Dynamic Dictionary Optimization for Sparse-representation-based Face Classification using Local Difference Images

Changbin Shao[a], Xiaoning Song[b,c,*], Zhen-Hua Feng[d], Xiao-Jun Wu[b], Yuhui Zheng[e]

[a]*School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, P.R. China*
[b]*School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, P.R. China*
[c]*Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing University of Science and Technology, Nanjing, 210094, P.R. China*
[d]*Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford GU2 7XH, U.K.*
[e]*School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, 210044, P.R. China*

Abstract

In this study, we present a new sparse-representation-based face-classification algorithm that exploits dynamic dictionary optimization on an extended dictionary using synthesized faces. More specifically, given a dictionary consisting of face examples, we first augment the dictionary with a set of virtual faces generated by calculating the image difference of a pair of faces. This results in an extended dictionary with hybrid training samples, which enhances the capacity of the dictionary to represent new samples. Second, to reduce the redundancy of the extended dictionary and improve the classification accuracy, we use a dictionary-optimization method. We truncate the extended dictionary with a more compact structure by discarding the original samples with small contributions to represent a test sample. Finally, we perform sparse-representation-based face classification using the optimized dictionary. Experimental results obtained using the AR and FERRET face datasets demonstrate the superiority of the proposed method in terms of accuracy, especially for small-sample-size problems.

Keywords: Sparse representation, face classification, data augmentation, dictionary optimization

*Corresponding author
Email address: xnsong@hotmail.com, x.song@jiangnan.edu.cn (Xiaoning Song)

1. Introduction

In recent years, sparse representation has received extensive attention for its wide applications in signal processing, such as super-resolution reconstruction [51, 52], image segmentation [38, 41], signal encoding [12, 18], color image restoration [21], image denoising [9], and pattern recognition [34, 35, 40]. In contrast with traditional methods [4, 19, 24, 25, 26, 27, 55, 60], sparse representation has introduced a number of new methodologies with promising results to the aforementioned areas. A survey of sparse representation and its applications is presented in [59]. For pattern recognition and classification, the most well-known sparse representation method is believed to be sparse-representation-based classification (SRC) [43, 44], which performs classification using an approach that is different from that used by conventional classifiers, *e.g.*, support vector machines (SVMs) [14, 15, 39, 45], AdaBoost [29, 42], and other statistical methods [16, 30].

Given a test sample and a dictionary consisting of a number of training samples/atoms with labels, the aim of SRC is to reconstruct the test sample using a sparse linear combination of all the training samples. In order to achieve the sparsity of reconstruction coefficient vectors, the ℓ_1-norm regularization is used in SRC. By design, SRC considers all possible supports (samples from all the classes in the dictionary) and adaptively selects the minimal number of training samples to represent the test sample, in which the resulting reconstruction residuals provide discriminative information for classification among the training classes. In the classification step, SRC evaluates the reconstruction error of each class in the dictionary with the test sample, and assigns the label to the class with the minimum reconstruction error. As opposed to the ℓ_1-norm, the ℓ_2-norm regularization has also been studied in collaborative-representation-based classification (CRC) [57], which shows that the use of ℓ_2-norm regularization is more efficient and accurate for face recognition.

Sparse representation is naturally discriminative [3, 10, 20, 56, 62]. The accuracy of SRC is directly related to the representation capability of a dictionary and the sparsity of reconstruction coefficient vectors. Of the total number of training samples, SRC selects a subset that compactly reconstructs an input signal and rejects all the other less relative samples. Therefore, the ability of sparse representation to explore discriminative information depends on the design of a sparse-constrained signal reconstruction algorithm, and the construction of an effective dictionary that is capable of preserving important properties of various types of signals.

In this paper, we focus on the problem of face classification using the sparse-representation-based framework. Face classification/recognition is one of the most popular and challenging tasks in machine vision and pattern recognition. Despite the capability of SRC in exploring intrinsic characteristics using a training dataset, it cannot effectively address the issues posed by appearance variations. Moreover, in industrial applications, some products such as CCTV security systems usually offer few or even a single face image per subject for face verification, which is the well-known 'small-sample-size' (SSS) problem [6]. To mitigate these issues, a number of new approaches have been developed, which can be roughly divided into two categories. The first one is to improve the underlying model of the classical SRC pipeline. For example, Xu *et al.* [47] integrated conventional and inverse representations for face recognition, and obtained impressive face recognition results. Yang *et al.* and Zhu *et al.* [54, 61] introduced the similarity and distinctiveness of features to CRC, and presented a more general model. Chen *et al.* [5] extended SRC to 2D-SRC, which can directly handle 2D matrices of face images rather than converting them into vectors. The other important category is to augment an available dictionary with synthesized virtual faces. The use of synthesized face images has been shown to be beneficial for a number of face-analysis tasks, such as face recognition [32, 46, 49] and facial landmark detection [11, 13, 17]. For SRC-based face classification, Deng *et al.* augmented the original dictionary by generating intraclass variant faces [7, 8]. Ryu *et al.* [31] used the distribution of a given gallery set to generate virtual training samples, while Beymer *et al*. [1] and Vetter *et al*. [2] constructed new frontalized face images for face recognition. Finally, the symmetry property of human faces has also been widely used in SRC-based face classification and other face analysis applications [12, 36, 50].

The aforementioned methods indicate that virtual training samples can be generated in two ways: the first one is to extract non-trivial knowledge hidden in the intrinsic variation mode among training samples, and the second is to synthesize new training samples by performing image perturbation on original samples. Nevertheless, the use of synthesized faces may lead to issues in decision making owing to the information redundancy among original and synthesized faces. The reason is that the use of a large number of virtual samples may lead to over-fitting, and the relationship between different virtual samples cannot be precisely described in the reconstruction step. Thereby, heuristic or adaptive strategies are expected to discover a subset with the most competitive training samples for signal reconstruction.

In general, traditional sparse-representation-based classification methods exploit all of the training samples to represent a query image for classification, and are referred to as 'global-representation-based' approaches. In contrast, we refer to a method using only a subset of the original training samples as a 'local-representation-based' approach. As demonstrated by the two-phase CRC [48] and linear-regression-based classification (LRC) [23] methods, local-representation-based methods perform better than global methods in terms of both accuracy and efficiency. By design, a robust local-representation-based method should be able to convert a difficult face classification task into an easier one using an optimized dictionary. It can also be regarded as a specific evaluation method that uses merely a subset of training samples to represent and classify a query sample. In fact, the underlying rationale behind the use of local-representation-based methods has already been empirically proven, and is widely admitted. If a test sample highly correlates to the training samples obtained from a specific subject, it should be reasonable to assign the label of the test sample to the subject. Therefore, it is important to develop a new method that enables us to acquire more inherent sparse fidelity information from insufficient (even a single training sample per subject) training samples. To this end, we propose a two-step local-sparse-representation-based classification method, namely 'Two-Step LSRC'. In contrast, we use the term 'Two-Step SRC' for the classical SRC method that performs classification on an optimized dictionary without synthesized faces. The proposed method includes three main contributions to the field:

- We construct an extended dictionary using the original training samples and a set of synthesized virtual samples. A virtual face is generated using the local difference image between a pair of faces. With the help of the extended dictionary, we can better represent a test sample.

- Because an extended dictionary may over-fit test samples and lead to inaccurate decision making, we optimize the dictionary to reduce the information redundancy. To this end, we discard the training samples with small contributions to representing a test sample. Note that in general, we only discard the original training samples, while retaining all of the synthesized ones during the optimization step. We discard synthesized samples only when we have a very large number of synthesized faces. This is different from the elimination strategy used in [32, 33], in which they always discard synthesized samples.

