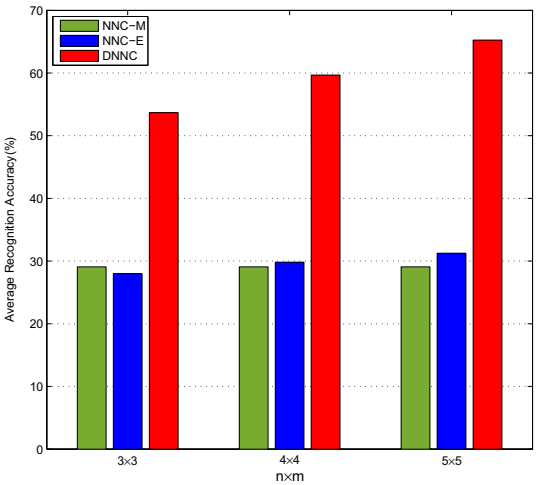

Fig. 4 Comparison of the average accuracy of three methods with object occlusion

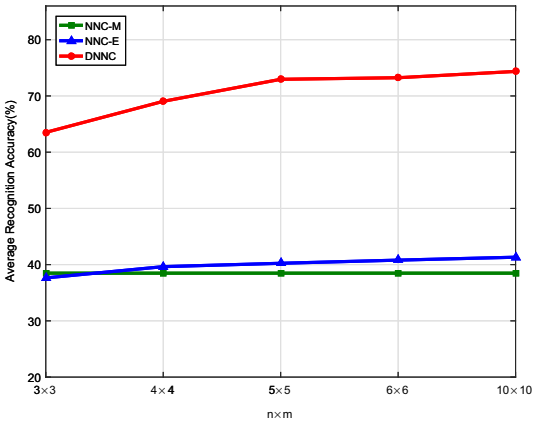

Fig. 5 Effect of patch number on the average accuracy of the three methods

sion), and subsets J and M (illumination+occlusion), where the highest values are in bold type. From Table 6, we can see that the recognition accuracy of DNNC is much better than other methods.

From Tables 4, 5 and 6, we can see that for the test images with object occlusion, the accuracy of DNNC is decreased. However, DNNC exhibits a relatively good performance when compared with other methods. In other words, the proposed dissimilarity measurement is effective for SSFR tasks, especially for the situation of occlusion.

4.4 Experiments on random subsets

4.4.1 Analysis of patch size for NNC methods

From the experimental results mentioned above, we know that the patch size has an effect on the performance of NNC-based methods. Therefore, we discuss the effect of the patch number N on the accuracy of DNNC, NNC-M, and NNC-E under the same situations. Consider all 2600 original images and five partition schemes. Let $n \times m$ be 3×3, 4×4, 5×5, 6×6, and 10×10, respectively.

Table 6 Comparison on occlusion and illumination+cclusion sets (session1)

Method	Occlusion	Illumination+occlusion	Average accuracy (%)	Year
CRC	58.10	23.80	40.95	2011
SVDL	86.30	79.40	82.85	2013
DMMA	46.90	30.90	38.90	2013
SeetaFace	63.13	55.63	59.38	2016
DNNC	**92.50**	**79.50**	**86.00**	

Table 7 Comparison of eight algorithms on random subsets

Method	# Test subsets				
	5	10	15	20	24
DMMA	66.52 ± 11.55	66.49 ± 3.15	65.27 ± 3.64	65.01 ± 2.54	64.15 ± 1.01
SRC	65.94 ± 10.41	67.08 ± 6.51	65.29 ± 3.81	71.20 ± 2.18	69.95 ± 0.78
CRC	69.22 ± 9.82	69.96 ± 5.07	68.43 ± 3.37	73.74 ± 1.91	72.52 ± 0.71
Block-LDA	56.00 ± 5.69	54.44 ± 3.99	55.09 ± 2.52	55.51 ± 2.14	55.98 ± 0.85
AGL	70.12 ± 6.56	69.50 ± 5.10	67.80 ± 3.61	68.32 ± 1.81	68.88 ± 0.96
ESRC	72.76 ± 7.61	74.57 ± 5.57	72.57 ± 2.63	73.74 ± 2.43	73.70 ± 0.87
DPC-1	75.90 ± 7.65	71.17 ± 3.45	71.95 ± 3.02	72.72 ± 2.59	72.49 ± 1.16
DNNC	**76.84** ± 4.94	**75.08** ± 6.46	**74.33** ± 3.85	**74.05** ± 2.61	**74.09** ± 0.79

The subset A is the training set, and the other 25 subsets are the test set. Experimental average results on 25 subsets are shown in Fig. 5. Observation on Fig. 5 implies that DNNC and NNC-E may need a larger patch size to achieve good performance, such as 10×10. In addition, the image patch size has a relatively small impact on the average accuracy of NNC-M.

