

Research Article

Patch Based Collaborative Representation with Gabor Feature and Measurement Matrix for Face Recognition

Zhengyuan Xu,¹ Yu Liu,² Mingquan Ye,³ Lei Huang,¹ Hao Yu,⁴ and Xun Chen ⁵

¹Department of Medical Engineering, Wannan Medical College, Wuhu 241002, China

²Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China

³School of Medical Information and Research Center of Health Big Data Mining and Applications, Wannan Medical College, Wuhu 241002, China

⁴Anhui Province Key Laboratory of Active Biological Macro-Molecules Research, Central Laboratory of Wannan Medical College, Wuhu 241002, China

⁵Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230027, China

Correspondence should be addressed to Xun Chen; xunchen@ece.ubc.ca

Received 13 December 2017; Accepted 29 March 2018; Published 10 May 2018

Academic Editor: Nibaldo Rodríguez

Copyright © 2018 Zhengyuan Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, sparse representation based classification (SRC) has emerged as a popular technique in face recognition. Traditional SRC focuses on the role of the l_1 -norm but ignores the impact of collaborative representation (CR), which employs all the training examples over all the classes to represent a test sample. Due to issues like expression, illumination, pose, and small sample size, face recognition still remains as a challenging problem. In this paper, we proposed a patch based collaborative representation method for face recognition via Gabor feature and measurement matrix. Using patch based collaborative representation, this method can solve the problem of the lack of accuracy for the linear representation of the small sample size. Compared with holistic features, the multiscale and multidirection Gabor feature shows more robustness. The usage of measurement matrix can reduce large data volume caused by Gabor feature. The experimental results on several popular face databases including Extended Yale B, CMU_PIE, and LFW indicated that the proposed method is more competitive in robustness and accuracy than conventional SR and CR based methods.

1. Introduction

Face recognition (FR) is one of the most classical and challenging problems in pattern classification, computer vision, and machine learning [1]. Although face recognition technology has made a series of achievements, it still confronts many challenges caused by the variations of illumination, pose, facial expression, and noise in real-world [2, 3]. In real applications, the small sample size problem of FR is also a more difficult issue due to the limitations in availability of training samples.

In terms of classification schemes, several widespread pattern classification methods are used in FR. Generally, there are two types of pattern classification methods [4, 5]: parametric methods and nonparametric methods. Parametric methods such as support vector machine (SVM) [6, 7]

center on how to learn the parameters of a hypothesis classification model from the training samples and then use them to identify the class labels of test samples. In contrast, the nonparametric methods, such as nearest neighbor (NN) [8] and nearest subspace (NS) [9], use the training samples directly to identify the class labels of test samples. Recent works have revealed an advantage by the nonparametric methods over the parametric methods [4, 10, 11]. The distance based classifiers are widely used in nonparametric methods for FR [11], such as the nearest subspace classifier (NSC) [12]. A key issue in distance based nonparametric classifiers is how to represent the test sample [4]. Recently, Wright et al. pioneered by using the SRC for robust FR [13]. First the test sample was sparsely coded by the training samples; then the class labels of test samples were identified by choosing which class yields the smallest coding error. Although sparse

representation related methods [13, 14] achieved a great success in FR, those methods focus on the role of the l_1 -norm while ignoring the role of CR [15], which uses all classes of training samples to represent the target test sample. In this study [16], Zhu et al. argued that both SRC and collaborative representation based classifier (CRC) suffer serious performance degradation when the training sample size is very small, because the test sample cannot be well represented. In order to solve the small sample size problem, they proposed to conduct CRC on the patch and named it the patch based CRC (PCRC).

The PCRC and some related works [17, 18] have demonstrated their effects on small sample size problem of FR; however, some key issues remain to be further optimized. On one hand, all the PCRC related works used the original face feature, but the original feature cannot effectively handle the variations of illumination, pose, facial expression, and noise [19]. On the other hand, the data redundancy existing in these methods leads to poor performance in classification accuracy and computational cost. The face feature problem has been noticed by some recent works in which an efficient and effective image representation has been proposed by using local and holistic features. The Eigenface [20, 21], Randomface [22], and Fisherface [21] are all classical holistic features [23], but some other works argued that those holistic features can be easily affected by variables such as illumination, pose, facial expression, and noise. Therefore, they introduced some local features such as LBP [24] and Gabor filter [25, 26]. Gabor filter has been successfully and widely used in FR [26, 27]. Gabor feature could effectively extract the face local features at multiple scales and multiple directions; however, this may lead to a sharp rise in data volumes [28]. The key to solving the large data volume problem is dimensionality reduction. Numerous dimensionality reduction methods have been put forward to find projections that better separate the classes in low-dimensional spaces, among which linear subspace analysis method has received more and more attention owing to its good properties, including principal component analysis (PCA) [28], linear discriminant analysis (LDA) [29], and independent component correlation algorithm (ICA) [30]. A lot of works showed that PCA has the optimal performance in FR [31–33].

In this paper, we first attempted to alleviate the influence of unreliability environment on small sample size in FR; therefore, we proposed to use Gabor feature and applied it to PCRC. Then to improve the computational efficiency of GPCRC, we proposed to use PCA for dimension reduction and then use the measurement matrix, including Random Gaussian matrices [34], Toeplitz and cyclic matrices [35], and Deterministic sparse Toeplitz matrices ($\Delta = 4, 8, 16$) [36], to reduce the dimension of the transformed signal and used the low-dimensional data to accurately represent the face. The experimental results showed that the GPCRC and its improved methods are effective.

Section 2 briefly reviewed SRC and CRC. Section 3 described the proposed GPCRC and its improved methods. Section 4 illustrated the experiments and the results. And Section 5 concluded the paper.

2. Related Works

2.1. Sparse Representation Based Classification. Recently, SRC was first reported by Wright et al. [13] for robust FR. In SRC, let $\mathbf{x}_i \in \mathbf{R}^{m \times n}$ denote the i -th face dataset, and each column of \mathbf{x}_i is a sample of the face of the i -th individual. Assuming there are k classes of face samples, let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k]$. When identifying a target face test sample, $\mathbf{y} \approx \mathbf{X}\hat{\boldsymbol{\alpha}}$ is used for coding, $\hat{\boldsymbol{\alpha}} = [\hat{\boldsymbol{\alpha}}_1, \hat{\boldsymbol{\alpha}}_2, \dots, \hat{\boldsymbol{\alpha}}_k]$, and the coefficient vector $\hat{\boldsymbol{\alpha}}_i$ is the encoding of the i -th individual sample. If \mathbf{y} is from the i -th class, then $\mathbf{y} \approx \mathbf{x}_i\hat{\boldsymbol{\alpha}}_i$ is the best reservation, which means that a large number of coefficients are close to zero in $\hat{\boldsymbol{\alpha}}_k$ ($k \neq i$); only $\hat{\boldsymbol{\alpha}}_i$ remains intact. Thus, the classification (ID) of the target face test sample \mathbf{y} can be decoded by the sparse nonzero coefficient in $\hat{\boldsymbol{\alpha}}$.

The SRC methods [13] are summarized as follows.

Step 1. Given k -class face training sample \mathbf{X} and test sample \mathbf{y} .

Step 2 (dimension adjustment). \mathbf{X} , \mathbf{y} are projected onto the corresponding low-dimensional feature space $\tilde{\mathbf{X}}$, $\tilde{\mathbf{y}}$ using the traditional dimensionality (PCA) reduction technique.

Step 3 (l_2 -norm). Obtain the normalized $\hat{\tilde{\mathbf{X}}}$ column and $\hat{\tilde{\mathbf{y}}}$, respectively.

Step 4. Solve the l_1 -minimization problem.
 $(\hat{\boldsymbol{\alpha}}) = \arg \min \|\boldsymbol{\alpha}\|_1$ satisfies $\hat{\tilde{\mathbf{X}}}\boldsymbol{\alpha} = \hat{\tilde{\mathbf{y}}}$ or $\|\hat{\tilde{\mathbf{X}}}\boldsymbol{\alpha} - \hat{\tilde{\mathbf{y}}}\|_2 < \varepsilon$.