- We optimized the extended dictionary by performing adaptive contribution measurement, and we used the optimized dictionary for robust face classification. We determined the class of a test sample using the class that yields the minimum reconstruction error. Experimental results obtained on the AR and FERET datasets demonstrate that the proposed method achieves significant improvement in accuracy for face classification, especially for the small-sample-size problem.

The remainder of the paper is organized as follows: In Section 2, we present a brief overview of the sparse-representation-based classification method, which serves as a prerequisite of the proposed Two-Step LSRC approach in Section 3. In Section 4, we give a comprehensive analysis of the proposed method, and in Section 5, we present the experimental results obtained for the AR and FERET face datasets. Finally, we conclude the paper in Section 6.

2. Sparse-representation-based classification

Given $K \times M$ training samples $\{\mathbf{x}_1, ..., \mathbf{x}_{KM}\}$, where $\mathbf{x} \in \mathbb{R}^D$, K is the number of classes, M is the number of samples of each class, and D is the dimensionality of a sample, SRC reconstructs a test sample $\mathbf{y} \in \mathbb{R}^D$ using a linear combination of all the training samples:

$$\mathbf{y} = \sum_{i=1}^{KM} \alpha_i \mathbf{x}_i + \mathbf{e}, \tag{1}$$

where α_i is the reconstruction coefficient corresponding to the ith training sample $\mathbf{x}_i$ and $\mathbf{e}$ is the residual. The above equation can be compactly rewritten as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{e}, \tag{2}$$

where $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_{KM}] \in \mathbb{R}^{D\times KM}$ is the training sample matrix, also known as the dictionary, and $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_{KM}]^T$ is the reconstruction coefficient vector.

In fact, SRC encodes the sample $\mathbf{y}$ using the dictionary $\mathbf{X}$. The sparsity of the obtained reconstruction coefficient vector $\boldsymbol{\alpha}$ is achieved by optimizing the ℓ_0-norm constrained loss:

$$min\|\boldsymbol{\alpha}\|_0 \;\; s.t. \;\; \mathbf{y} = \mathbf{X}\boldsymbol{\alpha}. \tag{3}$$

However, this is an NP-hard problem that is very difficult to solve. To mitigate this issue, the classical SRC [43, 44] method uses the ℓ_1-norm for regularization:

$$min\|\boldsymbol{\alpha}\|_1 \;\; s.t. \;\; \mathbf{y} = \mathbf{X}\boldsymbol{\alpha}. \tag{4}$$

By solving the above optimization problem, the classical SRC algorithm represents a test sample using an over-complete dictionary, and achieves promising results for occlusion- and illumination-invariant face recognition. However, the dimensionality of a face image is usually much higher than the number of atoms/samples in a dictionary. Hence, it is practically very difficult to construct an over-complete dictionary. Moreover, facial appearance variations lead to difficulties in SRC-based face classification, and the solution of the ℓ_1-norm regularized optimization problem is time-consuming.

More recently, the direct use of the ℓ_2-norm constraint has shown superior results in face classification in terms of both accuracy and speed [57, 58]:

$$min\|\boldsymbol{\alpha}\|_2^2 \;\; s.t. \;\; \mathbf{y} = \mathbf{X}\boldsymbol{\alpha}. \tag{5}$$

For this optimization problem, we can use the efficient closed-form solution:

$$\boldsymbol{\alpha} = (\mathbf{X}^T\mathbf{X} + \mu\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}, \tag{6}$$

where μ is the weight of the regularization term and $\mathbf{I}$ is the identity matrix.

Given a test sample and the reconstruction coefficient vector, we can calculate the contribution of the *i*th training sample to the reconstruction of the test sample. We define the ‘contribution’ of the *i*th training sample to the test sample as $\alpha_i \mathbf{x}_i$, and then calculate the contribution of each class in the dictionary to the test sample. For example, for the training samples of the *k*th class $\{\mathbf{x}_{(k-1)M+1}, ..., \mathbf{x}_{kM}\}$, the contribution of the *k*th class to the test sample is:

$$\mathbf{c}_k = \alpha_{(k-1)M+1}\mathbf{x}_{(k-1)M+1} + ... + \alpha_{kM}\mathbf{x}_{kM}. \tag{7}$$

To obtain the classification result of the test sample, the label of the test sample is assigned to the label of the class with the minimum reconstruction error:

$$Label(\mathbf{y}) = \arg\min_k E_k(\mathbf{y}), \quad k \in \{1, ..., K\}, \tag{8}$$

where

$$E_k(\mathbf{y}) = \| \mathbf{y} - \mathbf{c}_k \|_2^2 . \tag{9}$$

The assigned label k of the test sample indicates that the training samples of the kth class in the dictionary best represent the test sample.

3. Proposed method

Most studies mentioned in Section 1 demonstrate that the use of sparse-representation-based methods provides useful information for face classification. However, most of these approaches are not able to deal with the case where samples of the same class (subject) exhibit wide appearance variations, especially when there is a small number of training samples. In this section, we present a new sparse-representation-based face classification method, namely Two-Step LSRC, that dynamically optimizes an augmented dictionary consisting of synthesized local face differences. Fig. 1 shows the schematic diagram of the proposed method. A further discussion and analysis of the proposed method are given in Section 4.

3.1. Dictionary augmentation and optimization

For sparse-representation-based face classification, it is difficult to deal with the difficulties caused by intrinsic and external variations, including illumination, pose, expression, and occlusion. To address this issue, we assume that the image difference

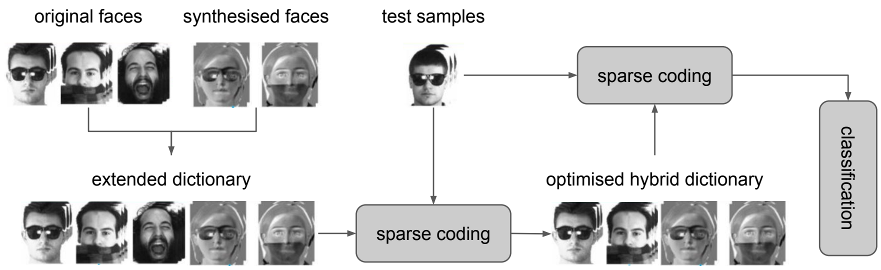


Figure 1: A schematic diagram of the proposed Two-Step LSRC framework.

between two original training samples reflects appearance variations of human faces to some extent; hence, it shares a certain proportion of the contribution to a test sample. This hypothesis has also been used and testified in [7]. Therefore, in this paper, we synthesize a set of auxiliary virtual samples that provide an effective offset for the adverse influence posed by appearance variations. With synthesized virtual samples, the augmented dictionary can better represent a test sample. To further reduce the information redundancy and improve the discriminative information of reconstruction coefficient vectors, we perform online dictionary optimization for robust face classification.