4.4.2 More compared methods

To validate the performance of DNNC on all 25 subsets, we compare other more methods: DMMA [19], SRC [45], CRC [29], Block-LDA [42], AGL [15], ESRC [46], and DPC [12].

The results of the first six methods are from [47], because we exactly follow the experimental setting of [47]. For DNNC and DPC, the original image is divided into 10×10 patches. Considering the running time, DPC uses one layer of voting, called DPC-1. We choose the subset A for training and randomly select 5, 10, 15, 20, 24 subsets from the rest 25 subsets for test. The goal of experiments is to verify the stability of the compared methods. In theory, the number of test samples dose not have a large effect on the performance of algorithms.

Each random selection of test samples is repeated 10 times. The experimental results are shown in Table 7, which gives the average recognition accuracy and the corresponding standard deviation of different methods with different test sample numbers. In Table 7, the bold values indicate the best results under the same conditions. From Table 7, we can see that DNNC has the highest average recognition accuracy than all other methods on the AR facial set. Furthermore, DNNC is relatively more stable compared to other methods. In summary, the proposed DNNC is effective for SSFP tasks under complex changing situations, including varying light, expression, and occlusion.

5 Conclusions

For SSFR tasks, we introduce a dissimilarity measurement to NNC and propose a DNNC method. DNNC adopts the dissimilarity measurement to calculate the differences between patches of the given test image and training images. Three combination rules are designed for combining the multiple patch differences to obtain the dissimilarity between the test image and training images. Among three combination rules, we find that the average rule is the best for DNNC through experiments on the expression variation subsets. In the case of expression change, a lot of comparative experiments have proved the superiority of our proposed method. In the environment of object occlusion, a contrast test using other methods highlights the insensitivity of our method to object occlusion, which proves that DNNC has a good performance on the situation of object occlusion. These experiments show that our method can effectively overcome the difficulties in practical application scenarios. In order to further validate that DNNC is suitable for complex practical application scenarios, random sampling experiments performed on all 25 subsets further show the stability and superiority of the proposed DNNC method.

Experimental results show that DNNC with the average rule performs well. Actually, the average rule is a common way in ensemble learning. Since the minimum and maximum rules obtain bad performance, some patches may play a negative role in measuring dissimilarity of whole images. Therefore, sparse ensemble or select ensemble [48–50] for these patches maybe need further work.

Acknowledgements We would like to thank three anonymous reviewers and Editor Nadia Magnenat-Thalmann for their valuable comments and suggestions, which have significantly improved this paper. This work was supported in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant No. 19KJA550002, by the Six Talent Peak Project of Jiangsu Province of China under Grant No. XYDXX-054, by the Priority Academic Program Development of Jiangsu Higher Education Institutions, and by the Collaborative Innovation Center of Novel Software Technology and Industrialization.