Step 5. Compute the residuals to identify the following:
 $r_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{x}_i\hat{\boldsymbol{\alpha}}_i\|_2$, $\text{ID}(\mathbf{y}) = \arg \min_i [r_i(\mathbf{y})]$, $i = 1, 2, \dots, k$.

2.2. Collaborative Representation Based Classification. Zhang et al. [15] argued that it is the CR but not the l_1 -norm sparsity that makes SRC powerful for face classification. Collaborative representation uses all classes (individuals) of training samples to label the target test samples. In order to reduce the complexity of face detection by \mathbf{X} -coordinate, a regularized least-squares method is proposed by Zhang et al., which is

$$(\hat{\boldsymbol{\rho}}) = \arg \min_{\boldsymbol{\rho}} \{ \|\mathbf{y} - \mathbf{X} \cdot \boldsymbol{\rho}\|_2^2 + \lambda \|\boldsymbol{\rho}\|_2^2 \}, \quad (1)$$

where λ is a regularization parameter. Equation (1) has dual roles, first of which is stabilizing the least-squares method, and, secondly, it proposes a “sparsity” which is much weaker than l_1 -norm to solve $\hat{\boldsymbol{\rho}}$. The CR for the regularized least-squares method of (1) can be solved as follows:

$$\hat{\boldsymbol{\rho}} = (\mathbf{X}^T \mathbf{X} + \lambda \cdot \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}. \quad (2)$$

Let $\mathbf{P} = (\mathbf{X}^T \mathbf{X} + \lambda \cdot \mathbf{I})^{-1} \mathbf{X}^T$. Obviously, since \mathbf{P} and \mathbf{y} have little relevance, they can be precalculated as a projection matrix. When a target test sample \mathbf{y} comes in to be identified, \mathbf{y} is projected onto \mathbf{P} by $\mathbf{P}\mathbf{y}$, thus making CR very fast. The classification by $\hat{\boldsymbol{\rho}}$ is very similar to the classification by $\hat{\boldsymbol{\alpha}}$ in the SRC method. In addition to representing the classification residuals $\|\mathbf{y} - \mathbf{x}_i\hat{\boldsymbol{\rho}}_i\|_2$, $\hat{\boldsymbol{\rho}}_i$ is a coefficient vector

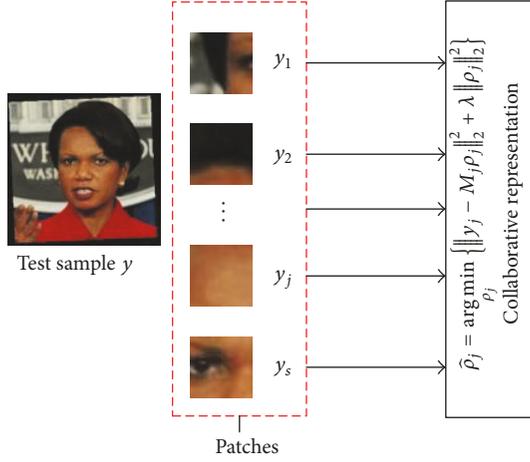


FIGURE 1: Patch based collaborative representation.

associated with class i ; the l_2 -norm “sparsity” $\|\hat{\rho}_i\|_2$ also contains abundant information for classification. The CRC method [15] is summarized as follows.

Step 1. Give k classes of face training sample \mathbf{X} and test sample \mathbf{y} .

Step 2 (reduce the dimension). Use PCA to reduce \mathbf{X} and \mathbf{y} to low-dimensional feature space and obtain $\hat{\mathbf{X}}$ and $\hat{\mathbf{y}}$.

Step 3 (normalization). Normalize the columns of $\hat{\mathbf{X}}$ and $\hat{\mathbf{y}}$ using the unit l_2 -norm.

Step 4. Encode $\hat{\mathbf{y}}$ on $\hat{\mathbf{X}}$:

$$\mathbf{P} = (\mathbf{X}^T \mathbf{X} + \lambda \cdot \mathbf{I})^{-1} \mathbf{X}^T, \hat{\rho} = \mathbf{P} \hat{\mathbf{y}}.$$

Step 5. Compute the regularization for identification:

$$r_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{x}_i \hat{\rho}_i\|_2 / \|\hat{\rho}_i\|_2; \text{ID}(\mathbf{y}) = \arg \min_i [r_i(\mathbf{y})], i = 1, 2, \dots, k.$$

2.3. Patch Based Collaborative Representation. In the equation of sparse representation and collaborative representation, it can be seen that if the linear system determined by the training dictionary \mathbf{X} is underdetermined, the linear representation of the target test sample \mathbf{y} over \mathbf{X} can be very accurate, but in reality available samples in each target are limited; the sparse representation and the cooperative representation method may fail because the linear representation of the target test sample may not be accurate enough. In order to alleviate this problem, Zhu et al. proposed a PCRC method for FR [16], as shown in Figure 1, the target face test sample \mathbf{y} is divided into a set of overlapping face image patches $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_s$ according to the patch size. Each of the divided face image patches \mathbf{y}_j is collaboratively represented on the local dictionary \mathbf{M}_j at the corresponding position of the patch \mathbf{y}_j extracted from \mathbf{X} . In this case, since the linear system determined by the local dictionary \mathbf{M}_j is often underdetermined, the patch based representation is more accurate than the overall face image representation.

3. Patch Based Collaborative Representation Using Gabor Feature and Measurement Matrix for Face Recognition

3.1. Gabor Feature. Gabor feature has been widely used in FR because of its robustness in illumination, expression, and pose compared to holistic feature. Yang and Zhang have applied multiscale and multidirectional Gabor dictionary to the SRC for FR [19], which further improves the robustness of the algorithm. Inspired by previous works, in this paper, we integrate Gabor feature into the PCRC framework to improve its robustness.

A Gabor filter with multidirection μ and multiscale ν is defined as follows [25, 26]:

$$\Psi_{\mu, \nu}(\mathbf{z}) = \frac{\|\mathbf{k}_{\mu, \nu}\|^2}{\delta^2} e^{(-\|\mathbf{k}_{\mu, \nu}\|^2 \|\mathbf{z}\|^2 / 2\delta^2)} \left[e^{i\mathbf{k}_{\mu, \nu} \cdot \mathbf{z}} - e^{-\delta^2/2} \right] \quad (3)$$

$$\mathbf{k}_{\mu, \nu} = \mathbf{k}_\nu \cdot e^{i\nu\mu}, \quad \psi_\mu = \frac{\pi\mu}{8}, \quad \mathbf{k}_\nu = \frac{k_{\max}}{f^\nu},$$

where the coordinates of the pixel are $\mathbf{z} = (x, y)$, k_{\max} is the maximum frequency, and the interval factor of the kernel function distance is denoted as f . The bandwidth of the filter is determined by δ . The convolution of the target image Img and the wavelet kernel $\Psi_{\mu, \nu}$ is expressed as

$$\mathbf{G}_{\mu, \nu} = \text{Img}(\mathbf{z}) * \Psi_{\mu, \nu}(\mathbf{z}) = \mathbf{M}_{\mu, \nu} \cdot \exp(i\theta_{\mu, \nu}(\mathbf{z})), \quad (4)$$

where $\mathbf{M}_{\mu, \nu}$ ($\mu = 0, 1, \dots, 7, \nu = 0, 1, \dots, 4$) is the amplitude of Gabor, $\theta_{\mu, \nu}(\mathbf{z})$ is the phase of Gabor, and the local energy change in the image is expressed by amplitude information. Because the Gabor phase changes periodically with the space position and the amplitude is relatively smooth and stable [19, 25, 26], only the magnitude of Gabor was used in this paper, such as Figure 2.