The first stage of the proposed Two-Step LSRC is to synthesize a number of virtual training samples as an auxiliary dictionary for the original one. In this step, the initial difference images are synthesized using the within-class deformation of a subject, *i.e.*, the intensity difference caused by illumination, pose, expression, and occlusion variations. In this study, we used five to ten additional subjects that were not included in the original dictionary to construct the initial auxiliary training set. Some original face images in the AR face dataset and the corresponding synthesized difference images are shown in Fig. 2. The difference image for each face is synthesized by subtracting the natural face from an original face image with appearance variations in expression, illumination, or occlusion, which represents a specific appearance variation type. Then, we obtained the extended dictionary by concatenating the original and auxiliary dictionary matrices into a larger one. With the extended dictionary, both the original and synthesized training samples are jointly used to represent a test sample by linear combination. In the proposed approach, the synthesized samples have equal weights as the

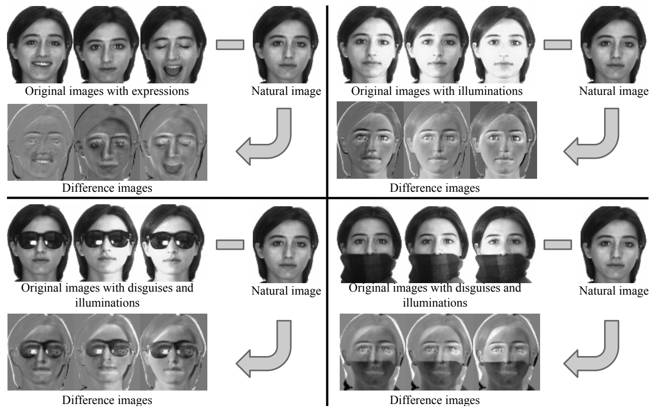


Figure 2: Some examples of the original faces and synthesized faces in the AR dataset. The synthesized face difference images represent local appearance variations by subtracting the natural face from faces with illumination, expression, and changes in disguise.

original ones when we estimate the reconstruction coefficient vector for a test sample.

Despite the outstanding representative capability of the extended dictionary, its use is not without difficulties. The extended dictionary is somewhat redundant; thus, it may over-fit test images and inject uncertainty into decision making. To address this issue, we applied an online dictionary-optimization approach to construct a more compact dictionary by discarding the samples of the classes with smaller contributions to representing a test sample. The final goal of dictionary optimization is to select the most representative training samples from the extended dictionary. Concretely, we first acquire the reconstruction coefficient vector of a test sample using the extended dictionary. Then, we can calculate the reconstruction error of each training class for the test sample, and a smaller reconstruction error indicates a larger contribution to the representation of the test sample. Thus, we can discard a certain proportion of training samples with larger reconstruction errors because they are useless for reconstructing the test sample. A similar idea has also been used in previous studies [32, 33]. However, the difference is that we only discard original training samples, while retaining all of the synthesized ones in the optimized dictionary. We discard synthesized faces only when the number of synthesized faces is larger than those of the original ones. In addition, we do not involve synthesized samples in decision making. The synthesized faces are only used to establish a partnership with the original samples when representing a test sample. Details of the proposed dictionary optimization are presented in the next section.

3.2. Proposed Two-Step LSRC

The Two-Step LSRC algorithm is designed to better represent a test sample $\mathbf{y}$ and obtain its label using sparse-representation-based classification with ℓ_2-norm regularization. The pipeline of our Two-Step LSRC is described as follows.

Step 1: Initialization and dictionary augmentation

Given a dictionary with KM samples $\{\mathbf{x}_1, ..., \mathbf{x}_{KM}\}$, we first generate N synthesized training samples $\{\hat{\mathbf{x}}_1, ..., \hat{\mathbf{x}}_N\}$ as introduced in Section 3.1. The extended dictionary is expressed by the matrix $\mathbf{Z} = [\mathbf{x}_1, ..., \mathbf{x}_{KM}, \hat{\mathbf{x}}_1, ..., \hat{\mathbf{x}}_N]$.

Step 2: Dynamic optimization of the extended dictionary

(1) ℓ_2-norm regularized coefficient vector encoding:

Calculate the coefficient vector by $\boldsymbol{\alpha} = (\mathbf{Z}^T\mathbf{Z} + \mu\mathbf{I})^{-1}\mathbf{Z}^T\mathbf{y}$, where $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_{KM}, \hat{\alpha}_1, ..., \hat{\alpha}_N]^T$ consists of the coefficients for all the column vectors in $\mathbf{Z}$;

(2) Calculate the contribution of each class in the original dictionary to the test sample using Eq. (7), and evaluate the reconstruction error of each class $E_k(\mathbf{y})$ using Eq. (9).

(3) Update the atoms in the extended dictionary $\mathbf{Z}$:

Find the labels $L(P) = \{l_1, l_2, ...\}$ of the classes with relatively larger reconstruction errors, where P is a pre-defined proportion to the total number of classes in the dictionary. Update the extended dictionary by removing the original training samples with the labels in the set $L(P)$. Perform the same procedure on the synthesized samples if $N > KM$.

Step 3: Classification

First, we obtained the reconstruction coefficient vector using the optimized hybrid dictionary, as in Eq. (6). Assume the number of training classes remaining in the optimized dictionary is $R = K - |LP|$ and the original training samples of the rth class are $\{\mathbf{x}_{(r-1)M+1}, ..., \mathbf{x}_{rM}\}$. the contribution of the rth class to the test sample is obtained by:

$$\mathbf{c}_r = \alpha_{(r-1)M+1}\mathbf{x}_{(r-1)M+1} + ... + \alpha_{rM}\mathbf{x}_{rM}, \tag{10}$$

and the reconstruction error of the rth class is:

$$E_r(\mathbf{y}) = \| \mathbf{y} - \mathbf{c}_r \|_2^2 . \tag{11}$$

Output the label of $\mathbf{y}$ using the label of the class with the minimum reconstruction error:

$$Label(\mathbf{y}) = \arg\min_r \{E_r(\mathbf{y})\}. \tag{12}$$

4. Analysis of the proposed Two-Step LSRC algorithm

In this section, we first discuss how the proposed method can reduce the residual in sparse-representation-based signal reconstruction, and then we give a probability explanation with an empirical illustration of the proposed method. Finally, we highlight the contribution and novelty of the proposed algorithm.

4.1. Advantages of using synthesized samples

As discussed in the last section, given a dictionary, the objective of the proposed algorithm is to reconstruct a new sample using a linear combination of the atoms/samples in the dictionary, in which the reconstruction coefficient vector is regularized by ℓ_2-norm for the purposes of efficiency and accuracy. However, owing to the limited volume of training samples in the dictionary and facial appearance variations, it is usually difficult to obtain an over-complete dictionary and to perfectly reconstruct a new face. In such a case, we used the residual $\mathbf{e}$ to measure the reconstruction error, *i.e.*, $\mathbf{e} = \mathbf{y} - \mathbf{X}\boldsymbol{\alpha}$.

In the proposed algorithm, we attempted to reduce the norm of the residual $\mathbf{e}$ by introducing a set of local difference faces. The advantage of the proposed dictionary augmentation method is that the new dictionary is capable of dealing with appearance variations such as illumination, pose, expression, and disguise. More specifically, we rewrote the extended dictionary as $\mathbf{Z} = [\mathbf{X}, \hat{\mathbf{X}}]$, where $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_{KM}]$ is the original dictionary, $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, ..., \hat{\mathbf{x}}_N]$ is the synthesized dictionary consisting of N local difference faces. The linear reconstruction of a test sample $\mathbf{y}$ using the extended dictionary can be expressed as:

$$\mathbf{y} \approx \mathbf{X}\boldsymbol{\alpha} + \hat{\mathbf{X}}\hat{\boldsymbol{\alpha}}. \tag{13}$$

As the reconstruction error $\mathbf{e}$ of the original dictionary can be calculated by $\mathbf{e} = \mathbf{y} - \mathbf{X}\boldsymbol{\alpha}$, the above equation is modified to:

$$\mathbf{e} = \hat{\mathbf{X}}\hat{\boldsymbol{\alpha}} + \hat{\mathbf{e}}, \tag{14}$$

which means that the synthesized local difference faces are used to reconstruct the residual between the test sample and the original dictionary. In addition, we can conclude that the new reconstruction error of the extended dictionary is smaller than that of the original dictionary, *i.e.*, $||\hat{\mathbf{e}}||_2^2 \leq ||\mathbf{e}||_2^2$. However, it is uncertain whether a smaller reconstruction error would lead to more accurate classification results because the auxiliary training samples affect the contributions of both intra-class and external-class samples when representing a test sample. To further improve the accuracy of face classification, we dynamically optimized the extended dictionary in the proposed algorithm using an elimination strategy.