References

1. Cai, Y., Xu, C.Y., Fan, J.L.: Application research of PCA algorithm for face recognition. Inf. Technol. **11**, 129–132 (2017)
2. Zhou, S.F., Dai, S.G.: Research on improved PCA face recognition algorithm. Softw. Guide **2**, 15–18 (2018)
3. Zhou, B., He, Y.Q., Wang, J.: Face recognition based on adaptive near-part local preserving projection algorithm. In: Advances in Laser and Optoelectronics, vol. 3 (2018)
4. Lin, J., Wu, X.A.: Research and application of improved face recognition algorithm based on virtual sample. Comput. Eng. Appl. **23**, 123–128 (2017)
5. Wu, H.P., Dai, S.K.: Face recognition of 2 DLDA based on ULBP eigensubspace. Pattern Recog. Artif. Intell. **10**, 894–899 (2014)
6. Qin J., He Z.S.: A SVM face recognition method based on Gabor-featured key points. In: International Conference on Machine Learning and Cybernetics, vol. 8 (2005)
7. Cheng, Y., Jiao, L.B., Cao, X.H., Li, Z.Y.: Illumination-insensitive features for face recognition. Vis. Comput. **33**, 1483–1493 (2017)
8. Liu, F., Bi, Y., Cui, Y., Tang, Z.M.: Local similarity based linear discriminant analysis for face recognition with single sample per person. In: Asian Conference on Computer Vision, vol. 9010, pp. 85–95 (2015)
9. Li, J.X., Zhang, S.M., Wu, H.S.: Single training sample face recognition based on image blocking and feature selection. Comput. Appl. Softw. **9**, 310–313 (2015)
10. Pan, J., Wang, X.S., Cheng, Y.H.: Single-sample face recognition based on LPP feature transfer. IEEE Access **4**, 2873–2884 (2016)
11. Wei, M.J., Xu, D.Y., Qin, Y.B.: Face recognition based on automatic weighted K nearest patches for single training sample. J. Front. Comput. Sci. Technol. **9**, 1505–1512 (2017)
12. Pei, T.W., Zang, L., Wang, B.J., et al.: Decision pyramid classifier for face recognition under complex variations using single sample per person. Pattern Recogn. **64**, 305–313 (2017)
13. Ding, C., Bao, T., Karmoshi, S., et al.: Single sample per person face recognition with KPCANet and a weighted voting scheme. Signal Image Video Process. **11**(7), 1213–1220 (2017)
14. Chu, Y.J., Zhao, L.D., Ahmad, T.: Multiple feature subspaces analysis for single sample per person face recognition. Vis. Comput. **35**(2), 239–256 (2019)
15. Yu, Y., Wang, J.H., Sun, H.J.: On applying adaptive generic learning framework to face recognition. Comput. Appl. Softw. **7**, 173–176 (2014)
16. Zhang, Z., Li, F.Z., Chow, T., et al.: Sparse codes auto-extractor for classification: a joint embedding and dictionary learning framework for representation. IEEE Trans. Signal Process. **64**(14), 3790–3805 (2016)
17. Zhang, Z., Jiang, W.M., Qin, J., et al.: Jointly learning structured analysis discriminative dictionary and analysis multiclass classifier. IEEE Trans. Neural Netw. Learn. Syst. **29**(8), 3798–3814 (2018)
18. Zhang, Z., Ren, J.H., Jiang, W.M., et al.: Joint subspace recovery and enhanced locality driven robust flexible discriminative dictionary learning. IEEE Trans. Circuits Syst. Video Technol. (2019)
19. Lu, J., Tan, Y.P., Wang, G.: Discriminative multimanifold analysis for face recognition from a single training sample per person. IEEE Trans. Pattern Anal. Mach. Intell. **35**(1), 39–51 (2013)
20. Mehrasa, N., Aghagolzadeh, A., Motameni, H.: A supervised multimanifold method with locality preserving for face recognition using single sample per person. J. Cent. South Univ. **24**(12), 2853–2861 (2017)
21. Zhuang, L., Chan, T.H., Yang, A.Y., et al.: Sparse illumination learning and transfer for single-sample face recognition with image corruption and misalignment. IEEE Comput. Vis. Pattern Recognit. **114**(2–3), 3546–3553 (2013)
22. Ji, H.K., Sun, Q.S., Ji, Z.X., et al.: Collaborative probabilistic labels for face recognition from single sample per person. Pattern Recogn. **62(C)**, 125–134 (2017)
23. Yu, Y.F., Dai, D.Q., Ren, C.X., et al.: Discriminative multi-scale sparse coding for single-sample face recognition with occlusion. Pattern Recogn. **66**, 320–312 (2017)
24. Wang, H., Zhang, D.S., Miao, Z.H.: Face recognition with single sample per person using HOG-LDB and SVDL. Signal Image Video Process. **13**(5), 985–992 (2019)
25. Kano, M., Hasebe, S., Hashimoto, I., et al.: Statistical process monitoring based on dissimilarity of process data. AIChE J. **6**, 1231–1240 (2002)
26. Ding, S.F., Xu, X., Fan, S.Y., Xue, Y.: Locally adaptive multiple kernel k-means algorithm based on shared nearest neighbors. Soft. Comput. **22**, 4573–4583 (2018)
27. Gou, J.P., Ma, H.X., Ou, W.H., et al.: A generalized mean distance-based k-nearest neighbor classifier. Expert Syst. Appl. **115**, 356–372 (2019)
28. Martinez A.M., Benavente R.: The AR face database. CVC Technical Report #24 (1998)
29. Zhang, L., Yang, M., Feng, X.: Sparse representation or collaborative representation: which helps face recognition? In: International Conference on Computer Vision. IEEE, pp. 471–478 (2012)
30. Yang, M., Van, L., Zhang, L.: Sparse variation dictionary learning for face recognition with a single training sample per person. In: IEEE International Conference on Computer Vision, pp. 689–696 (2013)
31. Seetafaceengine. https://github.com/seetaface/SeetaFaceEngine (2016)
32. Zhu, P., Yang, M., Zhang, L., Lee, I.Y.: Local generic representation for face recognition with single sample per person. In: Asian Conference on Computer Vision, pp. 34–50 (2014)
33. Turk, M., Pentland, A.: Eigenfaces for recognition. Cognit. Neurosci. **3**(1), 71–86 (1991)
34. Wu, J., Zhou, Z.: Face recognition with one training image per person. Pattern Recogn. Lett. **23**(14), 1711–1719 (2002)
35. Chen, S., Zhang, D., Zhou, Z.: Enhanced $(PC)^2A$ for face recognition with one training image per person. Pattern Recogn. Lett. **25**(10), 1173–1181 (2004)
36. Yang, J., Zhang, D., Frangi, A., et al.: Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Anal. Mach. Intell. **26**(1), 131–137 (2004)
37. Zhang, D., Zhou, Z.: $(2D)^2PCA$: two-directional two-dimensional PCA for efficient face representation and recognition. Neurocomputing **69**(1–3), 224–231 (2005)
38. Tan, X., Chen, S., Zhou, Z., et al.: Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft K-NN ensemble. IEEE Trans. Neural Netw. **16**(4), 875–886 (2005)
39. He, X., Yan, S., Hu, Y., et al.: Face recognition using Laplacianfaces. IEEE Trans. Pattern Anal. Mach. Intell. **27**(3), 328–340 (2005)
40. Zhang, D., Chen, S., Zhou, Z.: A new face recognition method based on SVD perturbation for single example image per person. Appl. Math. Comput. **163**(2), 895–907 (2005)
41. Gottumukkal, R., Asari, V.: An improved face recognition technique based on modular PCA approach. Pattern Recogn. Lett. **25**(4), 429–436 (2004)
42. Chen, S., Liu, J., Zhou, Z.H.: Making FLDA applicable to face recognition with one sample per person. Pattern Recogn. **37**(7), 1553–1555 (2004)
43. Deng, W., Hu, J., Guo, J., et al.: Robust, accurate and efficient face recognition from a single training image: a uniform pursuit approach. Pattern Recogn. **43**(5), 1748–1762 (2010)

