Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary

Meng Yang, Lei Zhang [*]

Biometric Research Center, Dept. of Computing, The Hong Kong Polytechnic University, Hong Kong, {csmyang,cslzhang}@comp.polyu.edu.hk

Abstract. By coding the input testing image as a sparse linear combination of the training samples via l_1-norm minimization, sparse representation based classification (SRC) has been recently successfully used for face recognition (FR). Particularly, by introducing an identity occlusion dictionary to sparsely code the occluded portions in face images, SRC can lead to robust FR results against occlusion. However, the large amount of atoms in the occlusion dictionary makes the sparse coding computationally very expensive. In this paper, the image Gabor-features are used for SRC. The use of Gabor kernels makes the occlusion dictionary compressible, and a Gabor occlusion dictionary computing algorithm is then presented. The number of atoms is significantly reduced in the computed Gabor occlusion dictionary, which greatly reduces the computational cost in coding the occluded face images while improving greatly the SRC accuracy. Experiments on representative face databases with variations of lighting, expression, pose and occlusion demonstrated the effectiveness of the proposed Gabor-feature based SRC (GSRC) scheme.

1 Introduction

Automatic face recognition (FR) is one of the most visible and challenging research topics in computer vision, machine learning and biometrics [1], [2], [3]. Although facial images have a high dimensionality, they usually lie on a lower dimensional subspace or sub-manifold. Therefore, subspace learning and manifold learning methods have been dominantly and successfully used in appearance based FR [4], [5], [6], [7], [8], [9], [10], [11]. The classical Eigenface and Fisherface [4], [5], [6] algorithms consider only the global scatter of training samples and they fail to reveal the essential data structures nonlinearly embedded in high dimensional space. The manifold learning methods have been proposed to overcome this limitation [7], [8], and the representative manifold learning methods include locality preserving projection (LPP) [9], local discriminant embedding (LDE) [10], unsupervised discriminant projection (UDP) [11], etc.

The success of manifold learning implies that the high dimensional face images can be sparsely represented or coded by the representative samples on the

---

[*] Corresponding author

manifold. Very recently, an interesting work was reported by Wright *et al.* [12], where the sparse representation (SR) technique is employed for robust FR. In Wright *et al.*'s pioneer work, the training face images are used as the dictionary to code an input testing image as a sparse linear combination of them via l_1-norm minimization. The SR based classification (SRC) of face images is conducted by evaluating which class of training samples could result in the minimum reconstruction error of the input testing image with the sparse coding coefficients. To make the l_1-norm sparse coding computationally feasible, in general the dimensionality of the training and testing face images should be reduced. In other words, a set of features could be extracted from the original image for SRC. In the case of FR without occlusion, Wright *et al.* tested different types of features, including Eigenface, Randomface and Fisherface, for SRC, and they claimed that SRC is insensitive to feature types when the feature dimension is large enough. To solve the problem of FR with occlusion or corruption, an occlusion dictionary was introduced to code the occluded or corrupted components [12]. Since the occluded face image can be viewed as a summation of non-occluded face image and the occlusion error, with the sparsity constrain the non-occluded part is expected to be sparsely coded by the training face dictionary only, while the occlusion part is expected to be coded by the occlusion dictionary only. Consequently, the classification can be performed based on the reconstruction errors using the SR coefficients over the training face dictionary. Such a novel idea has shown to be very effective in overcoming the problem of face occlusion.

Although the SRC based FR scheme proposed in [12] is very creative and effective, there are two issues to be further addressed. First, the features of Eigenface, Randomface and Fisherface tested in [12] are all holistic features. Since in practice the number of training samples is often limited, such holistic features cannot effectively handle the variations of illumination, expression, pose and local deformation. The claim made in [12] that feature extraction is not so important to SRC actually holds only for holistic features. Second, the occlusion matrix proposed in [12] is an orthogonal matrix, such as the identify matrix, Fourier bases or Haar wavelet bases. However, the number of atoms required in the orthogonal occlusion matrix is very high. For example, if the dimensionality of features used in SRC is 3000, then a 3000×3000 occlusion matrix is needed. Such a big occlusion matrix makes the sparse coding process very computationally expensive, and even prohibitive.

In this paper, we propose to solve the above two problems by adopting Gabor local features into SRC. The Gabor filter was first introduced by David Gabor in 1946 [13], and was later shown as models of simple cell receptive fields [14]. The Gabor filters, which could effectively extract the image local directional features at multiple scales, have been successfully and prevalently used in FR [15], [16], leading to state-of-the-art results. Since the Gabor features are extracted in local regions, they are less sensitive to variations of illumination, expression and pose than the holistic features such as Eigenface and Randomface. As in other Gabor-feature based FR works [15], [16], we will see that the Gabor-feature based SRC (GSRC) improves much the FR accuracy over original SRC. More importantly,

the use of Gabor filters in feature extraction makes it possible to obtain a much more compact occlusion dictionary. A Gabor occlusion dictionary computing algorithm is then presented. Compared with the occlusion dictionary used in original SRC, the number of atoms is significantly reduced (often with a ratio $40:1 \sim 50:1$ in our experiments) in the computed Gabor occlusion dictionary. It can not only greatly reduce the computational cost in coding the occluded face images, but also greatly improve the SRC accuracy. Our experiments on benchmark face databases clearly validate the performance of the proposed GSRC method.

The rest of the paper is organized as follows. Section 2 briefly reviews SRC and Gabor filters. Section 3 presents the proposed GSRC algorithm. Section 4 conducts experiments and Section 5 concludes the paper.