3.2. Measurement Matrix. Unfortunately, although Gabor feature can be used to enhance the robustness of face image representation, it brings higher dimensions to the training sets than holistic feature does. In other words, the computation cost and computation time are increased. In order to solve the problem caused by higher dimension of the training sets, a further dimension reduction is necessary. We proposed to use PCA [31–33] for our method. The steps of PCA are as follows: assuming N sample images $\mathbf{x}_i \in \mathbf{R}^m$, $i = 1, 2, \dots, N$. Firstly, normalize each sample (subtract the mean, and then divide the variance), convert vector \mathbf{x}_i^* , which accord with normal distribution $(0, 1)$. Secondly, compute the eigenvectors of covariance matrix $\mathbf{Q} = \mathbf{X}\mathbf{X}^T$, $\mathbf{X} = [\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_N^*]$: $\lambda \mathbf{E} = \mathbf{Q}\mathbf{E}$, \mathbf{E} is the eigenvector corresponding to the eigenvalue λ . Thirdly, the eigenvector is sorted according to the size of the eigenvalues; the first p eigenvectors are extracted to form a linear transformation matrix $\mathbf{W}_{\text{PCA}} \in \mathbf{R}^{m \times p}$. We can use $\mathbf{y}_i = \mathbf{W}_{\text{PCA}}^T \mathbf{x}_i$ to reduce dimension. But, if $m \gg N$, the dimension of $\mathbf{X}\mathbf{X}^T \in \mathbf{R}^{m \times m}$ would be very large. In order to solve this problem, singular value decomposition is usually used. The eigenvalues

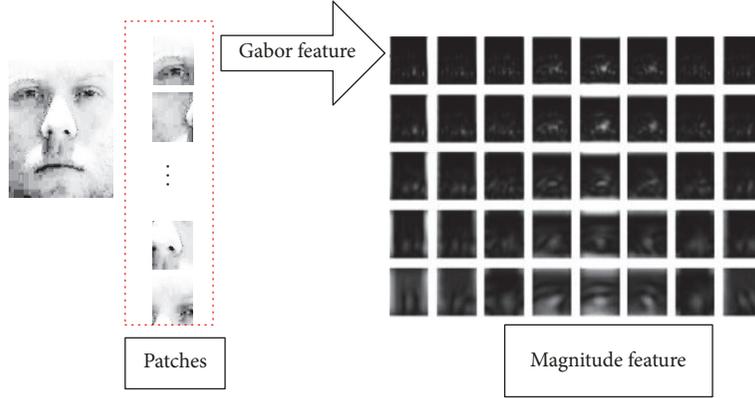


FIGURE 2: Patch based collaborative representation using Gabor feature.

λ and eigenvectors μ_i of XX^T are obtained by calculating the eigenvalues δ_i and eigenvectors v_i of $X^T X$ (the first p): $\mu_i = (1/\sqrt{\delta_i})Xv_i, i = 1, 2, \dots, p$.

In summary, we found that PCA and its related algorithms have two obvious shortcomings [37, 38]. Firstly, the leading eigenvectors encode mostly illumination and expression, rather than discriminating information. Secondly, in the actual calculation, the amount of calculation is very large and will fail in small samples.

Inspired by the method of using random face for feature extraction illustrated in this literature [38], we used the Random Gaussian matrices $\Phi \in \mathbf{R}^{M \times N}$ ($M \ll N$) as a measurement matrix to measure face images. The measurement matrix Φ is used to measure the redundant dictionary \mathbf{X} to obtain $A = \Phi\mathbf{X} \in \mathbf{R}^{M \times S}$, where $\mathbf{X} = [x_1, x_2, \dots, x_S] \in \mathbf{R}^{N \times S}$. In Figure 3(a), for any test image x , measurements y were obtained by $y = \Phi x \in \mathbf{R}^{M \times 1}$. In essence, utilizing the measurement matrix to reduce the dimension of the image is different from the sparse representation theory. The dimension M of the measurements y was measured by the measurement matrix Φ and was not limited by the number of training samples.

However, some literatures [35, 39] suggested that the Random Gaussian matrices are uncertain and limit its practical application. And Toeplitz and cyclic matrices were proposed for signal reconstruction. The Toeplitz and cyclic matrices rotate the row vectors to generate all matrices. Usually, the value of the vector in Toeplitz and cyclic matrices is ± 1 , and each element is independent of the Bernoulli distribution. Therefore, it is easy to implement hardware in practical application. Based on the above analysis, we further used Toeplitz and cyclic matrices and their improved method (namely, the Deterministic sparse Toeplitz matrices [36]) for our method. Some of the relevant measurement matrices used in this paper will be described in detail. The operating mechanism of each measurement matrix is shown in Figure 3.

3.2.1. Random Gaussian Matrices. The format of $m \times n$ Random Gaussian matrices is expressed as follows [34, 38]:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (5)$$

each element $a_{i,j}$ is independently subject to a Gaussian distribution whose mean is 0 and variance is $1/m$.

3.2.2. Toeplitz and Cyclic Matrices. The concrete form of $M \times N$ Toeplitz and cyclic matrices [35, 39] is presented below:

$$\begin{bmatrix} a_N & a_{N-1} & \cdots & a_1 \\ a_{N+1} & a_N & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{N+M-1} & a_{N+M-2} & \cdots & a_M \end{bmatrix}, \quad (6)$$

$$\begin{bmatrix} a_N & a_{N-1} & \cdots & a_1 \\ a_1 & a_N & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{M-1} & a_{M-2} & \cdots & a_M \end{bmatrix}. \quad (7)$$

Equation (6) is the Toeplitz matrices; the main diagonal $a_{i,j} = a_{i+1,j+1}$ are constants. If $a_i = a_{N+i}$, the additional condition of (6) is satisfied; then it becomes a cyclic matrices, and its element $\{a_i\}_{i=1}^{N+M-1}$ follows a certain probability distribution $P(a)$.

3.2.3. Deterministic Sparse Toeplitz Matrices. The construction of Deterministic sparse Toeplitz matrices [36] is based on the Toeplitz and cyclic matrices, which is illustrated in this paper by the example of the random spacing Toeplitz matrices with an interval of $\Delta = 2$. The independent elements in the first row and the first column of (6) constitute the vector \mathbf{T} :

$$[a_N \ a_{N-1} \ \cdots \ a_1 \ a_{N+1} \ \cdots \ a_{N+M-1}]. \quad (8)$$

Conducting random sparse spacing to (7), one can see that T contains all the independent elements in (8). Then