4.2. Probability explanation of the dictionary-optimization approach

As introduced in Section 4.1, a test sample $\mathbf{y}$ is represented by the extended dictionary $\mathbf{Z} = [\mathbf{X}, \hat{\mathbf{X}}]$. In fact, the proposed Two-Step LSRC algorithm assumes that the auxiliary dictionary, which consists of synthesized faces, provides complementary contributions to the representation of $\mathbf{y}$.

Concretely, the linear representation of $\mathbf{y}$ using the original samples can be rewritten as:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\alpha}_1 + ... + \mathbf{X}_k\boldsymbol{\alpha}_k + ... + \mathbf{X}_K\boldsymbol{\alpha}_K + \mathbf{e}, \tag{15}$$

where $\mathbf{X}_k$ is the kth class-specific sub-dictionary consisting of the original samples of the kth class, and $\boldsymbol{\alpha}_k$ is the corresponding class-specific reconstruction coefficient vector. Theoretically, all of the entries of $\boldsymbol{\alpha}_k$ should be zero except for the one belonging to the same class of $\mathbf{y}$. With the original dictionary, the ideal result of the obtained coefficient vector should be:

$$\boldsymbol{\alpha} = [(0, ..., 0), ..., \boldsymbol{\alpha}_k^T, ..., (0, ..., 0)]^T, (\boldsymbol{\alpha}_k^T \neq (0, ..., 0)^T), \tag{16}$$

where the elements are all zero except those in $\boldsymbol{\alpha}_k$. The traditional SRC algorithm considers the classes with coefficients of zero that do not affect the final classification procedure. However, in practice, it is very difficult for sparse-representation-based classification methods to obtain a coefficient vector satisfying Eq. (16), especially when we use the ℓ_2-norm regularization. The noise in a reconstruction coefficient vector introduces uncertainty in decision making, and may result in inaccurate face classification.

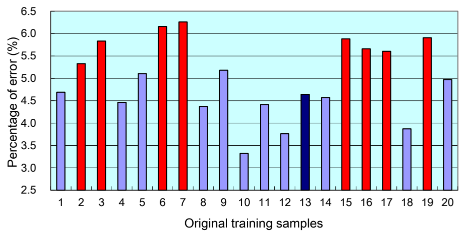
To address the above problem and construct an optimal dictionary, we interpret the sparse representation as a variable selection problem. By design, we ignored the training samples with smaller contributions when representing a test sample; this is done by setting the corresponding coefficients to zero. In other words, we attempted to filter the classes with smaller reconstruction coefficients using an elimination strategy for dictionary optimization. Because the synthesized local difference samples reflect the intrinsic knowledge of a test sample to some extent, the auxiliary dictionary $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, ..., \hat{\mathbf{x}}_N]$ can share a certain proportion of the contribution to the test sample $\mathbf{y}$. Using the extended hybrid dictionary, the coefficient vector is modified as $[\boldsymbol{\alpha}^T, \hat{\boldsymbol{\alpha}}^T]^T$, where $\boldsymbol{\alpha}$ and $\hat{\boldsymbol{\alpha}}$ contain the coefficients corresponding to the original and synthesized samples, respectively. Although the process of solving $\boldsymbol{\alpha}$ satisfying Eq. (16) is still very difficult, using the synthesized local difference faces, the elimination of less-representative training classes in dictionary optimization enlarges the elements in the corresponding class-specific reconstruction vector that have the same label as $\mathbf{y}$. The probability explanation for this assumption is given as follows.

We used T_i to indicate the event for which a test sample $\mathbf{y}$ belongs to the ith class and $P(T_i/\mathbf{y})$ is the probability of an event where $\mathbf{y}$ belongs to the ith class. More specifically, when $\boldsymbol{\alpha}_i \approx (0, ..., 0)^T, (i \neq k, i \in [1, ..., K])$, we can obtain $P(T_i/\mathbf{y}) \approx 0$, which means that the test sample does not belong to the ith class. As denoted by Eq. (9), we used $E_k(\mathbf{y})$ as the reconstruction error between the kth original subdictionary $\mathbf{X}_k$ and $\mathbf{y}$. In fact, we can assume that $P(T_k/\mathbf{y}) \propto \sum_{i=1}^{K} E_i(\mathbf{y})/E_k(\mathbf{y})$, which means that we can obtain higher posterior probability $P(T_k/\mathbf{y})$, while $E_k(\mathbf{y})$ is relatively smaller than the reconstruction errors of other classes. Thus, we assign $P(T_i/\mathbf{y}) = 0$ when $E_i(\mathbf{y})$ belongs to the subsets with larger reconstruction errors. Finally, we used the remaining atoms in the optimized dictionary to better represent $\mathbf{y}$ and perform robust classification. The joint use of the synthesized faces and the elimination strategy reduces the relative reconstruction error of the correct training class; hence, improving the accuracy in sparse-representation-based face classification.

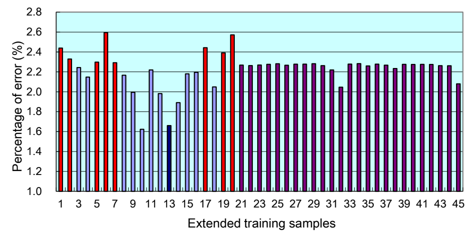
4.3. An empirical illustration of the proposed method

To better illustrate how the proposed method works, we compared the relative reconstruction errors between all the training classes and a test sample in sparse-representation-based classification using a) the original dictionary without elimination, b) the original dictionary with elimination, c) the extended dictionary without elimination, and d) the extended dictionary with elimination, as in Fig. 3. It should be noted that the term 'relative reconstruction error' refers to the percentage of the reconstruction error derived from a specific class relative to the sum of the reconstruction errors over all the training classes and auxiliary training samples. In Fig. 3, the test sample belongs to the 13th class of all the 20 classes (each class has eight training samples), and we used 25 auxiliary local difference faces in the extended dictionary. The relative reconstruction errors derived from the training samples of the correct class, *i.e.*, the 13th class, are highlighted using dark blue bars. The red bars indicate the training classes with higher relative reconstruction errors, which should be eliminated during the dictionary-optimization step. The remaining classes that are different from the label of the test sample are plotted using light blue bars, and the relative reconstruction errors derived from the auxiliary local difference faces are plotted using purple bars (ID 21-45).