2   Related Work

2.1   Sparse representation based classification for face recognition

Denote by $A_i = [\boldsymbol{s}_{i,1}, \boldsymbol{s}_{i,2}, ..., \boldsymbol{s}_{i,n_i}] \in \mathbb{R}^{m \times n_i}$ the set of training samples of the ith object class, where $\boldsymbol{s}_{i,j}, j = 1, 2, \cdots, n_i$, is an m-dimensional vector stretched by the jth sample of the ith class. For a test sample $\boldsymbol{y}_0 \in \mathbb{R}^m$ from this class, intuitively, $\boldsymbol{y}_0$ could be well approximated by the linear combination of the samples within A_i, i.e. $\boldsymbol{y}_0 = \sum_{j=1}^{n_i} \alpha_{i,j} \boldsymbol{s}_{i,j} = A_i \boldsymbol{\alpha}_i$, where $\boldsymbol{\alpha}_i = [\alpha_{i,1}, \alpha_{i,2}, ..., \alpha_{i,n_i}]^T \in \mathbb{R}^{n_i}$ are the coefficients. Suppose we have K object classes, and let $A = [A_1, A_2, \cdots, A_K]$ be the concatenation of the n training samples from all the K classes, where $n = n_1 + n_2 + \cdots + n_K$, then the linear representation of $\boldsymbol{y}_0$ can be written in terms of all training samples as $\boldsymbol{y}_0 = A\boldsymbol{\alpha}$, where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1; \cdots; \boldsymbol{\alpha_i}; \cdots; \boldsymbol{\alpha}_K] = [0, \cdots, 0, \alpha_{i,1}, \alpha_{i,2}, \cdots, \alpha_{i,n_i}, 0, \cdots, 0]^T$ [12].

In the case of occlusion or corruption, we can rewrite the test sample $\boldsymbol{y}$ as

$$\boldsymbol{y} = \boldsymbol{y}_0 + \boldsymbol{e}_0 = A\boldsymbol{\alpha} + \boldsymbol{e}_0 = [A, \ A_e] \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}_e \end{bmatrix} \doteq B\boldsymbol{\omega} \tag{1}$$

where $B = [A, A_e] \in \mathbb{R}^{m \times (n+n_e)}$, and the clean face image $\boldsymbol{y}_0$ and the corruption error $\boldsymbol{e}_0$ have sparse representations over the training sample dictionary A and occlusion dictionary $A_e \in \mathbb{R}^{m \times n_e}$, respectively. In [12], the occlusion dictionary A_e was set as an orthogonal matrix, such as identity matrix, Fourier bases, Haar wavelet bases, etc. The SRC algorithm [12] is summarized in Algorithm 1.

2.2   Gabor filters

The Gabor filters (kernels) with orientation μ and scale ν are defined as [15]:

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{\left(-\|k_{\mu,\nu}\|^2 \|z\|^2 / 2\sigma^2\right)} \left[e^{ik_{\mu,\nu}z} - e^{-\sigma^2/2} \right] \tag{6}$$

where $z = (x, y)$ denotes the pixel, and the wave vector $k_{\mu,\nu}$ is defined as $k_{\mu,\nu} = k_\nu e^{i\phi_\mu}$ with $k_v = k_{\max}/f^v$ and $\phi_\mu = \pi\mu/8$. k_{max} is the maximum frequency, and

---

Algorithm 1 The SRC algorithm in [12]

1: Normalize the columns of A (in the case of non-occlusion) or B (in the case of occlusion) to have unit l_2-norm.

2: Solve the l_1-minimization problem:

$$\hat{\boldsymbol{\alpha}}_1 = \arg \min_{\boldsymbol{\alpha}} \left\{ \|\boldsymbol{y}_0 - A\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \right\} \tag{2}$$

or

$$\hat{\boldsymbol{\omega}}_1 = \arg \min_{\boldsymbol{\omega}} \left\{ \|\boldsymbol{y} - B\boldsymbol{\omega}\|_2^2 + \lambda \|\boldsymbol{\omega}\|_1 \right\} \tag{3}$$

where $\hat{\boldsymbol{\omega}}_1 = [\hat{\boldsymbol{\alpha}}_1; \hat{\boldsymbol{\alpha}}_{e1}]$, and λ is a positive scalar number that balances the reconstructed error and coefficients' sparsity.

3: Compute the residuals:

$$r_i\left(\boldsymbol{y}_0\right) = \|\boldsymbol{y}_0 - A\delta_i\left(\hat{\boldsymbol{\alpha}}_1\right)\|_2, \ \ \text{for } i = 1, \cdots, k. \tag{4}$$

or

$$r_i\left(\boldsymbol{y}\right) = \|\boldsymbol{y} - A_e\hat{\boldsymbol{\alpha}}_{e1} - A\delta_i\left(\hat{\boldsymbol{\alpha}}_1\right)\|_2, \ \ \text{for } i = 1, \cdots, k. \tag{5}$$

where $\delta_i\left(\cdot\right) : \mathbb{R}^n \to \mathbb{R}^n$ is the characteristic function which selects the coefficients associated with the ith class.

4: Output that identity$(\boldsymbol{y}_0) = \arg \min r_i\left(\boldsymbol{y}_0\right)$ or identify$(\boldsymbol{y}) = \arg \min r_i\left(\boldsymbol{y}\right)$.

---

f is the spacing factor between kernels in the frequency domain. In addition, σ determines the ratio of the Gaussian window width to wavelength.

