Face Recognition Using Improved Extended Sparse Representation Classifier and Feature Descriptor

Mengmeng Liao and Xiaodong Gu$^{(\boxtimes)}$

Department of Electronic Engineering, Fudan University,
Shanghai 200433, China
xdgu@fudan.edu.cn

Abstract. Representing query samples, many methods based on sparse representation do not take into account the different importance of atoms. In this paper, we propose a new extended sparse weighted representation classifier (ESWRC). In ESWRC, we introduce a representativeness estimator, and use it to estimate the atom representativeness. The atom representativeness is used to construct the weights of atoms. The weighted atoms are used to represent the query samples. In addition, we propose a distinctive feature descriptor, called logarithmic weighted sum (LWS) feature descriptor, which combines the advantages of discrete orthonormal S-transform feature, Gabor feature, covariance and logarithmic operation. We combine ESWRC and LWS for face recognition and call it improved extended sparse representation classifier and feature descriptor (IESRCFD) method. Experimental results show that IESRCFD outperforms many state-of-the-art methods.

Keywords: Image recognition · LWS feature descriptor
Extended sparse weighted representation classifier

1 Introduction

Face recognition is a hot topic in the fields of computer vision and pattern recognition, and many researchers have been working on it [1]. Various techniques have been used for face recognition such as principal component analysis (PCA) [2], linear discriminant analysis (LDA) [3], local preserving projections (LPP) [4] and Eigenface [5] during the past few years. However, these early methods can only deal with simple face recognition problems well. Extreme learning machine (ELM) [6] is a single hidden layer network which can solve face recognition problems very quickly. But its recognition rate is limited. Although sparse representation classifier (SRC) [7] performs better than methods like PCA and LDA, it has no robustness for large continuous occlusion. Yang et al. [8] proposed a regularized robust coding (RRC) model to improve the robustness of SRC, which could regress a given signal with regularized regression coefficients. By assuming that the coding residual and the coding coefficient are respectively independent and identically distributed, RRC seeks for a maximum

posterior solution of the coding problem. However, it can only mitigate the occlusion in face recognition to some extent.

In face recognition, the intra-class difference which is caused by variable expressions, illuminations, and disguises, can be shared across different subjects [9]. Based on this idea, Deng et al. [9] proposed the extended sparse representation-based classifier (ESRC). ESRC uses the intra-class variant dictionary and training samples to represent the test samples, which improves the recognition performance for occlusion or non-occlusion face images. In ESRC, a test sample is represented by several training samples that come from the same class. For training samples within a same class, ESRC holds that they have equal importance when representing other samples together. Because different samples even from the same class holds different amount of information or content, which suggests that they should be considered distinctively.

The statistical properties of features such as covariance matrix and distribution characteristics can also be regarded as a new feature and can be used for face recognition. Because they can reflect the characteristics of the original feature. Logarithmic image processing is a mathematical framework based on abstract linear mathematics [10]. It replaces the linear arithmetic operations with a non-linear one which more accurately characterizes the response of human eyes [11]. Texture feature is a kind of global feature. Discrete orthonormal S-transform (DOST) [12] can measure it. Hence, the coefficient of DOST can be view as a global feature. Furthermore, DOST can preserve the phase information. Gabor feature is a local feature. It has a good spatial locality and directional selectivity. Hence, DOST and Gabor feature can be used to construct the discriminative features.

In this paper, we propose an improved extended sparse representation classifier and feature descriptor (IESRCFD) method. In IESRCFD, we first define a logarithmic weighted sum (LWS) feature descriptor which combines advantages of global feature, local feature, statistical properties of features and logarithmic operation. Next, we estimate the atom representativeness using the proposed representativeness estimator. Then we propose an extended sparse weighted representation classifier (ESWRC) by using the representativeness. In ESWRC, the representativeness is used to reflect the importance of the atom when representing other samples. The atom representativeness is incorporated into the sparse representation process as a weight coefficient. Finally, IESRCFD is obtained by combining ESWRC and LWS feature descriptor.

The main contributions of this paper are as follows.

(1) Defined a logarithmic weighted sum (LWS) feature descriptor which describes images in a more accurate way. It has combined the advantages of discrete orthonormal S-transform feature, Gabor feature, covariance and logarithmic operation. Gabor feature plays the most important role in our algorithm among them.
(2) Proposed ESWRC considering the importance of each atom in representing the query samples by assigning a weight to each atom, which improves the recognition rate.
(3) Proposed to use ESWRC and LWS together as IESRCFD to achieve a very high recognition rate.