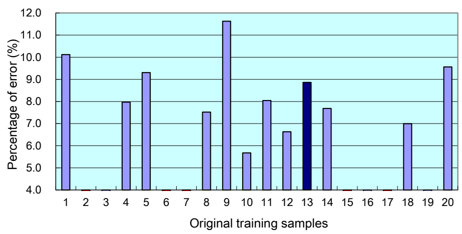
Because the percentage of the reconstruction error derived from a class of training samples is inversely proportional to the probability $P(T_i/\mathbf{y})$, the error derived from the correct class should have the smallest value for the purpose of accurate decision making. In contrast, a larger error indicates that the training samples in the corresponding



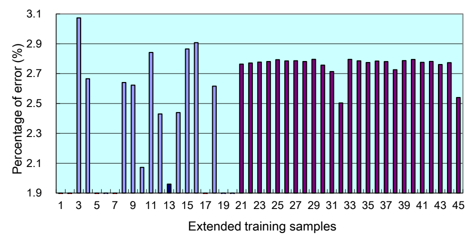
(a) Original dictionary without elimination



(b) Extended dictionary without elimination



(c) Original dictionary with elimination



(d) Extended dictionary with elimination

Figure 3: A comparison of the percentage of the reconstruction error derived from each class for a test sample using: (a) the original dictionary without elimination; (b) the extended dictionary without elimination; (c) the original dictionary with elimination; and (d) the extended dictionary with elimination. To obtain the meanings of the bars with different colors, refer to Section 4.3. The horizontal axes indicate the ID of the training classes (1-20) and synthesized samples (21-45), while the vertical axes indicate the percentage of the reconstruction error derived from a specific class (or sample). We performed the experiment on the AR database, and we selected the test sample using the 15th image of the 13th subject. We selected the training samples using the first eight images of the first 20 subjects. Then, we generated the auxiliary local difference faces by subtracting the first image from other images of the 22nd subject.

class are incapable of representing the test sample, and they should be discarded from the optimized dictionary.

In Fig. 3a, the results are obtained using the original dictionary, which is similar to the classical SRC algorithm [44]. The correct class does not provide the smallest reconstruction error, and leads to an inaccurate classification result. The relative reconstruction errors derived from the 10th, 12th, and 18th classes are all smaller than that of the correct class. With the auxiliary local difference faces, the relative reconstruction errors derived from the correct class (13th) is decreased, which is smaller than that derived from the 12th and 18th classes, as shown in Fig. 3b. However, the 10th class still has the smallest error and the final classification result is incorrect. This method is similar to the ESRC [7] algorithm, which also uses auxiliary samples.

We obtained Fig. 3c using the original dictionary with the elimination strategy, which demonstrates that the simple use of the elimination strategy cannot improve the accuracy for sparse-representation-based classification. This method can be viewed as the Two-Step SRC algorithm. Finally, in Fig. 3d, the proposed Two-Step LSRC algorithm uses the extended dictionary with auxiliary training samples and the elimination strategy, which significantly reduces the relative reconstruction error derived from the correct class and obtains an accurate face-classification result.

In this section, we demonstrate that the single use of synthesized faces or the dictionary-optimization approach is not able to obtain the minimum reconstruction error of the training samples in the correct class for a test sample. In contrast, the proposed Two-Step LSRC, which combines dictionary augmentation and optimization methods, significantly reduces the reconstruction error of the correct class and improves the face-classification accuracy.

4.4. Novelty of the proposed method

As discussed above, the superiority of the proposed Two-Step LSRC algorithm stems from two aspects:

1) A simple but effective data augmentation method is proposed to synthesize local difference faces in our Two-Step LSRC. With the generated local difference faces, we can obtain a more robust sparse representation for a new sample. As mentioned in Section 4.1, the reconstruction error can be reduced using the synthesized local difference faces. The main reason is that these local difference faces reflect some local appearance variations in pose, expression, illumination, and occlusion. For a new test sample, the synthesized local difference faces are used as a complementary dictionary

that collaborates with the original one to represent the test sample. This is very important for successful compress sensing and sparse representation. In addition, according to Section 4.3, the use of synthesized faces reduces the relative reconstruction error derived from the class that has the same label as a test sample, which is beneficial for decision making (Fig. 3b). Although the single use of an extended dictionary does not provide accurate classification in the example shown in Section 4.3, in practical applications, it improves the accuracy of face classification when compared to that of using the original dictionary. This is validated in our experimental results in the next section.

2) To further improve the accuracy of the proposed algorithm in face classification, we apply dictionary optimization to the augmented training samples. The elimination strategy works as a filter mechanism that selects the most representative samples by assessing the contribution of each class to the test sample. The two-step mechanism successively reduces the size of candidate classes by discarding a number of insignificant classes to represent the test sample. With this mechanism, we can select the best class that matches the test sample by reducing the relative reconstruction error between the correct class and a test sample, as discussed in Section 4.3. In contrast, the traditional SRC method makes decisions from all the training classes. Therefore, it is not able to identify all of the external classes that make smaller contributions to the reconstruction of a test sample. In summary, the assessment method in our proposed algorithm is used to evaluate the effectiveness of the samples of each class when representing a test sample, and hence, it obtains more accurate face-classification results. Finally, it should be noted that the use of the ℓ_2-norm constraint in the proposed algorithm is more efficient when compared to the use of the ℓ_1-norm constraint in the classical SRC [44].

5. Experimental Results

In this section, we present our experimental results obtained on the FERET [28] and AR [22] face datasets, which have been widely used to benchmark face classification algorithms. The face images of these two databases were captured with illumination, pose, and expression variations. Moreover, disguised face images were also included in

Figure 4: Example faces of the AR dataset.

Figure 5: Face examples of the FERET dataset.

the AR dataset. In our experiments, we used the closed-form solution of $\boldsymbol{\alpha}$ in Eq. (6), and the regularization term μ was set to 0.01.

For the AR dataset, we selected 3120 images from 120 subjects (26 images per subject). These images were captured over two sessions. The selected subset has also been widely used in previous studies [46, 49, 50]. Each image was down-sampled to 40×50 and converted to a 2000×1 vector. Some examples faces of the AR dataset are shown in Fig. 4.

The FERET dataset was obtained from the FERET program, and it has become a standard benchmarking dataset for evaluating state-of-the-art face-recognition algorithms. The proposed algorithm was evaluated on a subset of FERET. The subset has 1400 images of 200 individuals (seven images per individual). Each face image in FERET was also resized and converted to a 2000×1 vector. Some example faces of the FERET dataset are shown in Fig. 5.

5.1. Face classification with insufficient training samples

This experiment is designed to evaluate the accuracy of the proposed Two-Step L-SRC algorithm when we have insufficient training samples (few training samples per subject). In the experiment, we compared the proposed method with various face-recognition approaches such as the well-known PCA, LDA, SRC, and ESRC methods. The nearest-neighbor classifier with Euclidean distance was used to perform face classification for PCA and LDA.

5.1.1. Results obtained using the AR dataset

We first evaluated the proposed Two-Step LSRC algorithm using the AR face dataset. In this experiment, we selected a subset of the first 100 subjects in the database. Because the first 20 images of each subject contain all of the variation types, we selected two images per individual from these 20 images for training, and the remaining images for testing purposes. More specifically, for the 1st to 10th subjects, the first two images were selected as training samples; for the 11th to 20th subjects, the 3rd and 4th images were used for training, and in the same manner, the last two images were selected as training samples for the 91st to 100th subjects. Thus, we developed a training set of 200 images and a test set of 2400 images. The auxiliary dictionary with local difference faces was constructed as presented in Section 3. In order to evaluate the generalization capacity of the proposed algorithm, we used the images from the 101st to 105th subjects (not included in the training or test subsets) to synthesize 125 virtual faces.