The convolution of an image Img with a Gabor kernel $\psi_{\mu,\nu}$ outputs $G_{\mu,\nu}\left(z\right) = Img\left(z\right) * \psi_{\mu,\nu}\left(z\right)$, where "$*$" denotes the convolution operator. The Gabor filtering coefficient $G_{\mu,\nu}\left(z\right)$ is a complex number, which can be rewritten as $G_{\mu,\nu}\left(z\right) = M_{\mu,\nu}\left(z\right) \cdot \exp\left(i\theta_{\mu,\nu}\left(z\right)\right)$ with $M_{\mu,\nu}\left(z\right)$ being the magnitude and $\theta_{\mu,\nu}\left(z\right)$ being the phase. It is known that magnitude information contains the variation of local energy in the image. In [15], the augmented Gabor feature vector $\boldsymbol{\chi}$ is defined via uniform down-sampling, normalization and concatenation of the Gabor filtering coefficients:

$$\boldsymbol{\chi} = \left(\boldsymbol{a}_{0,0}^{(\rho)^t} \ \boldsymbol{a}_{0,1}^{(\rho)^t} \ \cdots \ \boldsymbol{a}_{4,7}^{(\rho)^t}\right)^t \tag{7}$$

where $\boldsymbol{a}_{\mu,\nu}^{(\rho)}$ is the concatenated column vector from down-sampled magnitude matrix $M_{\mu,\nu}^{(\rho)}$ by a factor of ρ, and t is the transpose operator.

3 Gabor-feature based SRC with Gabor occlusion dictionary

3.1 Gabor-feature based SRC (GSRC)

Images from the same face, taken at (nearly) the same pose but under varying illumination, often lie in a low-dimensional linear subspace known as the

harmonic plane or *illumination cone* [17], [18]. This implies that if there are only variations of illumination, SRC can work very well. However, SRC with the holistic image features is less efficient when there are local deformations of face images, such as certain amount of variations of expressions and pose.

The augmented Gabor face feature vector $\boldsymbol{\chi}$, which is a local feature descriptor, can not only enhance the face feature but also tolerate to image local deformation to some extent. So we propose to use $\boldsymbol{\chi}$ to replace the holistic face features in the SRC framework, and the Gabor-feature based SR without face occlusion is

$$\boldsymbol{\chi}(\mathbf{y}_0) = X(A_1)\boldsymbol{\alpha}_1 + X(A_2)\boldsymbol{\alpha}_2 + \cdots + X(A_K)\boldsymbol{\alpha}_K = X(A)\boldsymbol{\alpha} \qquad (8)$$

where $X(A) = [X(A_1) X(A_2) \cdots X(A_K)]$ and $X(A_i) = [\boldsymbol{\chi}(\boldsymbol{s}_{i,1}), \cdots, \boldsymbol{\chi}(\boldsymbol{s}_{i,n_i})]$. With Eq. (8) and replacing $\boldsymbol{y}_0$ and A in Eq. (2) and Eq. (4) by $\boldsymbol{\chi}(\boldsymbol{y}_0)$ and $X(A)$ respectively, the Gabor-feature based SRC (GSRC) can be achieved.

When the query face image is occluded, similar to original SRC, an occlusion dictionary will be introduced in the GSRC to code the occlusion components, and the SR in Eq. (8) is modified to:

$$\boldsymbol{\chi}(\boldsymbol{y}) = [X(A), \; X(A_e)] \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha}_e \end{bmatrix} \doteq X(B)\boldsymbol{\omega} \qquad (9)$$

where $X(A_e)$ is the Gabor-feature based occlusion dictionary, and $\boldsymbol{\alpha}_e$ is the representation coefficient vector of the input Gabor feature vector $\boldsymbol{\chi}(\boldsymbol{y})$ over $X(A_e)$. So in the case of occlusion, GSRC can be achieved by Algorithm 1 through replacing $\boldsymbol{y}$, B, A and A_e in Eq. (3) and Eq. (5) by $\boldsymbol{\chi}(\boldsymbol{y})$, $X(B)$, $X(A)$ and $X(A_e)$ respectively. Clearly, the remaining key problem is how to process $X(A_e)$ to make the GSRC more efficient.

3.2 Discussions on occlusion dictionary

SRC is successful in solving the problem of face occlusion by introducing an occlusion dictionary A_e to code the occluded face components; however, one fatal drawback of SRC is that the number of atoms in the occlusion dictionary is very big. Specifically, the orthogonal occlusion dictionary, such as the identity matrix, was employed in [12] so that the number of atoms equals to the dimensionality of the image feature vector. For example, if the feature vector has a dimensionality of 3000, then the occlusion dictionary is of size 3000×3000. Such a high dimensional dictionary makes the sparse coding very expensive, and even computationally prohibitive. The empirical complexity of the commonly used l_1-regularized sparse coding methods (such as l_{1-} ls [19], l_{1-} magic [20], PDCO-LSQR [21] and PDCO-CHOL [21]) to solve Eq. (2) is $O(n^\varepsilon)$ with $\varepsilon \approx 2$ [19]. So if the number of atoms (i.e. n) in the occlusion dictionary is too big, the computational cost will be huge.

By using Gabor-feature based SR, the face image dictionary A and the occlusion dictionary A_e in Eq. (1) will be transformed into the Gabor feature

dictionary $X(A)$ and the Gabor-feature based occlusion dictionary $X(A_e)$ in Eq. (9). Fortunately, $X(A_e)$ is compressible, as can be illustrated by Fig. 1.