The rest of paper is organized as follows. Section 2 introduces logarithmic weighted sum feature descriptor. Section 3 introduces the proposed representativeness estimator. Section 4 is the proposed extended sparse weighted representation classifier. Section 5 is the proposed improved extended sparse representation classifier and feature descriptor (IESRCFD) method. Section 6 gives the experimental results. Section 7 is the conclusion.

2   Logarithmic Weighted Sum Feature Descriptor

In this paper, we propose a logarithmic weighted sum (LWS) feature descriptor which combines those advantages of discrete orthonormal S-transform (DOST) feature, Gabor feature, covariance and logarithmic operation. Figure 1 is the construction process of LWS feature descriptor. In Fig. 1, w_1 and w_2 are the weights.
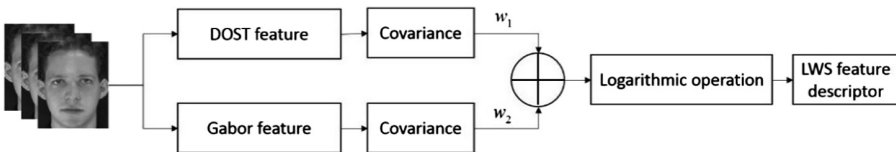

Fig. 1. The construction process of LWS feature descriptor.

2.1   Discrete Orthonormal S-Transform (DOST) Feature

Recent work shows that a discrete orthonormal S-transform (DOST) basis can be used to accelerate the calculation of S-transform (ST) [13] and eliminate the redundancy in the space-frequency domain [14]. Besides, DOST can preserve the phase information and allows for an arbitrary partitioning of the frequency domain, which achieves the zero redundancy. Those advantages make DOST has a wide range of applications, such as texture classification [12] and signal analysis [14].

DOST is a pared-down version of the fully redundant ST [14]. The main idea of DOST is to form N orthogonal unit-length basis vectors. Each base vector corresponds to a specific region in the time-frequency domain. And each specific region can be determined by the following three parameters: v, γ and τ. v is the center of each frequency domain band, γ is the width of that band, τ is the location in time domain. Thus the λ th basis vector $D[\lambda]_{[v,\gamma,\tau]}$ can be defined as

$$D[\lambda]_{[v,\gamma,\tau]} = \begin{cases} ie^{-i\pi\tau} \dfrac{e^{-i2\alpha(v-\gamma/2-1/2)} - e^{-i2\alpha(v+\gamma/2-1/2)}}{2\sqrt{\gamma}\sin\alpha}, & \alpha \neq 0 \\ -\sqrt{\gamma}ie^{-i\pi\tau}, & \alpha = 0 \end{cases} \tag{1}$$

where α equals to $\pi(\lambda/N - \tau/\gamma)$, and it represents the center of temporal window. $\lambda = 0, 1, \ldots, N-1$.

For any one dimensional signal $h[\lambda]$, its length is N. Its DOST coefficients $S_{[v,\gamma,\tau]}$ for the region corresponding to the choice of $[v, \gamma, \tau]$ can be obtained by formula (2).

$$S_{[v,\gamma,\tau]} = \frac{1}{\sqrt{\gamma}} \sum_{f=v-\gamma/2}^{v+\gamma/2-1} e^{-i\pi\tau} e^{i2\pi\frac{\tau}{\gamma}f} \left[\sum_{\lambda=0}^{N-1} e^{-i2\pi\frac{\lambda}{N}f} h[\lambda] \right] \tag{2}$$

where the content in the square brackets is $H[f]$.

For each input image, the DOST coefficient matrix can be obtained by using discrete orthonormal S-transform. The DOST coefficient matrix has the same size of the input image.

2.2 Gabor Wavelet Feature

Two dimensional Gabor wavelet transform is a kind of wavelet transform. It has a good time-frequency localization characteristic. Besides, the function of two dimensional Gabor wavelet is similar to enhancing bottom image features including edge and peak, as well as local features. The extracted Gabor feature not only has a good spatial locality and directional selectivity, but also is robust to illumination and pose variations. Hence, in this paper, we choose the two dimensional Gabor wavelet to extract the local feature of an image.

For each pixel in an image, a vector $F_{x,y}$ can be obtained by

$$F_{x,y} = [I(x,y), x, y, |G_{0,0}(x,y)|, \cdots, |G_{0,7}(x,y)|, |G_{1,0}(x,y)|, \cdots, |G_{4,7}(x,y)|] \tag{3}$$

where $I(x,y)$ represents the intensity value of position (x,y), $G_{\varphi,\sigma}(x,y)$ is the response of a two dimensional Gabor wavelet centered at (x,y) with orientation φ and scale σ:

$$G_{\varphi,\sigma}(x,y) = \frac{k_\sigma^2}{4\pi^2} \sum_{t,s} e^{-\frac{k_\sigma^2}{8\pi^2}((x-s)^2 + (y-t)^2)} \left(e^{ik_\sigma((x-t)cos(\theta_\varphi) + (y-s)sin(\theta_\varphi))} - e^{-2\pi^2} \right) \tag{4}$$

where $k_\sigma = \frac{1}{\sqrt{2^{\sigma-1}}}$, $\theta_\varphi = \frac{\pi\varphi}{8}$. Let $\varphi = 8$, $\sigma = 5$.