Table 1 shows the recognition rates of PCA [37], PCA+LDA [53], SRC [44], ESRC [7], Two-Step SRC, and Two-Step LSRC using the AR Database. Note that the traditional methods including PCA, PCA+LDA, SRC, and ESRC were performed without dictionary optimization. As shown in Table 1, the proposed Two-Step LSRC method outperforms all of the other algorithms, regardless of the number of training samples that were discarded in the elimination step. Second, when compared with the classical PCA and PCA+LDA algorithms, the sparse-representation-based algorithms obtained higher recognition rates for the AR face dataset. This validates the effectiveness of the use of sparse-representation-based face-classification methods. Third, the simple use of synthesized local difference faces improves the performance in terms of accuracy even without the elimination strategy, and this is shown by comparing ESRC to Two-Step SRC and SRC. Then, by using only the dictionary-optimization method, we cannot improve the accuracy by comparing the recognition rates between Two-

Step SRC and SRC. However, the joint use of the proposed dictionary augmentation and optimization methods achieves superior recognition results compared to the other methods. Finally, when we discard 70% of training classes, the proposed Two-Step LSRC method achieves a 71.9% recognition rate, which is 2.88% higher than that of the state-of-the-art ESRC algorithm.

5.1.2. Results obtained using the FERET dataset

In this experiment, we selected a subset that contains the first 190 subjects in the FERET database. More specifically, for each subject, we selected the first two images for training, and the remaining images were used as test samples. Thus, we created a training set of 380 images and a test set of 950 images. Meanwhile, to generate the auxiliary dictionary with synthesized faces, we selected the 191st to the 200th subjects in order to synthesize 60 virtual samples. In Table 2, we compared the proposed method with PCA [37], PCA+LDA [53], SRC [44], ESRC [7], and the Two-Step SRC.

The performance of different face-classification algorithms in Table 2 is similar to corresponding values in Table 1. First, all of the spare-representation-based face-classification algorithms outperform the classical PCA and PCA+LDA algorithms in terms of accuracy. Second, the single use of auxiliary training samples (ESRC vs SRC) performs better than the single use of the elimination strategy (Two-Step SRC vs SRC). Finally, the proposed Two-Step LSRC algorithm that uses both the dictionary augmentation and optimization methods is superior to the others in terms of face-recognition rate. In particular, our Two-Step LSRC achieves a 69.47% recognition rate, which is 10.73% higher than ESRC when we discard 80% or 90% of training samples. The FERET face dataset is easier than AR for face recognition. This is why the improvement of the proposed algorithm is more significant than the results obtained for the AR dataset.

According to the experimental results obtained for the AR and FERET face datasets, we can conclude that the proposed Two-Step LSRC improves the performance of face classification in terms of accuracy, when compared to the SRC and ESRC methods as well as the classical PCA and PCA+LDA algorithms.

Table 1: Comparison of the face-classification accuracy of PCA, PCA+LDA, SRC, ESRC, Two-Step SRC, and Two-Step LSRC using the AR database, measured as the recognition rate (%).

Method	Elimination proportion						Remarks
	0%	10%	30%	50%	70%	90%	
Two-Step LSRC	69.04	70.08	70.71	71.62	**71.92**	70.62	With synthesized faces
ESRC	69.04	-	-	-	-	-	With synthesized faces
Two-Step SRC	66.25	66.00	65.75	65.54	63.54	57.79	Without synthesized faces
SRC	66.25	-	-	-	-	-	Without synthesized faces
PCA+LDA T(100,90)	56.95	-	-	-	-	-	T is the No. of transform axes
PCA T(200)	41.41	-	-	-	-	-	T is the No. of transform axes

Table 2: A comparison of the face-classification accuracy of PCA, PCA+LDA, SRC, ESRC, Two-Step SRC, and Two-Step LSRC for the FERET database, measured as the recognition rate (%).

Method	Elimination proportion									
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Two-Step LSRC	58.74	61.58	62.42	63.16	64.53	65.37	66.74	67.79	69.47	**69.47**
ESRC	58.74	-	-	-	-	-	-	-	-	-
Two-Step SRC	54.74	55.05	55.58	56.53	56.53	57.47	58.95	59.89	59.37	59.16
SRC	54.74	-	-	-	-	-	-	-	-	-
PCA+LDA T(100,90)	52.52	-	-	-	-	-	-	-	-	-
PCA T(200)	50.05	-	-	-	-	-	-	-	-	-

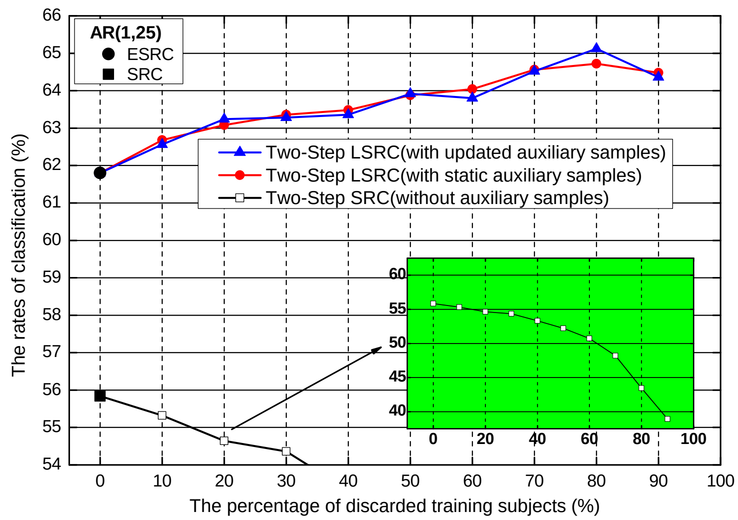
5.2. Face classification using a single training sample

In this section, we design two other experiments and present the face-recognition results obtained for different methods when we have only a single training sample per subject. We used the experiments to verify the effectiveness of the proposed method when dealing with the small-sample-size (SSS) problem.

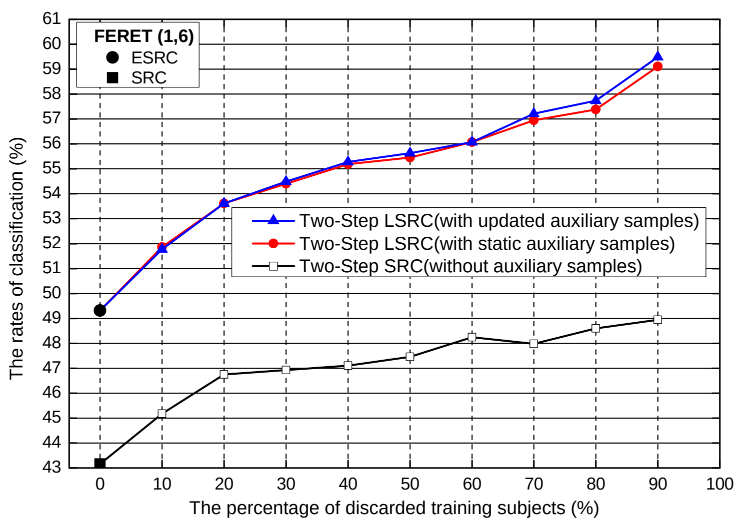
In this section, for the first experiment, we selected the first 100 subjects of the AR database to generate training and test subsets. In practice, we selected one image from the first 13 images of each subject for training. More specifically, for the 1st to 8th subjects, the first image was selected as the training sample; for the 9th to 16th subjects, the second image was selected for training, and in the same manner, the 13th image was selected as the training sample for the 97th to 100th subjects. Meanwhile, we used the remaining 25 images of each subject for testing. Thus, we generated a training set of 100 images and a test set of 2500 images. Moreover, the auxiliary dictionary was constructed using the 101st to 105th subjects. In total, we generated 125 auxiliary virtual samples. For the second experiment, we selected the first 190 subjects from FERET to generate both training and test samples. More specifically, for each subject, the first image was selected for training and the remaining images were selected for testing. Thus, we created a training set of 190 images and a test set of 1140 images. The auxiliary dictionary of FERET was generated in the same way as in Section 5.1.2.