After the band-pass Gabor filtering of the face images, a uniform down-sampling with a factor ρ is conducted to form the augmented Gabor feature vector $\boldsymbol{\chi}$, as indicated by the red pixels in Fig. 1. The spatial down-sampling is performed for all the Gabor filtering outputs along different orientations and at different scales. Therefore, the number of (spatial) pixels in the augmented Gabor feature vector $\boldsymbol{\chi}$ is $1/\rho$ times that of the original face image; meanwhile, at each position, e.g. P1 or P2 in Fig. 1, it contains a set of directional and scale features extracted by Gabor filtering in the neighborhood (e.g. the circles centered on P1 and P2). Certainly, the directional and scale features at the same spatial location are in general correlated. In addition, there are often some overlaps between the supports of Gabor filters, which makes the Gabor features at neighboring positions also have some redundancies.
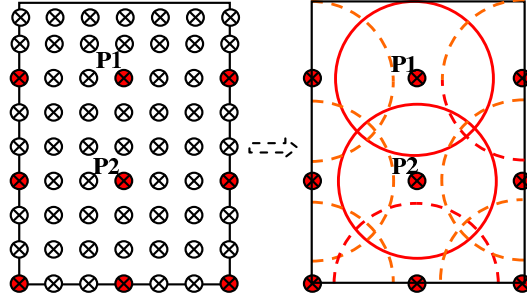

Fig. 1. The uniform down-sampling of Gabor feature extraction after Gabor filtering.
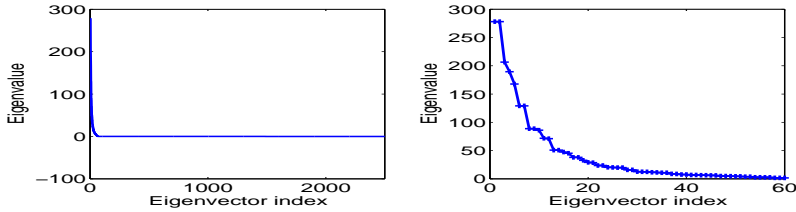

Fig. 2. The eigenvalues (left: all the eigenvalues, right:the first 60 eigenvalues) of Gabor feature-based occlusion matrx.

Considering that "occlusion" is a phenomenon of spatial domain, a spatial down-sampling of the Gabor features with a factor of ρ implies that we can use approximately $1/\rho$ times the occlusion bases to code the Gabor features of the occluded face image. In other words, the Gabor-feature based occlusion dictionary $X(A_e)$ can be compressed because the Gabor features are redundant as we discussed above. To validate this conclusion, we suppose that the image

size is 50×50, and in the original SRC the occlusion dictionary is an identity matrix $A_e = I \in \mathbb{R}^{2500\times2500}$. Then the Gabor-feature based occlusion matrix $X(A_e) \in \mathbb{R}^{2560\times2500}$, where we set $\rho=36$, $\mu = \{0,\cdots,7\}$, $\nu = \{0,\cdots,4\}$. Fig. 2 shows the eigenvalues of $X(A_e)$. Though all the basis vectors of identity matrix I (i.e. A_e) have equal importance, only a few (i.e. 60, with energy proportion of 99.67 %) eigenvectors of $X(A_e)$ have significant eigenvalues, as shown in Fig. 2. This implies that $X(A_e)$ can be much more compactly represented by using only a few atoms generated from $X(A_e)$, often with a compression ratio slightly over ρ:1. For example, in this experiment we have 2500/60=41.7 $\approx \rho$=36. Next we present an algorithm to compute a compact Gabor occlusion dictionary under the framework of SRC.

3.3 Gabor occlusion dictionary computing

Now that $X(A_e)$ is compressible, we propose to compute a compact occlusion dictionary from it with the sparsity constraint required by sparse coding. We call this compact occlusion dictionary the Gabor occlusion dictionary and denote it as Γ. Then we could replace $X(A_e)$ by Γ in the GSRC based FR.

For the convenience of expression, we denote by $Z = X(A_e) = [\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_e}] \in \mathbb{R}^{m_\rho \times n_e}$ the uncompressed Gabor-feature based occlusion matrix, with each column $\boldsymbol{z}_i$ being the augmented Gabor-feature vector generated from each atom of the original occlusion dictionary A_e. The compact occlusion dictionary to be computed is denoted by $\Gamma = [\boldsymbol{d}_1, \boldsymbol{d}_2, ..., \boldsymbol{d}_p] \in \mathbb{R}^{m_\rho \times p}$, where p can be set as slightly less than n_e/ρ in practice. It is required that each occlusion basis $\boldsymbol{d}_j, j = 1, 2, \cdots, p$, is a unit column vector, i.e. $\boldsymbol{d}_j^T \boldsymbol{d}_j = 1$. Since we want to replace Z by Γ, it is expected that the original dictionary Z can be well represented by Γ, while the representation being as sparse as possible. With such consideration, our objective function in determining Γ is defined as:

$$J_{\Gamma,\Lambda} = \arg\min_{\Gamma,\Lambda} \left\{ \|Z - \Gamma\Lambda\|_F^2 + \zeta \|\Lambda\|_1 \right\} \quad \text{s.t.} \quad \boldsymbol{d}_j^T \boldsymbol{d}_j = 1, \forall j \qquad (10)$$

where Λ is the representation matrix of Z over dictionary Γ, and ζ is a positive scalar that balances the F-norm term and the l_1-norm term.