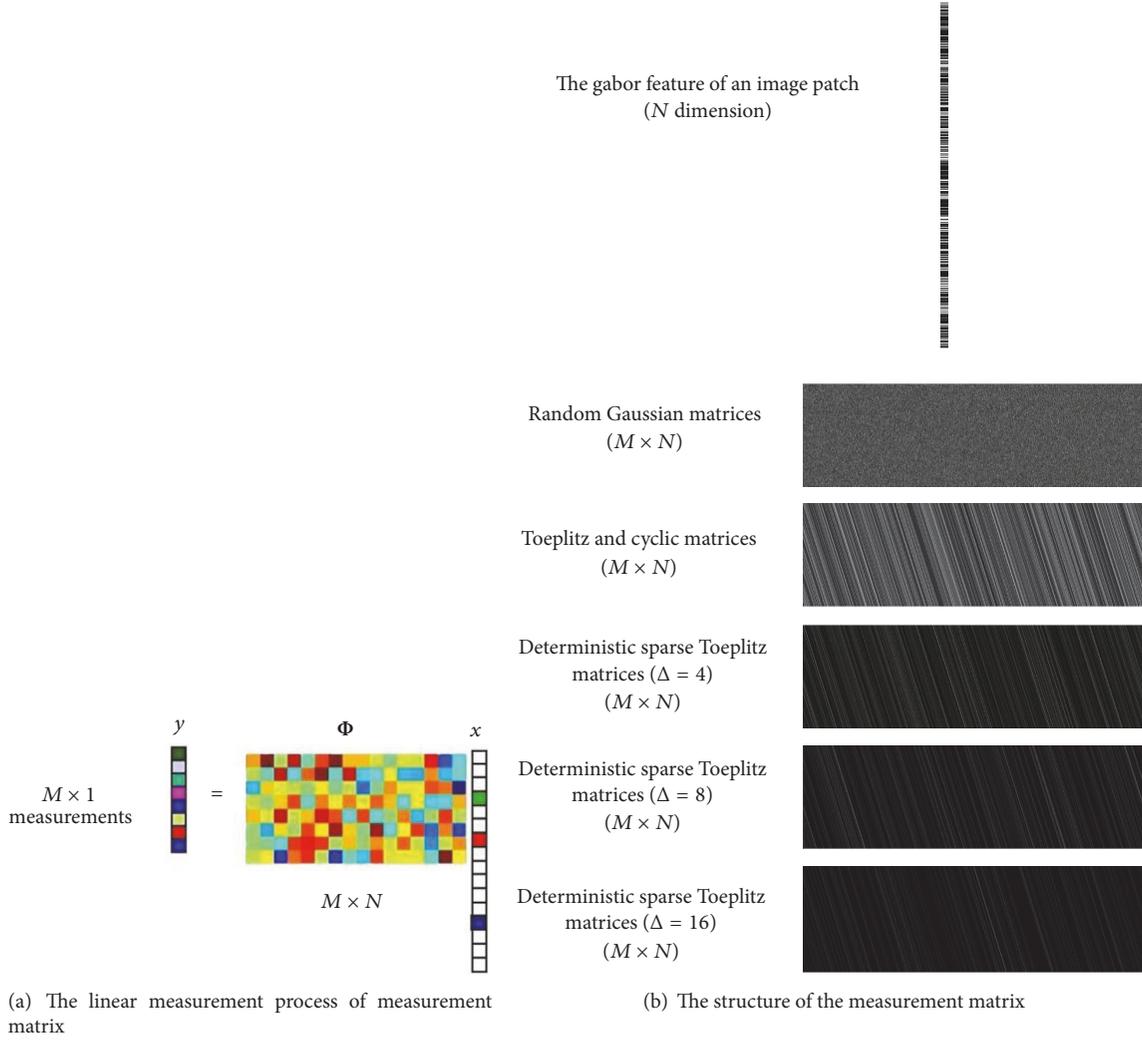


FIGURE 3: The operating mechanism of measurement matrix.

assignment operations are performed on T , where the elements a_i ($i \in \Lambda$, Λ is the $(N + M - 1)/\Delta$ indexes randomly chosen from the index sequence $1 \rightarrow N + M - 1$) obey the independent and identically distributed (i.i.d.) Random Gaussian distribution, while the other elements are 0. Finally, we obtained the Deterministic sparse Toeplitz matrices $a_{i+1,j+1} = a_{i,j}$ according to the characteristics of construction of the Toeplitz and cyclic matrices.

3.3. The Proposed Face Recognition Approach. Although the PCRC can indeed solve the problem of small sample size, this method is still based on the original feature of the patch, and the robustness and accuracy are yet to be improved. Based on the above analysis, Gabor feature and measurement matrix are infused into the PCRC for FR, which not only solves the problem of small sample size but also enhances the robustness and efficiency of the method.

Our proposed method for FR is summarized as follows.

Step 1 (input). Face training sample \mathbf{X} and test sample \mathbf{y} .

Step 2 (patch). Divide the k face training samples \mathbf{X} into s patches and divide the test sample \mathbf{y} into s patches, where \mathbf{y}_i and \mathbf{M}_{ij} are the position corresponding to the sample patches:

$$\mathbf{X} = \begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_i \\ \vdots \\ \mathbf{M}_s \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{11} & \cdots & \mathbf{M}_{1j} & \cdots & \mathbf{M}_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{M}_{i1} & \cdots & \mathbf{M}_{ij} & \cdots & \mathbf{M}_{ik} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{M}_{s1} & \cdots & \mathbf{M}_{sj} & \cdots & \mathbf{M}_{sk} \end{bmatrix}, \quad (9)$$

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_i \\ \vdots \\ \mathbf{y}_s \end{bmatrix}.$$

Step 3 (extract and measure features). (1) Extract Gabor feature from patches of training samples and test sample, respectively, to obtain

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_i \\ \vdots \\ \mathbf{G}_s \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{11} & \cdots & \mathbf{G}_{1j} & \cdots & \mathbf{G}_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{G}_{i1} & \cdots & \mathbf{G}_{ij} & \cdots & \mathbf{G}_{ik} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{G}_{s1} & \cdots & \mathbf{G}_{sj} & \cdots & \mathbf{G}_{sk} \end{bmatrix}, \quad (10)$$

$$\mathbf{Gy} = \begin{bmatrix} \mathbf{Gy}_1 \\ \vdots \\ \mathbf{Gy}_i \\ \vdots \\ \mathbf{Gy}_s \end{bmatrix}.$$

(2) Carry out the measurement and reduce the dimension of each patch \mathbf{G}_{ij} ($i = 1, 2, \dots, s$, $j = 1, 2, \dots, k$) using the measurement matrix Φ , to obtain measured signal:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_i \\ \vdots \\ \mathbf{A}_s \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{11} & \cdots & \mathbf{a}_{1j} & \cdots & \mathbf{a}_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_{i1} & \cdots & \mathbf{a}_{ij} & \cdots & \mathbf{a}_{ik} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_{s1} & \cdots & \mathbf{a}_{sj} & \cdots & \mathbf{a}_{sk} \end{bmatrix}$$

$$= \begin{bmatrix} \Phi \mathbf{G}_{11} & \cdots & \Phi \mathbf{G}_{1j} & \cdots & \Phi \mathbf{G}_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Phi \mathbf{G}_{i1} & \cdots & \Phi \mathbf{G}_{ij} & \cdots & \Phi \mathbf{G}_{ik} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Phi \mathbf{G}_{s1} & \cdots & \Phi \mathbf{G}_{sj} & \cdots & \Phi \mathbf{G}_{sk} \end{bmatrix}, \quad (11)$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_i \\ \vdots \\ \mathbf{b}_s \end{bmatrix} = \begin{bmatrix} \Phi \mathbf{Gy}_1 \\ \vdots \\ \Phi \mathbf{Gy}_i \\ \vdots \\ \Phi \mathbf{Gy}_s \end{bmatrix}.$$

Step 4. Collaborative representation of each patch measurement signal \mathbf{b}_i of the test sample over the training sample measurement signal local dictionary \mathbf{A}_i :

$$\hat{\rho}_i = \arg \min_{\rho_i} \left\{ \|\mathbf{b}_i - \mathbf{A}_i \rho_i\|_2^2 + \lambda \|\rho_i\|_2^2 \right\} \quad i = 1, 2, \dots, s. \quad (12)$$

Equation (12) can be easily derived as follows:

$$\hat{\rho}_i = \mathbf{P} \mathbf{b}_i, \quad \mathbf{P} = (\mathbf{A}_i^T \mathbf{A}_i + \lambda \mathbf{I})^{-1} \mathbf{A}_i^T \quad i = 1, 2, \dots, s, \quad (13)$$

where \mathbf{I} is unit matrix.

Step 5. Recognition results of patch \mathbf{y}_i are

$$r_{ij}(\mathbf{y}_i) = \frac{\|\mathbf{b}_i - \mathbf{a}_{ij} \hat{\rho}_i\|_2}{\|\hat{\rho}_i\|_2} \quad i = 1, 2, \dots, s, \quad j = 1, 2, \dots, k, \quad (14)$$

$$\text{ID}(r_{ij}) = \arg \min_j [r_{ij}(\mathbf{y}_i)].$$

Step 6. Use voting to obtain the final recognition result.

4. Experimental Analysis

The proposed approaches, GPCRC and its improved methods, were evaluated in three publicly available databases: Extended Yale B [40–42], CMU_PIE [43, 44], and LFW [45, 46]. To show the effectiveness of the GPCRC, we compared our methods with four classical methods and their improved methods, namely, SRC [13], CRC [15], SSRC [14], and PCRC [16]. Then, we tested our GPCRC using several measurement matrix, namely, Random Gaussian matrices, Toeplitz, and cyclic matrices and Deterministic sparse Toeplitz matrices ($\Delta = 4, 8, 16$) to reduce the dimension of the transformed signal and compare their performance with the conventional PCA methods. In all the following experiments, we ran the MATLAB 2015b on a typical Intel(R) Core(TM) i5-3470 CPU 3.20 GHz, Windows7 x64 PC. In our implementation of Gabor filters, the parameters were set as $f = \sqrt{2}$, $K_{\max} = \pi/2$, $\sigma = \pi$, $\mu = 0, 1, \dots, 7$, $\nu = 0, 1, \dots, 4$ for all the experiments below. All the programs were run 20 times on each database and the mean and variance of each method were reported. In order to report the best result of each method, the parameter λ used in SRC, SSRC, CRC, and PCRC was set to 0.001 [13], 0.001 [14], 0.005 [15], and 0.001 [16, 17], respectively. The patch sizes were used in PCRC and GPCRC and its improved methods are 8×8 [16, 17] and 16×16 , respectively.