44. Ji, H.K., Sun, Q.S., Ji, Z.X., et al.: Collaborative probabilistic labels for face recognition from single sample per person. Pattern Recogn. **62(C)**, 125–134 (2017)
45. Wright, J., Yang, A., V, A., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31(2)**, 210–227 (2009)
46. Deng, W., Hu, J., Guo, J.: Extended SRC: undersampled face recognition via intraclass variant dictionary. IEEE Trans. Pattern Anal. Mach. Intell. **34**(9), 1864–70 (2012)
47. Ding, R.X., Du, D.K., Huang, Z.H., et al.: Variational feature representation-based classification for face recognition with single sample per person. J. Visual Commun. Image Represent. **30**, 35–45 (2015)
48. Zhang, L., Zhou, W.D.: Sparse ensembles using weighted combination methods based on linear programming. Pattern Recogn. **44**(1), 97–106 (2011)
49. Fakhari, F., Mosavi, M.R., Lajvardi, M.M.: Image fusion based on multi-scale transform and sparse representation: an image energy approach. IET Image Proc. **11**(11), 1041–1049 (2017)
50. Xiang, F.T., Jian, Z., Liang, P., et al.: Robust image fusion with block sparse representation and online dictionary learning. IET Image Proc. **12**(3), 345–353 (2018)

Zhengqi Zhang was born in Suzhou, Jiangsu Province. He is currently pursuing the Master's degree at the School of Computer Science and Technology, Soochow University, Suzhou, China. His research interests include machine learning, data mining, and face recognition.

Li Zhang received the B.S. degree in 1997 and the Ph.D. degree in 2002 in electronic engineering from Xidian University, Xi'an, China. Now she is a full professor with the School of Computer Science and Technology, Soochow University, Suzhou, China. She was a postdoctor at the Institute of Automation, Shanghai Jiao Tong University, Shanghai, China, from 2003 to 2005. She worked as an associate professor at the Institute of Intelligent Information Processing, Xidian University, Xi'an, China, from 2005 to 2010. She was a visiting professor at Yuan Ze University, Taiwan, from February to May 2010. Her research interests have been in the areas of machine learning, pattern recognition, neural networks, and intelligent information processing. She published more than 100 papers in journals and conferences and holds 23 patents. She is a member of the IEEE, a Senior Member of China Computer Federation (CCF), and a Senior Member of Chinese Association for Artificial Intelligence (CAAI).

Meng Zhang received the B.S. degree in the School of Computer Science and Technology, Soochow University, Suzhou, China. Now, she is pursuing the Master's degree. Her research interests include machine learning, data mining, and pattern recognition.