Eq. (10) is a joint optimization problem of the occlusion dictionary Γ and the representation matrix Λ. Like in many multi-variable optimization problems, we solve Eq. (10) by optimizing Γ and Λ alternatively. The optimization procedures are described in the following Algorithm 2.

It is straightforward that the proposed Gabor occlusion dictionary computing algorithm converges because in each iteration $J_{\Gamma,\Lambda}$ will decrease, as illustrated in Fig. 3. Consequently, in GSRC we use Γ to replace the $X(A_e)$ in Eq. (9). Finally, the sparse coding problem in GSRC with face occlusion is

$$\boldsymbol{y}_\Gamma = B_\Gamma \boldsymbol{\omega}_\Gamma, \quad where \ \boldsymbol{y}_\Gamma = \boldsymbol{\chi}(\boldsymbol{y}), \ B_\Gamma = [\boldsymbol{X}(A), \ \Gamma], \ \boldsymbol{\omega}_\Gamma = [\boldsymbol{\alpha}; \boldsymbol{\alpha}_\Gamma] \quad (18)$$

Since the number of atoms in Γ is significantly reduced, the number of variables to be solved in $\boldsymbol{\omega}_\Gamma$ is much decreased, and thus the computational cost in solving Eq. (18) is greatly reduced compared with the original SRC.

Algorithm 2 Algorithm of Gabor occlusion dictionary computing

---

1: Initialize Γ.

   We initialize each column of Γ (i.e. each occlusion basis) as a random vector with unit l_2-norm.

2: Fix Γ and solve Λ.

   By fixing Γ, the objective function in Eq. (10) is reduced to

$$J_\Lambda = \arg\min_\Lambda \left\{ \|Z - \Gamma\Lambda\|_F^2 + \zeta \|\Lambda\|_1 \right\} \tag{11}$$

   The minimization of Eq. (11) can be achieved by some standard convex optimization technique. In this paper, we use the algorithm in [19].

3: Fix Λ and update Γ.

   Now the objective function is reduced to

$$J_\Gamma = \arg\min_\Gamma \left\{ \|Z - \Gamma\Lambda\|_F^2 \right\} \quad \text{s.t.} \quad \boldsymbol{d}_j^T \boldsymbol{d}_j = 1, \forall j \tag{12}$$

   We can write matrix Λ as $\Lambda = [\boldsymbol{\beta}_1; \boldsymbol{\beta}_2; \cdots ; \boldsymbol{\beta}_p]$, where $\boldsymbol{\beta}_j, j = 1, 2, \cdots, p$, is the row vector of Λ. We update $\boldsymbol{d}_j$ one by one. When updating $\boldsymbol{d}_j$, all the other columns of Γ, i.e. $\boldsymbol{d}_l, l \neq j$, are fixed. Then J_Γ in Eq. (12) is converted into

$$J_{\boldsymbol{d}_j} = \arg\min_{\boldsymbol{d}_j} \left\| Z - \sum_{l \neq j} \boldsymbol{d}_l \boldsymbol{\beta}_l - \boldsymbol{d}_j \boldsymbol{\beta}_j \right\|_F^2 \quad \text{s.t.} \quad \boldsymbol{d}_j^T \boldsymbol{d}_j = 1 \tag{13}$$

   Let $Y = Z - \sum_{l \neq j} \boldsymbol{d}_l \boldsymbol{\beta}_l$, Eq. (13) can be written as

$$J_{\boldsymbol{d}_j} = \arg\min_{\boldsymbol{d}_j} \left\| Y - \boldsymbol{d}_j \boldsymbol{\beta}_j \right\|_F^2 \quad \text{s.t.} \quad \boldsymbol{d}_j^T \boldsymbol{d}_j = 1 \tag{14}$$

   Using Langrage multiplier, $J_{\boldsymbol{d}_j}$ is equivalent to

$$J_{\boldsymbol{d}_j,\gamma} = \arg\min_{\boldsymbol{d}_j} tr \left(-Y\boldsymbol{\beta}_j^T \boldsymbol{d}_j^T - \boldsymbol{d}_j \cdot \boldsymbol{\beta}_j Y^T + \boldsymbol{d}_j \cdot (\boldsymbol{\beta}_j \boldsymbol{\beta}_j^T - \gamma)\boldsymbol{d}_j^T + \gamma \right) \tag{15}$$

   where γ is a scalar variable. Differentiating $J_{\boldsymbol{d}_j,\gamma}$ with respect to $\boldsymbol{d}_j$, and let it be 0, we have

$$\boldsymbol{d}_j = Y\boldsymbol{\beta}_j^T \left(\boldsymbol{\beta}_j \boldsymbol{\beta}_j^T - \gamma \right)^{-1} \tag{16}$$

   Since $\left(\boldsymbol{\beta}_j \boldsymbol{\beta}_j^T - \gamma \right)$ is a scalar and γ is a variable, the solution of Eq. (16) under constrain $\boldsymbol{d}_j^T \boldsymbol{d}_j = 1$ is

$$\boldsymbol{d}_j = Y\boldsymbol{\beta}_j^T \left/ \left\| Y\boldsymbol{\beta}_j^T \right\|_2 \right. \tag{17}$$

   Using the above procedures, we can update all the vectors $\boldsymbol{d}_j$, and hence the whole set Γ is updated.