4.1. Extended Yale B. To test the robustness of the proposed method on illumination, we used the classic Extended Yale B database [40–42], because faces from Extended Yale B database were acquired in different illumination conditions. The Extended Yale B database contains 38 human subjects under 9 poses and 64 illumination conditions. For obvious comparison, all frontal-face images marked with P00 were used in our experiment, and the face images size was down-sampled to 32×32 . 10, 20 faces from each individual were selected randomly as training sample; 30 others from each individual were selected as test sample. Figure 4 shows some P00 marked samples from the Extended Yale B database. The experimental results of each method are shown in Table 1. In Figure 5 we compared the recognition rates of

TABLE 1: Recognition rate (%) on Extended Yale B database.

Features	Methods	Training sample size	
		10	20
Intensity feature	SRC	76.15 ± 9.60	80.50 ± 4.32
	CRC	78.60 ± 10.58	83.00 ± 6.64
	SSRC	87.53 ± 7.91	90.01 ± 4.38
	PCRC	90.44 ± 1.09	94.25 ± 1.01
LBP feature	SRC	83.56 ± 11.32	89.74 ± 9.52
	CRC	83.02 ± 10.19	90.30 ± 3.27
	SSRC	90.58 ± 6.39	95.89 ± 3.01
	PCRC	79.67 ± 12.57	87.25 ± 5.08
Gabor feature	SRC	85.25 ± 5.36	91.50 ± 3.50
	CRC	86.18 ± 4.43	90.00 ± 2.80
	SSRC	87.06 ± 9.49	92.24 ± 5.06
	PCRC	94.34 ± 2.06	98.83 ± 0.94



FIGURE 4: Some marked samples of the Extended Yale B database.

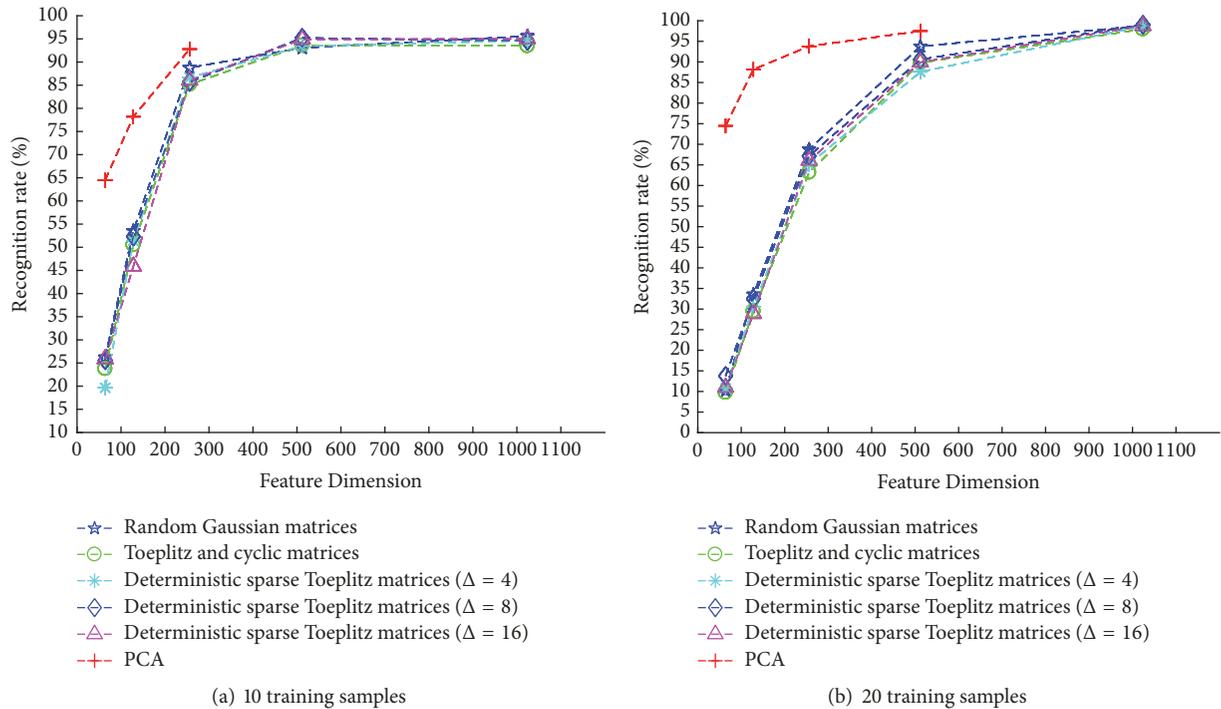


FIGURE 5: The performance of each dimension reduction method for GPCRC on Extended Yale B database.

TABLE 2: The speed of each dimension reduction methods.

Method	Random Gaussian	Toeplitz and cyclic	PCA
Time(s)	1.1131	1.0836	41.7295
Method	Deterministic sparse Toeplitz		
	($\Delta = 4$)	($\Delta = 8$)	($\Delta = 16$)
Time(s)	1.2643	1.2226	1.2678



FIGURE 6: Some marked samples of the CMU_PIE database.

various methods with the space dimensions 32, 64, 128, 256, 512, and 1024 for each patch in GPCRC, and those numbers correspond to downsampling ratios of 1/320, 1/160, 1/80, 1/40, 1/20, and 1/10 respectively. In Table 1, it can be clearly seen that GPCRC achieves the highest recognition rate performance in all experiments. In Figure 6, we can see the performance of various dimensionality reduction methods for GPCRC. 10 samples of each individual in Figure 5(a) were used as training samples. When the feature dimension is low (≤ 256), the best performance of PCA is 92.80% (256 dimension), which is significantly higher than that of the measurement matrix at 256 dimension: Random Gaussian matrices are 88.77%, Toeplitz and cyclic matrices are 85.17%, Deterministic sparse Toeplitz matrices ($\Delta = 4$) are 86.67%, Deterministic sparse Toeplitz matrices ($\Delta = 8$) are 85.69%, and Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 86.05%, but none of these has been as good as the performance of the original dimension (10240 dimension). So the reason why the performance of PCA is significantly better than the measurement matrix at low dimension can be analyzed through their operation mechanism: the PCA transforms the original data into a set of linearly independent representations of each dimension in a linear transformation, which can extract the principal component of the data. These principal components can represent image signals more accurately [31–33]. The theory of compressed sensing [34–36] states that when the signal is sparse or compressible in a transform domain, the measurement matrix which is noncoherent with the transform matrix can be used to project transform coefficients to the low-dimensional vector, and this projection can maintain the information for signal reconstruction. The compressed sensing technology can achieve the reconstruction with high accuracy or high probability using small number of projection data. In the case of extremely low dimension, the reconstruction information is insufficient, and the original signal cannot be reconstructed accurately. So the recognition rate is not high. When the dimension is increased and the reconstructed information can more accurately reconstruct the original signal, the