The recognition rates of the ESRC, SRC, Two-Step SRC, and Two-Step LSRC methods are shown in Fig. 6, and are parameterized by the percentage of discarded training classes (0% - 90%). In this experiment, we used two different elimination s-trategies for dictionary optimization. The first one only discards original training samples, which results in a static auxiliary sample set. The second one discards both the original training samples and synthesized samples, which provides an updated auxiliary sample set.

First, as shown in Fig. 6a and Fig. 6b, the use of the proposed dictionary-augmentation method improves the face classification accuracy for both the AR and FERET datasets, and we observe this by comparing the blue and red lines to the black ones. Second, the single use of the elimination strategy is unstable, and cannot always



(a) AR



(b) FERET

Figure 6: Recognition rates of different methods obtained using the (a) AR (1/25) and (b) FERET (1/6) face databases (No. Training / No. Test), when we have only one training sample per subject.

improve the accuracy of a sparse-representation-based face-classification method. Note that there is a sharp drop in the face-recognition rate of Two-Step SRC (without synthesized samples) for the AR database (Fig. 6a). The reason for this phenomenon may lie in the fact that the dictionary contains an insufficient number of training samples, *i.e.*, only a single training image per class in the dictionary. In contrast, the joint use of synthesized faces and the elimination strategy does not result in this issue. Moreover, we find that the accuracy of the proposed Two-Step LSRC algorithm increases as the percentage of the discarded training samples increases. For example, the proposed Two-Step LSRC improves the accuracy from 49.31% to 59.49% for the FERET database. The main reason is that we can realize more accurate decision making when we have fewer candidate classes, as discussed in Section 4.4. Finally, the use of updated auxiliary sample sets does not significantly improve the face-recognition accuracy when compared to the static one. This is why we only discard original training samples in dictionary optimization.

6. Conclusion

In this paper, we proposed a new sparse-representation-based face-classification method, namely Two-Step LSRC. The key innovation of this work is to perform face classification using an optimized dictionary with virtual training samples. The proposed method successfully utilizes local difference face images as an auxiliary dictionary, along with a dictionary-optimization strategy, which enhances the representation capacity of the original dictionary. According to our experimental results obtained using the AR and FERET datasets, our dictionary-argumentation method successfully improves the accuracy of sparse-representation-based classification methods in face classification. Of the many possible research directions associated with this work, the following two are worth noting: (1) to explore the selection of different fuzzy measurement methods in our Two-Step LSRC, and (2) to investigate the use of different fuzzy memberships of the synthesized samples, which allows them to dynamically contribute to the representation of a test sample.

Acknowledgements

We would like to thank the anonymous reviewers for their constructive suggestions. This work was funded by the Natural Science Foundation of China (Grant no. 61672265), the Natural Science Foundation of Jiangsu Province (Grants no. BK20161135, BK20130473), China Postdoctoral Science Foundation (Grant no. 2016M590407), the Fundamental Research Funds for the Central Universities (Grant No. JUSRP115A29, JUSRP51618B), and the Open Project Program of the Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of the Ministry of Education (No. JYB201603). This study was also supported by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions, Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET) and the Open Project Program of the Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University, No. MJUKF201709).

References

[1] D. Beymer, T. Poggio, Face recognition from one example view, in: Computer Vision, 1995. Proceedings., Fifth International Conference on, IEEE, 1995, pp. 500–507.

[2] V. Blanz, T. Vetter, Face recognition based on fitting a 3d morphable model, IEEE Transactions on pattern analysis and machine intelligence 25 (9) (2003) 1063–1074.

[3] E. J. Candes, T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies?, IEEE Transactions on Information Theory 52 (12) (2006) 5406–5425.

[4] B. Chen, H. Shu, G. Coatrieux, G. Chen, X. Sun, J. L. Coatrieux, Color image analysis by quaternion-type moments, Journal of Mathematical Imaging and Vision 51 (1) (2014) 124–144.

[5] J. Chen, J. Yang, L. Luo, J. Qian, W. Xu, Matrix variate distribution-induced sparse representation for robust image classification, IEEE transactions on neural networks and learning systems 26 (10) (2015) 2291–2300.

[6] L.-F. Chen, H.-Y. M. Liao, M.-T. Ko, J.-C. Lin, G.-J. Yu, A new LDA-based face recognition system which can solve the small sample size problem, Pattern recognition 33 (10) (2000) 1713–1726.

[7] W. Deng, J. Hu, J. Guo, Extended SRC: Undersampled face recognition via intraclass variant dictionary, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (9) (2012) 1864–1870.

[8] W. Deng, J. Hu, J. Guo, W. Cai, D. Feng, Robust, accurate and efficient face recognition from a single training image: A uniform pursuit approach, Pattern Recognition 43 (5) (2010) 1748–1762.

[9] W. Dong, X. Li, D. Zhang, G. Shi, Sparsity-based image denoising via dictionary learning and structural clustering, in: IEEE International Conference on Computer Vision and Pattern Recognition, 2011, pp. 457–464.

[10] E. Elhamifar, R. Vidal, Robust classification using structured sparse representation, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE, 2011, pp. 1873–1879.

[11] Z.-H. Feng, G. Hu, J. Kittler, W. Christmas, X.-J. Wu, Cascaded collaborative regression for robust facial landmark detection trained using a mixture of synthetic and real images with dynamic weighting, IEEE Transactions on Image Processing 24 (11) (2015) 3425–3440.

[12] Z.-H. Feng, P. Huber, J. Kittler, W. Christmas, X.-J. Wu, Random cascaded-regression copse for robust facial landmark detection, IEEE Signal Processing Letters 1 (22) (2015) 76–80.

[13] Z.-H. Feng, J. Kittler, W. Christmas, P. Huber, X.-J. Wu, Dynamic attention-controlled cascaded shape regression exploiting training data augmentation and fuzzy-set sample weighting, arXiv preprint arXiv:1611.05396.

[14] B. Gu, V. S. Sheng, A robust regularization path algorithm for v-support vector classification, IEEE Transactions on Neural Networks and Learning Systems PP (99) (2016) 1–8.

[15] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, S. Li, Incremental support vector learning for ordinal regression., IEEE Transactions on Neural Networks and Learning Systems 26 (7) (2014) 1403–1416.

[16] B. Gu, X. Sun, V. S. Sheng, Structural minimax probability machine, IEEE Transactions on Neural Networks and Learning Systems PP (99) (2016) 1–11.

[17] J. Kittler, P. Huber, Z.-H. Feng, G. Hu, W. Christmas, 3D Morphable Face Models and Their Applications, in: International Conference on Articulated Motion and Deformable Objects, Springer, 2016, pp. 185–206.

[18] H. Lee, A. Battle, R. Raina, A. Y. Ng, Efficient sparse coding algorithms, in: Advances in neural information processing systems, 2006, pp. 801–808.

[19] J. Li, X. Li, B. Yang, X. Sun, Segmentation-based image copy-move forgery detection scheme, IEEE Transactions on Information Forensics and Security 10 (3) (2015) 507–518.

[20] Y. Liu, F. Wu, Z. Zhang, Y. Zhuang, S. Yan, Sparse representation using nonnegative curds and whey, in: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, 2010, pp. 3578–3585.

[21] J. Mairal, M. Elad, G. Sapiro, Sparse representation for color image restoration, IEEE Transactions on image processing 17 (1) (2008) 53–69.

[22] A. M. Martinez, The AR face database, CVC Technical Report 24.