4: Go back to step 2 until the values of $J_{\Gamma,\Lambda}$ in adjacent iterations are close enough, or the maximum number of iterations is reached. Finally, output Γ.
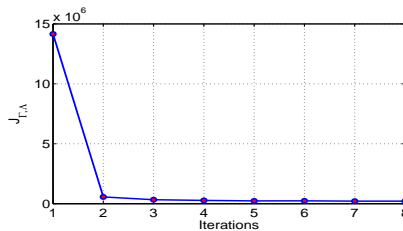

Fig. 3. Illustration of the convergence of Algorithm 2. A Gabor occlusion dictionary with 100 atoms is computed from the original Gabor-feature based occlusion matrix with 4980 columns. The compression ratio is nearly 50:1.

4 Experimental Results

In this section, we perform experiments on benchmark face databases to demonstrate the improvement of GSRC over SRC. To evaluate more comprehensively the performance of GSRC, in section 4.1 we first test FR without occlusion, and then in section 4.2 we demonstrate the robustness and efficiency of GSRC in FR with block occlusion. Finally in section 4.3 we test FR against disguise occlusion. In our implementation of Gabor filters, the parameters are set as $K_{max} = \pi/2$, $f = \sqrt{2}$, $\sigma = \pi$, $\mu = \{0, \cdots, 7\}$, $\nu = \{0, \cdots, 4\}$ by our experimental experiences and fixed for all the experiments below. Here we should also note that the regularization parameters in sparse coding are also tuned by experience (Actually, how to adaptively set the regularizatin parameters is still an open problem). In addition, all the face images are cropped and aligned by using the location of eyes, which is provided by the face databases. The code of our method is available at `http://www4.comp.polyu.edu.hk/~cslzhang/code.htm`.

4.1 Face recognition without occlusion

We evaluated the performance of the proposed algorithm on three representative facial image databases: Extended Yale B [22], [18], AR [23] and FERET [24]. In both the original SRC and the proposed GSRC, we used PCA to reduce the feature dimension. The dictionary size is set according to the image variability and the size of database. Some discussions on the dictionary size with respect to image variability are given using FERET database.

  1) Extended Yale B Database: As the experiment on Extended Yale B database [22], [18] in [12], for each subject, we randomly selected half of the images for training (i.e. 32 images per subject), and used the other half for testing. The images are normalized to 192×168, and the dimension of the augmented Gabor feature vector of each image is 19760. PCA is then applied to reduce their dimensionality for classification in SRC and GSRC. In our experiments, we set λ=0.001 (refer to Eq. (2)) in GSRC. The results of SRC are from the original paper [12]. Fig. 4(a) shows the recognition rates of GSRC versus feature dimension in comparison with those of SRC. It can be seen that GSRC is much better

than SRC in all the dimensions. On this database, the maximal recognition rate of GSRC is 99.17%, while that of SRC is 96.77%.
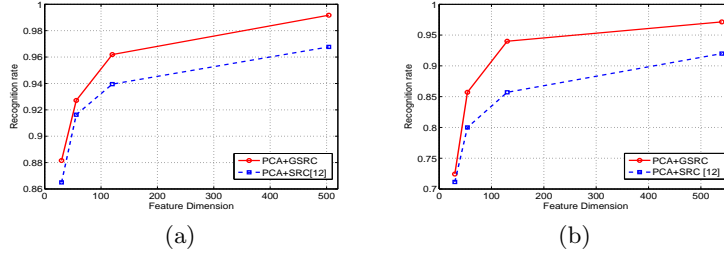

(a)   (b)

Fig. 4. Recognition rates by SRC and GSRC versus feature dimension on (a) Extended Yale B and (b) AR database.

2) AR database: As [12], we chose a subset (only with illumination changes and expressions) of AR dataset [23] consisting of 50 male subjects and 50 female subjects. For each subject, the seven images from Session 1 were used for training, with other seven images from Session 2 for testing. The size of original face image is 165×120, and the Gabor-feature vector is of dimension 12000. We set λ=0.001 in GSRC. The results of SRC are from the original paper [12]. The comparison of GSRC and SRC is shown in Fig. 4(b). Again we can see that GSRC performs much better than SRC under all the dimensions. On this database, the maximal recognition rate of GSRC and SRC are 97.14% and 91.19%, respectively.

The improvement brought by GSRC on AR database is bigger than that on Extended Yale B database. This is because in Extended Yale B, mostly there are only illumination variations between training images and testing images, and dictionary size (i.e. 32 atoms per subject) is big. Thus the original SRC works very well on it. However, the training and testing samples of the AR database have much more variations of expression, time and illumination, and dictionary size (i.e. 7 atoms per subject) is much smaller. Therefore, the local feature based GSRC is much more robust than global feature based SRC in this case.

3) FERET pose database: Here we used the pose subset of the FERET database [24], which includes 1400 images from 198 subjects (about 7 each). This subset is composed of the images marked with 'ba', 'bd', 'be', 'bf', 'bg', 'bj', and 'bk'. In our experiment, each image has the size of 80×80. Some sample images of one person are shown in the Fig. 5(a).