recognition rate will improve. When the feature dimension is higher (≥ 512), PCA cannot work because of the sample size. At 512 dimension, the measurement matrix (Deterministic sparse Toeplitz matrices ($\Delta = 8$) are 95.21%; Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 94.91%) has reached the performance of the original dimension (10240 dimension), and the dimension is only 1/20 of the original dimension. At the 1/10 of the original dimension, the performance of the measurement matrix has basically achieved the performance of the original dimension: Random Gaussian matrices are 95.60%, Deterministic sparse Toeplitz matrices ($\Delta = 4$) are 94.64%, Deterministic sparse Toeplitz matrices ($\Delta = 8$) are 94.56%, and Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 94.98%. In Figure 6(b), 20 samples of each individual were used as training samples. When the feature dimension ≤ 512 , the best performance is PCA which is 97.50% (512 dimension). At 1024 dimension, PCA cannot work and the performance of the measurement matrix has basically achieved the performance of the original dimension: Random Gaussian matrices are 98.80%, Deterministic sparse Toeplitz matrices ($\Delta = 4$) are 98.68%, Deterministic sparse Toeplitz matrices ($\Delta = 8$) are 99.01%, and Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 98.77%. In Table 2, we fixed the dimension of each reduction methods at 512 dimension in the case of 20 training samples. In general, the complexity of PCA is $O(n^3)$ [47], where n is the number of rows of the covariance matrix. The example we provided consists of 38 individuals, and each individual contains 20 samples. Since the number of Gabor features (10240) for each patch far outweighs the total number of training samples (760), the number of rows of the covariance matrix is determined by the total number of training samples, and the complexity of PCA is approximately $O(760^3)$. However, in the proposed measurement matrix algorithm, the measurement matrix is formed by a certain deterministic distribution (e.g., Gaussian distribution) and structure, according to the required dimension and length of the signal, so that the measurement matrix demonstrates low complexity [36, 48, 49]. The actual speed of each dimension reduction methods for one patch of all samples (including all training samples and all testing samples) is listed. Based on the above analysis, we can see that the PCA method has the most outstanding performance, but it is limited by the sample size and it is time-consuming. The measurement matrix does not perform well in low dimension, but is not limited by the sample size. When the dimension reaches a certain value, its performance was the same as that of the original dimension. Thus, the dimension of the data can be reduced without loss of recognition rate.

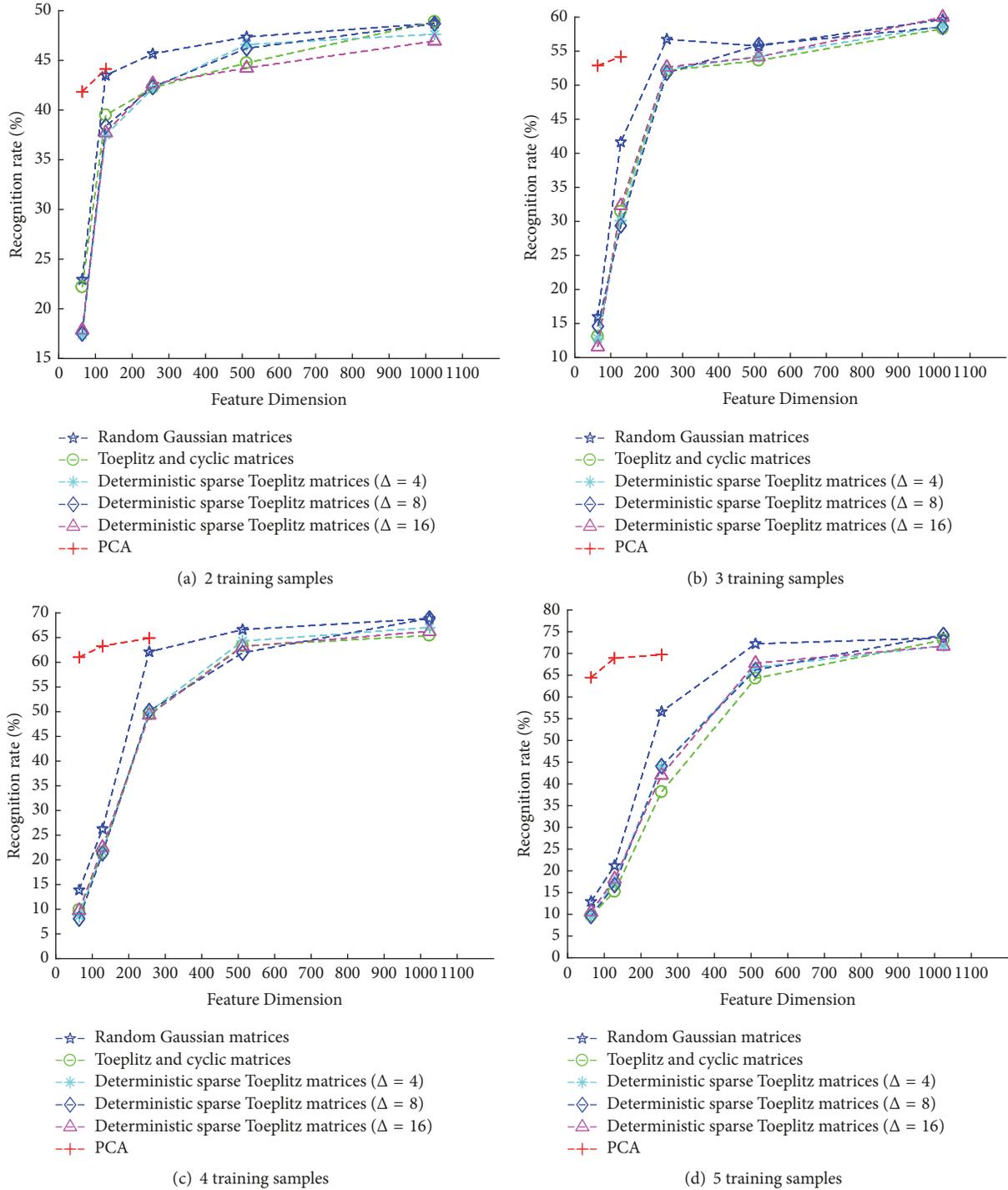


FIGURE 7: The performance of each dimension reduction method for GPCRC on CMU_PIE database.

4.2. *CMU_PIE*. In order to further test the robustness in illumination, pose, and expression, we utilized a currently popular database CMU_PIE [43, 44], a database consisting of 41,368 images of 68 people; for each person, there are 13 different poses, 43 different illumination conditions, and 4 different expressions. To testify the advantage of our method in small sample size condition, we randomly selected 2, 3, 4,

and 5 face samples from each individual as training sample and other 40 face samples from each individual as test samples. The face images size is downsampled to 32×32 . Figure 6 shows some marked samples of the CMU_PIE database. The experimental results of each method are shown in Table 3; GPCRC has the best results. In Figure 7 we compared the recognition rates of various dimension reduction methods

TABLE 3: Recognition rate (%) on CMU_PIE database.

Features	Method	Training sample size			
		2	3	4	5
Intensity feature	SRC	28.21 ± 5.43	39.36 ± 9.82	44.73 ± 15.29	60.54 ± 6.45
	CRC	30.34 ± 6.15	40.62 ± 9.20	44.35 ± 7.42	61.03 ± 7.82
	SSRC	37.52 ± 8.34	54.74 ± 8.50	57.32 ± 9.82	66.73 ± 8.85
	PCRC	42.32 ± 1.17	56.46 ± 1.88	61.16 ± 2.91	70.47 ± 3.87
LBP feature	SRC	38.42 ± 6.46	53.37 ± 7.10	53.67 ± 5.21	63.31 ± 7.24
	CRC	37.32 ± 7.75	54.83 ± 6.15	54.27 ± 8.31	64.25 ± 6.90
	SSRC	43.17 ± 5.71	56.84 ± 9.21	64.53 ± 4.76	68.25 ± 4.35
	PCRC	41.21 ± 7.10	52.21 ± 11.76	62.54 ± 6.21	70.22 ± 3.07
Gabor feature	SRC	41.43 ± 4.20	53.36 ± 5.40	53.58 ± 7.32	66.79 ± 5.45
	CRC	40.25 ± 5.15	55.35 ± 3.28	56.68 ± 7.42	66.24 ± 6.18
	SSRC	44.55 ± 6.00	55.96 ± 6.20	62.50 ± 3.16	70.41 ± 4.79
	PCRC	48.57 ± 1.29	58.45 ± 1.94	68.82 ± 2.56	73.18 ± 2.44

for GPCRC. Similarly, we can see that PCA cannot achieve the best recognition rate because of the sample size limit, and at 1024 dimension the performance of the measurement matrix has basically achieved the performance of the original dimension. The recognition rate of each measurement matrix is shown below:

- (i) Random Gaussian matrices are **48.73%** (2 training samples), **59.66%** (3 training samples), **68.86%** (4 training samples), and **73.60%** (5 training samples).
- (ii) Toeplitz and cyclic matrices are **48.89%** (2 training samples), 58.30% (3 training samples), 65.44% (4 training samples), and 73.12% (5 training samples).
- (iii) Deterministic sparse Toeplitz matrices ($\Delta = 4$) are 47.63% (2 training samples), **58.68%** (3 training samples), 67.03% (4 training samples), and 71.78% (5 training samples).
- (iv) Deterministic sparse Toeplitz matrices ($\Delta = 8$) are **48.69%** (2 training samples), **58.55%** (3 training samples), **68.91%** (4 training samples), and **74.27%** (5 training samples).
- (v) Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 46.94% (2 training samples), **60.00%** (3 training samples), 66.25% (4 training samples), and 71.69% (5 training samples).