[23] I. Naseem, R. Togneri, M. Bennamoun, Linear regression for face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (11) (2010) 2106–2112.

[24] Z. Pan, P. Jin, J. Lei, Y. Zhang, X. Sun, S. Kwong, Fast reference frame selection based on content similarity for low complexity hevc encoder, Journal of Visual Communication and Image Representation 40 (2016) 516–524.

[25] Z. Pan, J. Lei, Y. Zhang, X. Sun, Fast motion estimation based on content property for low-complexity h.265/hevc encoder, IEEE Transactions on Broadcasting (2016) 1–10Doi:10.1109/TBC.2016.2580920.

[26] Z. Pan, Y. Zhang, S. Kwong, Efficient motion and disparity estimation optimization for low complexity multiview video coding, IEEE Transactions on Broadcasting 61 (2) (2015) 166–176.

[27] S. C. Park, M. K. Park, M. G. Kang, Super-resolution image reconstruction: a technical overview, IEEE signal processing magazine 20 (3) (2003) 21–36.

[28] P. J. Phillips, H. Moon, S. A. Rizvi, P. J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (10) (2000) 1090–1104.

[29] G. Rätsch, T. Onoda, K.-R. Müller, Soft margins for adaboost, Machine learning 42 (3) (2001) 287–320.

[30] I. Rish, An empirical study of the naive bayes classifier, in: IJCAI 2001 workshop on empirical methods in artificial intelligence, vol. 3, 2001, pp. 41–46.

[31] Y.-S. Ryu, S.-Y. Oh, Simple hybrid classifier for face recognition with adaptively generated virtual data, Pattern recognition letters 23 (7) (2002) 833–841.

[32] X. Song, Z.-H. Feng, G. Hu, J. Kittler, W. Christmas, X.-J. Wu, Dictionary integration using 3d morphable face models for pose-invariant collaborative-representation-based classification, arXiv preprint arXiv:1611.00284.

[33] X. Song, Z.-H. Feng, G. Hu, X.-J. Wu, Half-face dictionary integration for representation-based classification, IEEE Transactions on Cybernetics 47 (1) (2017) 142–152.

[34] X. Song, Z.-H. Feng, G. Hu, X. Yang, J. Yang, Y. Qi, Progressive sparse representation-based classification using local discrete cosine transform evaluation for image recognition, Journal of Electronic Imaging 24 (5) (2015) 053010–053010.

[35] X. Song, Z.-H. Feng, X. Yang, X. Wu, J. Yang, Towards multi-scale fuzzy sparse discriminant analysis using local third-order tensor model of face images, Neurocomputing 185 (2016) 53–63.

[36] X. Song, Z. Liu, X. Yang, S. Gao, A new sparse representation-based classification algorithm using iterative class elimination, Neural Computing and Applications 24 (7-8) (2014) 1627–1637.

[37] X. Song, J. Yang, X. Wu, X. Yang, An optimal symmetrical null space criterion of Fisher discriminant for feature extraction and recognition, Soft Computing 15 (2) (2011) 281–293.

[38] M. W. Spratling, Image segmentation using a sparse coding model of cortical area v1, IEEE transactions on image processing 22 (4) (2013) 1631–1643.

[39] J. A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural processing letters 9 (3) (1999) 293–300.

[40] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, Y. Ma, Toward a practical face recognition system: Robust alignment and illumination by sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (2) (2012) 372–386.

[41] Y. Wang, X. Wang, C. Gan, C. Wan, Image segmentation by sparse representation, in: Audio, Language and Image Processing (ICALIP), 2012 International Conference on, IEEE, 2012, pp. 365–369.

[42] X. Wen, L. Shao, Y. Xue, W. Fang, A rapid learning algorithm for vehicle classification, Information Sciences 295 (2015) 395–406.

[43] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, S. Yan, Sparse representation for computer vision and pattern recognition, Proceedings of the IEEE 98 (6) (2010) 1031–1044.

[44] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2) (2009) 210–227.

[45] Z. Xia, X. Wang, X. Sun, Q. Liu, N. Xiong, Steganalysis of lsb matching using differences between nonadjacent pixels, Multimedia Tools and Applications 75 (4) (2014) 1–16.

[46] Y. Xu, X. Fang, X. Li, J. Yang, J. You, H. Liu, S. Teng, Data Uncertainty in Face Recognition, IEEE Transactions on Cybernetics 44 (10) (2014) 1950–1961.

[47] Y. Xu, X. Li, J. Yang, Z. Lai, D. Zhang, Integrating conventional and inverse representation for face recognition, IEEE Transactions on Cybernetics 44 (10) (2014) 1738–1746.

[48] Y. Xu, D. Zhang, J. Yang, J.-Y. Yang, A two-phase test sample sparse representation method for use with face recognition, IEEE Transactions on Circuits and Systems for Video Technology 21 (9) (2011) 1255–1262.

[49] Y. Xu, Q. Zhu, Z. Fan, D. Zhang, J. Mi, Z. Lai, Using the idea of the sparse representation to perform coarse-to-fine face recognition, Information Sciences 238 (2013) 138–148.

[50] Y. Xu, X. Zhu, Z. Li, G. Liu, Y. Lu, H. Liu, Using the original and 'symmetrical' face training samples to perform representation based two-step face recognition, Pattern Recognition 46 (4) (2013) 1151–1158.

[51] J. Yang, J. Wright, T. Huang, Y. Ma, Image super-resolution as sparse representation of raw image patches, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.

[52] J. Yang, J. Wright, T. S. Huang, Y. Ma, Image super-resolution via sparse representation, IEEE transactions on image processing 19 (11) (2010) 2861–2873.

[53] J. Yang, J. Yang, Why can LDA be performed in PCA transformed space?, Pattern recognition 36 (2) (2003) 563–566.

[54] M. Yang, D. Zhang, S. Wang, Relaxed collaborative representation for pattern classification, in: IEEE International Conference on Computer Vision and Pattern Recognition, 2012, pp. 2224–2231.

[55] C. Yuan, X. Sun, R. Lv, Fingerprint liveness detection based on multi-scale lpq and pca, China Communications 13 (7) (2016) 60–65.

[56] X.-T. Yuan, X. Liu, S. Yan, Visual classification with multitask joint sparse representation, IEEE Transactions on Image Processing 21 (10) (2012) 4349–4360.

[57] L. Zhang, M. Yang, X. Feng, Sparse representation or collaborative representation: Which helps face recognition?, in: 2011 International Conference on Computer Vision, IEEE, 2011, pp. 471–478.

[58] L. Zhang, M. Yang, X. Feng, Y. Ma, D. Zhang, Collaborative representation based classification for face recognition, arXiv preprint arXiv:1204.2358.

[59] Z. Zhang, Y. Xu, J. Yang, X. Li, D. Zhang, A survey of sparse representation: algorithms and applications, IEEE Access 3 (2015) 490–530.

[60] Y. Zheng, J. Byeungwoo, D. Xu, Q. M. J. Wu, H. Zhang, Image segmentation by generalized hierarchical fuzzy c-means algorithm, Journal of Intelligent and Fuzzy Systems 28 (2) (2015) 4024–4028.

[61] P. Zhu, W. Zuo, L. Zhang, S. C. Shiu, Image Set based Collaborative Representation for Face Recognition, IEEE Transactions on Information forensics and security (2014) 1120–1132.

[62] L. Zhuang, A. Y. Yang, Z. Zhou, S. Shankar Sastry, Y. Ma, Single-sample face recognition with image corruption and misalignment via sparse illumina-

tion transfer, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3546–3553.