Five tests with different pose angles were performed. In test 1 (pose angle is zero degree), images marked with 'ba' and 'bj' were used as training set, and images marked with 'bk' were used as testing set. In all the other four tests, we used images marked with 'ba', 'bj' and 'bk' as gallery, and used the images with 'bg', 'bf', 'be' and 'bd' as probes. Fig. 5(b) compares GSRC (λ=0.005 for best results) with SRC (λ=0.05 for best results) for different poses. The feature dimension in both methods is 350. Obviously, we can see that GSRC has much higher recognition rates than SRC. Especially, when the pose variation is moderate (0º and ±15º), GSRC's recognition rates are 98.5%, 89.5% and 96%,

respectively, about 20% higher than those of the SRC algorithm (83.5%, 57.5% and 70.5%, respectively). The results also show that good performance can be achieved with a small dictionary size when image variability is small (i.e. test 1). Meanwhile, with the same dictionary size, the performance drops as image variability increases (i.e. test $2 \sim 5$). It is undeniable that GSRC's performance also degrades much as pose variation becomes large (e.g. $\pm 25^{\circ}$). Nevertheless, GSRC can much improve the robustness to moderate pose variation.
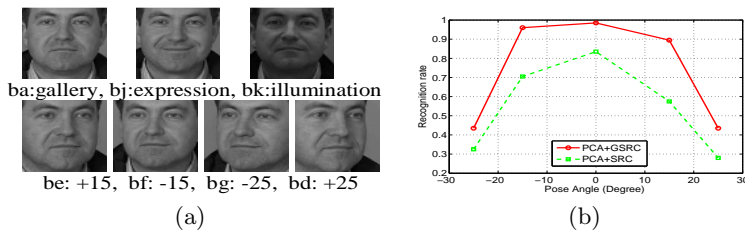

ba:gallery, bj:expression, bk:illumination

be: +15, bf: -15, bg: -25, bd: +25

(a)   (b)

Fig. 5. Samples and results on the FERET pose database. (a). Samples of one subject. (b). Recognition rates of SRC and GSRC versus pose variation.

4.2   Recognition against block occlusion

In this sub-section, we test the robustness of GSRC to the block occlusion using a subset of Extended Yale B face database. We chose Subsets 1 and 2 (717 images, normal-to-moderate lighting conditions) for training, and Subset 3 (453 images, more extreme lighting conditions) for testing. In accordance to the experiments in [12], the images were resized to 96×84, and the occlusion dictionary A_e in SRC is set to an identity matrix.
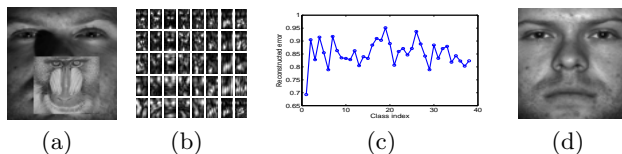

(a)   (b)   (c)   (d)

Fig. 6. An example of face recognition with block occlusion. (a). A 30% occluded test face image y from Extended Yale B. (b). Uniformly down-sampled Gabor features $\chi(y)$ of the test image. (c). Estimated residuals $r_i(y), i = 1, 2, \cdots, 38$. (d). One sample of the class to which the test image is classified.

With the above settings, in SRC the size of matrix B in Eq. (1) is 8064×8781. In the proposed GSRC, the dimension of augmented Gabor-feature vector is 8960 ($\rho \approx 40$). The Gabor occlusion dictionary Γ is then computed using Algorithm 2. In the experiment, we compress the number of atoms in Γ to 200 (i.e. p=200, with compression ratio about 40:1), and hence the size of dictionary B_Γ in Eq. (18) is 8960×917. Compared with the original SRC, the computational cost is reduced from about $O(\eta^2)$ with η=8781 to about $O(\kappa^2)$ with κ=917. Here the time consumption of Gabor feature extraction (about 0.26 second) could

be negligible, compared with that of l_1-norm minimization, which is about 90 seconds as reported in [12].

As in [12], we simulated various levels of contiguous occlusion, form 0% to 50%, by replacing a randomly located square block in each test image with an irrelevant image, whose size is determined by the occlusion percentage. The location of occlusion was randomly chosen for each test image and is unknown to the computer. We tested the performance of GSRC with λ=0.0005, and Fig. 6 illustrates the classification process by using an example. Fig. 6(a) shows a test image with 30% randomly located occlusion; Fig. 6(b) shows the argumented Gabor features of the test image. The residuals of GSRC are plotted in Fig. 6(c), and a template image of the identified subject is shown in Fig. 6(d). The detailed recognition rates of GSRC and SRC are listed in the Table 1, where the results of SRC are from the original paper [12]. We see that GSRC can correctly classify all the test images when the occlusion percentage is less than or equal to 30%. When the occlusion percentage becomes larger, the advantage of GSRC over SRC is getting higher. Especially, GSRC can still have a recognition rate of 87.4% when half of image is occluded, while SRC only achieves a rate of 65.3%.

Table 1. The recognition rates of GSRC and SRC under different levels of block occlusion.

Occlusion percentage	0%	10%	20%	30%	40%	50%
Recognition rate of GSRC	**1**	**1**	**1**	**1**	**0.965**	**0.874**
Recognition rate of SRC	1	1	0.998	0.985	0.903	0.653

Table 2. Recognition rates of GSRC and SRC on the AR database with disguise occlusion ('-p': partitioned, '-sg': sunglasses, and '-sc': scarves).

Algorithms	GSRC	SRC	GSRC-p	SRC-p
Recognition rate-sg	**93.0%**	87%	**100%**	97.5%
Recognition rate-sc	**79%**	59.5%	**99%**	93.5%

4.3   Recognition against disguise

A subset from the AR database consists of 1399 images from 100 subjects (14 samples each class except for a corrupted image w-027-14.bmp), 50 male and 50 female. 799 images (about 8 samples per subject) of non-occluded frontal views with various facial expressions were used for training, while the others for testing. The images are resized to 83×60. So in the original SRC, the size of matrix B in Eq. (1) is 4980×5779. In the proposed GSRC, the dimension of Gabor-feature vectors is 5200 ($\rho \approx 38$), and 100 atoms (with compression ratio about 50:1) are computed to form the Gabor occlusion dictionary by Algorithm 2. Thus the size of dictionary B_Γ in Eq. (18) is 5200×899, and the computational cost is roughly reduced from about $O(\eta^2)$ with η=5779 to about $O(\kappa^2)$ with κ=899, where Gabor feature extraction consumes very little time (about 0.19 second).