Through the above analysis, we further validated the feasibility of the measurement matrix in the premise of ensuring the recognition rate.

4.3. LFW. In practice, the number of training samples is very limited, and only one or two can be obtained from individual identity documents. To simulate the actual situation, we select the LFW database. The LFW database includes frontal images of 5,749 different subjects in unconstrained environment [45]. LFW-a is a version of LFW after alignment using commercial face alignment software [46]. We chose 158 subjects from LFW-a; each subject contains no less than ten samples. For each subject, we randomly choose 1 to 2 samples from each individual for training and another 5 samples from each



FIGURE 8: Some marked samples of the LFW database.

individual for testing. The face images size is downsampled to 32×32 . Figure 8 shows some marked samples of the LFW database. The experimental results of each method are shown in Table 4; it can be clearly seen that the GPCRC achieves the highest recognition rate performance in all experiments with the training sample size from 1 to 2. In Figure 9, we can also see the advantages of the measurement matrix in small sample size. At 1024 dimension, the performance of the measurement matrix in the single training sample is as follows: Random Gaussian matrices are **26.18%**, Toeplitz and cyclic matrices are 24.68%, Deterministic sparse Toeplitz matrices ($\Delta = 4$) are 24.43%, Deterministic sparse Toeplitz matrices ($\Delta = 8$) are **25.85%**, and Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 25.28%, and the performances of the measurement matrix in the two training samples are as follows: Random Gaussian matrices are **39.68%**, Toeplitz and cyclic matrices are 37.72%, Deterministic sparse Toeplitz matrices ($\Delta = 4$) are 37.56%, Deterministic sparse Toeplitz matrices ($\Delta = 8$) are **39.91%**, and Deterministic sparse Toeplitz matrices ($\Delta = 16$) are 38.25%.

5. Conclusion

In order to alleviate the influence of unreliability environment on small sample size in FR, in this paper we applied improved method for Gabor local features to PCRC; we proposed to use the measurement matrices to reduce the dimension of the transformed signal. Several important observations can be summarized as follows. (1) The proposed GPCPC method can effectively deal with the influence of unreliability environment in small sample FR. (2) The measurement

TABLE 4: Recognition rate (%) on LFW database.

Features	Method	Training sample size	
		1	2
Intensity feature	SRC	10.58 ± 7.20	19.02 ± 8.29
	CRC	9.08 ± 8.15	20.39 ± 7.20
	SSRC	17.03 ± 5.87	28.52 ± 7.50
	PCRC	22.29 ± 3.96	33.29 ± 5.28
LBP feature	SRC	15.52 ± 5.54	25.31 ± 7.42
	CRC	16.23 ± 6.21	26.22 ± 6.57
	SSRC	19.16 ± 3.41	32.10 ± 4.26
	PCRC	21.26 ± 4.73	31.59 ± 6.11
Gabor feature	SRC	18.68 ± 4.23	31.35 ± 6.40
	CRC	17.31 ± 8.15	30.83 ± 5.20
	SSRC	21.30 ± 7.21	33.67 ± 4.33
	PCRC	25.82 ± 3.42	39.60 ± 3.54

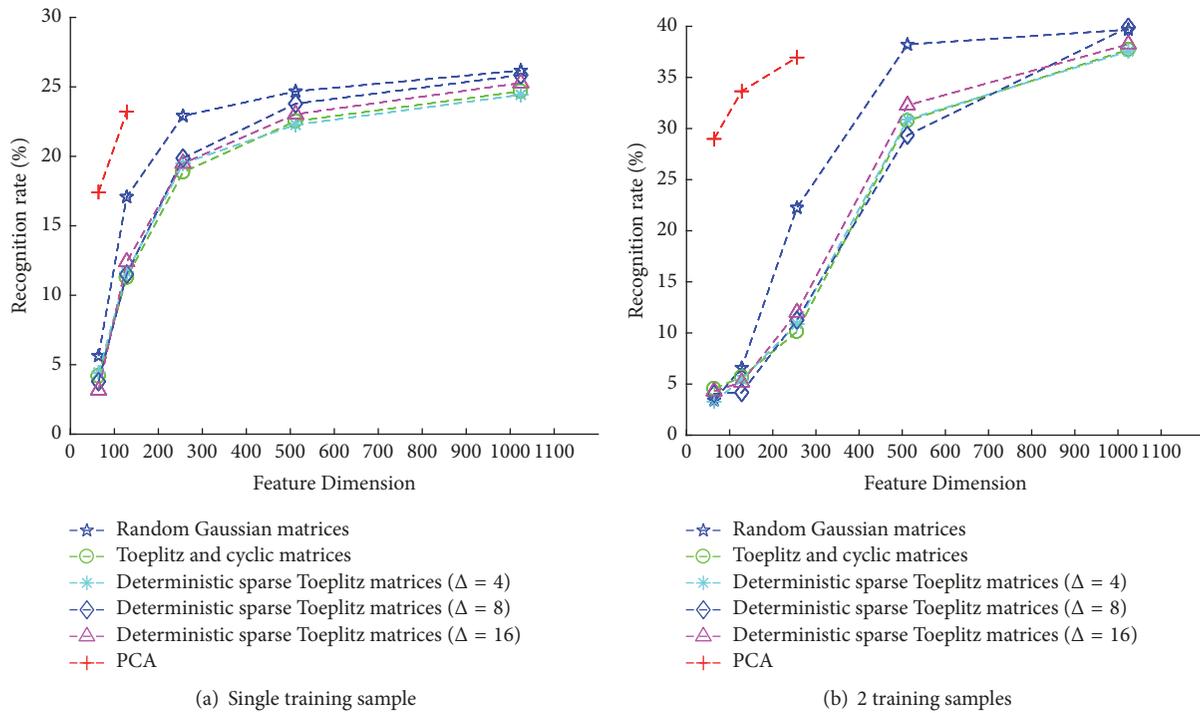


FIGURE 9: The performance of each dimension reduction method for GPCRC on LFW database.

matrix proposed to deal with the high dimension in GPCRC method can effectively improve the computational efficiency and computational speed, and they were able to overcome the limitations of the PCA method.

Conflicts of Interest

None of the authors have conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61672386), Humanities and Social

Sciences Planning Project of Ministry of Education (no. 16YJAZH071), Anhui Provincial Natural Science Foundation of China (no. 1708085MF142), University Natural Science Research Project of Anhui Province (no. KJ2017A259), and Provincial Quality Project of Anhui Province (no. 2016zy131).

References

[1] J. Galbally, S. Marcel, and J. Fierrez, “Biometric antispoofing methods: a survey in face recognition,” *IEEE Access*, vol. 2, pp. 1530–1552, 2014.
 [2] S. Nagpal, M. Singh, R. Singh, and M. Vatsa, “Regularized deep learning for face recognition with weight variations,” *IEEE Access*, vol. 3, pp. 3010–3018, 2015.