We consider two separate test sets of 200 images (1 sample each session and each subject, with neutral expression). The first one contains images of the subjects wearing sunglasses, which occlude roughly 20% of the image. The second one is composed of images of the subjects wearing a scarf, which occlude roughly 40% of the images. The results by GSRC (λ=0.0005) and SRC are listed in Table 2 (where the results of SRC are from the original paper [12]). We see that on faces occluded by sunglasses, GSRC achieves a recognition rate of 93.0%, over 5% higher than that of SRC, while for occlusion by scarves, the proposed GSRC achieves a recognition rate 79%, about 20% higher than that of SRC.

In [12], the authors partitioned image into blocks for face classification by assuming the occlusion is connected. Such an SRC scheme is denoted by SRC-p. Here, after partitioning the image into several blocks, we calculate the Gabor features of each block and then use GSRC to classify each block image. The final classification result is obtained by voting. We denote the GSRC with partitioning as GSRC-p. In experiments, we partitioned the images into eight (4×2) blocks of size 20×30. The Gabor-feature vector of each block is of dimension 800, and the number of atoms in the computed Gabor occlusion dictionary Γ is set to 20. Thus the dictionary B in SRC is of size 600×1379, while the dictionary B_Γ in GSRC is of size 800×819. The recognition rates of SRC-p and GSRC-p are also listed in Table 2. We see that with partitioning, GSRC can lead to recognition rates of 100% on sunglasses and 99% on scarves, also better than SRC.

5 Conclusion

In this paper, we proposed a Gabor-feature based SRC (GSRC) scheme, which uses the image local Gabor features for SRC, and proposed an associated Gabor occlusion dictionary computing algorithm to handle the occluded face images. Apart from the improved face recognition rate, one important advantage of GSRC is its compact occlusion dictionary, which has much less atoms than that of the original SRC scheme. This greatly reduces the computational cost of sparse coding. We evaluated the proposed method on different conditions, including variations of illumination, expression and pose, as well as block occlusion and disguise. The experimental results clearly demonstrated that the proposed GSRC has much better performance than SRC, leading to much higher recognition rates while spending much less computational cost. This makes it much more practicable to use than SRC in real world face recognition.

Acknowledgements

References

1. Zhao, W.Y., Chellppa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Survey **35** (2003) 399–459

2. Su, Y., g. Shan, S., Chen, X.L., Gao, W.: Hierarchical ensemble of global and local classifiers for face recognition. IEEE IP **18** (2009) 1885–1896
3. Zhang, W.C., g. Shan, S., Gao, W., Chen, X.L.: Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. ICCV (2005) 786–791
4. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cognitive Neuroscience **3** (1991) 71–86
5. Belhumeur, P.N., Hespanha, J.P., Kriengman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE PAMI **19** (1997) 711–720
6. Yang, J., Yang, J.Y.: Why can LDA be performed in PCA transformed space? Pattern Recognition **36** (2003) 563–566
7. Tenenbaum, J.B., deSilva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290** (2000) 2319–2323
8. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290** (2000) 2323–2325
9. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.J.: Face recognition using laplacianfaces. IEEE PAMI **27** (2005) 328–340
10. Chen, H.T., Chang, H.W., Liu, T.L.: Local discriminant embedding and its variants. CVPR (2005) 846–853
11. Yang, J., Zhang, D., Yang, J.Y., Niu, B.: Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics. IEEE PAMI **29** (2007) 650–664
12. Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y.: Robust face recognition via sparse representation. IEEE PAMI **31** (2009) 210–227
13. Gabor, D.: Theory of communication. J. Inst. Elect. Eng **93** (1946) 429–457
14. Jones, J.P., Palmer, L.A.: An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. Journal of Neurophysiology **58** (1987) 1233–1258
15. Liu, C., Wechsler, H.: Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE IP **11** (2002) 467–476
16. Shen, L., Bai, L.: A review on gabor wavelets for face recognition. Pattern Analysis and Application **9** (2006) 273–292
17. Basri, R., Jacobs, D.: Lambertian reflectance and linear subspaces. IEEE PAMI **25** (2003) 218–233
18. Georghiades, A.S., Belhumeur, P., Kriegman, D.: From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE PAMI **23** (2001) 643–660
19. Kim, S.J., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D.: A method for large-scale l_1-regularized least squares. IEEE Journal on Selected Topics in Signal Processing **1** (2007) 606–617
20. Cands, E., Romberg, J.: l_1-magic: A collection of matlab routines for solving the convex optimization programs central to compressive sampling (2006) `www.acm.caltech.edu/l1magic/`.
21. Saunders, M.: PDCO: Primal-dual interior method for convex objectives (2002) `http://www.stanford.edu/group/SOL/software/pdco.html`.
22. Lee, K., Ho, J., Kriegman, D.: Acquring linear subspaces for face recognition under variable lighting. IEEE PAMI **27** (2005) 684–698
23. Martinez, A., benavente, R.: The AR face database (1998)
24. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.: The FERET evaluation methodology for face recognition algorithms. IEEE PAMI **22** (2000) 1090–1104