- [3] J. W. Tanaka and D. Simonyi, "The parts and wholes of face recognition: a review of the literature," *The Quarterly Journal of Experimental Psychology*, no. 10, pp. 1–37, 2016.
- [4] W. Yang, Z. Wang, and C. Sun, "A collaborative representation based projections method for feature extraction," *Pattern Recognition*, vol. 48, no. 1, pp. 20–27, 2015.
- [5] O. Boiman, E. Shechtman, and M. Irani, "In defense of nearest-neighbor based image classification," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.
- [6] K. VenkataNarayana, V. Manoj, and K. K. Swathi, "Enhanced face recognition based on PCA and SVM," *International Journal of Computer Applications*, vol. 117, no. 2, pp. 40–42, 2015.
- [7] E. Owusu and Q. R. Mao, "An svm cadaboost-based face detection system," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 26, no. 4, pp. 477–491, 2014.
- [8] B. V. Dasarathy, "Nearest neighbor (NN) norms: NN pattern classification techniques," *Los Alamitos IEEE Computer Society Press*, vol. 13, no. 100, pp. 21–27, 1991.
- [9] Y. Liu, S. S. Ge, C. Li, and Z. You, " κ -NS: a classifier by the distance to the nearest subspace," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 22, no. 8, pp. 1256–1268, 2011.
- [10] J. Hamm and D. D. Lee, "Grassmann discriminant analysis: a unifying view on subspace-based learning," in *Proceedings of the 25th International Conference on Machine Learning*, pp. 376–383, July 2008.
- [11] R. Wang, S. Shan, X. Chen, and W. Gao, "Manifold-manifold distance with application to face recognition based on image set," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.
- [12] J. Chien and C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1644–1649, 2002.
- [13] J. Wright, A. Ganesh, Z. Zhou, A. Wagner, and Y. Ma, "Demo: robust face recognition via sparse representation," in *Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition (FG '08)*, pp. 1–2, IEEE, Amsterdam, The Netherlands, September 2008.
- [14] W. Deng, J. Hu, and J. Guo, "In defense of sparsity based face recognition," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 399–406, June 2013.
- [15] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 471–478, Barcelona, Spain, November 2011.
- [16] P. Zhu, L. Zhang, Q. Hu, and S. C. K. Shiu, "Multi-scale patch based collaborative representation for face recognition with margin distribution optimization," in *European Conference on Computer Vision*, vol. 7572 of *Lecture Notes in Computer Science*, pp. 822–835, 2012.
- [17] D. Yang, W. Chen, J. Wang, and Y. Xu, "Patch based face recognition via fast collaborative representation based classification and expression insensitive two-stage voting," *International Conference on Computational Science and Its Applications*, vol. 9787, pp. 562–570, 2016.
- [18] J. Lai and X. Jiang, "Classwise sparse and collaborative patch representation for face recognition," *IEEE Transactions on Image Processing*, vol. 25, no. 7, pp. 3261–3272, 2016.
- [19] M. Yang and L. Zhang, "Gabor feature based sparse representation for face recognition with gabor occlusion dictionary," in *Proceedings of the 11th European Conference on Computer Vision (ECCV '10)*, pp. 448–461, Springer, Crete, Greece, 2010.
- [20] M. R. D. Rodavia, O. Bernaldez, and M. Ballita, "Web and mobile based facial recognition security system using Eigenfaces algorithm," in *Proceedings of the 2016 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE '16)*, pp. 86–92, Bangkok, Thailand, December 2016.
- [21] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.
- [22] L. P. Yang, W. G. Gong, W. H. Li, and X. H. Gu, "Random sampling subspace locality preserving projection for face recognition," *Optics & Precision Engineering*, vol. 16, no. 8, pp. 1465–1470, 2008.
- [23] X. Sun, M. Chen, and A. Hauptmann, "Action recognition via local descriptors and holistic features," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 58–65, June 2009.
- [24] X. Wang, Y. Zhang, X. Mu, and F. S. Zhang, "The face recognition algorithm based on improved lbp," *International Journal of Optoelectronic Engineering*, vol. 1, no. 4, pp. 76–80, 2012.
- [25] J. Ahmad, U. Ali, and R. J. Qureshi, "Fusion of thermal and visual images for efficient face recognition using gabor filter," in *Proceedings of the IEEE International Conference on Computer Systems and Applications*, pp. 135–139, March 2006.
- [26] V. Štruc, R. Gajšek, and N. Pavešić, "Principal gabor filters for face recognition," in *Proceedings of the IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems (BTAS '09)*, pp. 1–6, Washington, DC, USA, September 2009.
- [27] M. Li, X. Yu, K. H. Ryu, S. Lee, and N. Theera-Umpon, "Face recognition technology development with gabor, PCA and SVM methodology under illumination normalization condition," *Cluster Computing*, no. 3, pp. 1–10, 2017.
- [28] J. Yang, D. Zhang, and A. F. Frangi, "Two-dimensional PCA: a new approach to appearance -based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.
- [29] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data: with application to face recognition," *Pattern Recognition*, vol. 34, no. 10, pp. 2067–2070, 2001.
- [30] J. Kim, J. Choi, J. Yi, and M. Turk, "Effective representation using ICA for face recognition robust to local distortion and partial occlusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, pp. 1977–1981, 2005.
- [31] C. G. Turhan and H. S. Bilge, "Class-wise two-dimensional PCA method for face recognition," *IET Computer Vision*, vol. 11, no. 4, pp. 286–300, 2017.
- [32] A. Mashhoori and M. Zolghadri Jahromi, "Block-wise two-directional 2DPCA with ensemble learning for face recognition," *Neurocomputing*, vol. 108, pp. 111–117, 2013.
- [33] M. P. Rajath Kumar, R. Keerthi Sravan, and K. M. Aishwarya, "Artificial neural networks for face recognition using PCA and BPNN," in *Proceedings of the 35th IEEE Region 10 Conference (TENCON '15)*, pp. 1–6, IEEE, Macao, China, November 2015.
- [34] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

- [35] W. U. Bajwa, J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak, "Toeplitz-structured compressed sensing matrices," in *Proceedings of the IEEE/SP 14th Workshop on Statistical Signal Processing (SSP '07)*, pp. 294–298, IEEE, Madison, WI, USA, August 2007.
- [36] C. Zhang, H.-R. Yang, and S. Wei, "Compressive sensing based on deterministic sparse toeplitz measurement matrices with random pitch," *Acta Automatica Sinica*, vol. 38, no. 8, pp. 1362–1369, 2012.
- [37] E. S. Takemura, M. R. Petraglia, and A. Petraglia, "An image super-resolution algorithm based on wiener filtering and wavelet transform," in *Proceedings of the 2013 IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE '13)*, pp. 130–134, August 2013.
- [38] M. Yang, *Face Recognition via Sparse Representation*, Wiley Encyclopedia of Electrical and Electronics Engineering, 2008.
- [39] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [40] A. S. Georghiadis, P. N. Belhumeur, and D. J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.
- [41] N. McLaughlin, J. Ming, and D. Crookes, "Largest matching areas for illumination and occlusion robust face recognition," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 796–808, 2017.
- [42] Y. Xu, X. Fang, J. Wu, X. Li, and D. Zhang, "Discriminative transfer subspace learning via low-rank and sparse representation," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 850–863, 2016.
- [43] H. Roy and D. Bhattacharjee, "Local-gravity-face (lg-face) for illumination-invariant and heterogeneous face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 7, pp. 1412–1424, 2016.
- [44] Y. Tai, J. Yang, Y. Zhang, L. Luo, J. Qian, and Y. Chen, "Face recognition with pose variations and misalignment via orthogonal procrustes regression," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 25, no. 6, pp. 2673–2683, 2016.
- [45] G. B. Huang, M. Narayana, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," *Month*, 2008.
- [46] L. Wolf, T. Hassner, and Y. Taigman, "Similarity scores based on background samples," in *Asian Conference on Computer Vision*, vol. 5995, pp. 88–97, 2009.
- [47] J. Wang, *Geometric Structure of High-Dimensional Data and Dimensionality Reduction*, Springer, 2012.
- [48] Q. Wang, J. Li, and Y. Shen, "A survey on deterministic measurement matrix construction algorithms in compressive sensing," *Acta Electronica Sinica*, vol. 41, no. 10, pp. 2041–2050, 2013.
- [49] W. Yin, S. Morgan, J. Yang, and Y. Zhang, "Practical compressive sensing with toeplitz and circulant matrices," in *Proceedings of the Visual Communications and Image Processing 2010*, vol. 7744, July 2010.



Hindawi

Submit your manuscripts at
www.hindawi.com