For an image Q, we assume that the row number of Q is $\hat{A}$, the column number of Q is $\hat{B}$. Then, its Gabor wavelet feature $\mathbf{O}$ is denoted by

$$\mathbf{O} = [F_{1,1}^T, \cdots, F_{1,\hat{A}}^T, F_{2,1}^T \cdots, F_{2,\hat{A}}^T, \cdots, F_{\hat{B},1}^T, \cdots, F_{\hat{B},\hat{A}}^T]^T \tag{5}$$

2.3 LWS Feature Descriptor

For image Q, its Gabor wavelet feature is $\mathbf{O}$. The covariance matrix of $\mathbf{O}$ is denoted by C_L. Meanwhile, image $\hat{Q}$ with the size of 43×43 is obtained by down-sampling the image Q. The DOST coefficient matrix $\tilde{\mathbf{O}}$ of image $\hat{Q}$ is obtained by formula (2). The covariance matrix of $\tilde{\mathbf{O}}$ is denoted by C_G. Taking image Q as an example, we define a new logarithmic weighted sum (LWS) feature descriptor.

$$FD = \log[\tilde{\lambda} \cdot C_G \ + \ (1-\tilde{\lambda}) \cdot C_L] \tag{6}$$

where FD is the LWS feature descriptor of Q, $\log[\cdot]$ represents the logarithmic operation, $\tilde{\lambda}$ is the weight coefficient, $0 \leq \tilde{\lambda} \leq 1$. The LWS feature descriptors of other images in data set are obtained in this way.

3   Representativeness Estimator

In methods based on sparse representation, a test sample is represented by several atoms. For different atoms, because they contain different amounts of information, they have different representativeness to represent other samples generally. In this paper, we propose a representativeness estimator, and use it to estimate the atom representativeness.

The main steps to estimate the representativeness by using representativeness estimator are as follows.

Firstly, for any sample (e.g., image), we remove its relativity by whitening. Then its representativeness is estimated by computing its information entropy. This process is as follows.

Given an arbitrary sample (e.g., image) S, its size is $M \times N$.

3.1   Remove the Relativity of Sample

(1) Remove the mean of S.

$$X = S - \bar{S} \tag{7}$$

where $\bar{S}$ is the mean of S.
(2) X is arranged into a column vector.
(3) Compute the covariance matrix $\mathbf{\Pi}$.

$$\mathbf{\Pi} = \mathrm{E}[XX^T] \tag{8}$$

where $\mathrm{E}[\ \cdot\ ]$ represents the mathematical expectation.
(4) Singular value decomposition.

$$\mathbf{\Pi} = \mathbf{U} * \mathbf{\Lambda} * \mathbf{U}^T \tag{9}$$

where $\mathbf{\Lambda}$ is a diagonal matrix.
(5) Compute the whitening matrix $\tilde{M}$.

$$\tilde{M} = \mathbf{\Lambda}^{-1/2} \mathbf{U}^T \tag{10}$$

(6) Compute the whitened matrix Z.

$$Z = \tilde{\mathrm{M}} * X \tag{11}$$

3.2 Compute the Information Entropy

(1) **Z** is arranged into a matrix $\tilde{\mathbf{Z}}$, and the size of $\tilde{\mathbf{Z}}$ is $M \times N$.
(2) Compute the information entropy of $\tilde{\mathbf{Z}}$.

$$H_{\tilde{\mathbf{Z}}} = -\sum_{i=0}^{255} p_i \log p_i \tag{12}$$

where $H_{\tilde{\mathbf{Z}}}$ is the information entropy of $\tilde{\mathbf{Z}}$, p_i represents the proportion of pixel whose grayscale is i.

In this paper, $H_{\tilde{\mathbf{Z}}}$ represents the representativeness of image S, and the greater the value of $H_{\tilde{\mathbf{Z}}}$, the stronger its representativeness.

4 Extended Sparse Weighted Representation Classifier

In ESRC, for atoms with the same class, because they contain different amounts of information, they should have different representativeness to represent other samples. That is to say, their importance is not the same when they together represent other samples. In order to reflect their importance, we propose an extended sparse weighted representation classifier (ESWRC).

We assume that $\boldsymbol{\psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \boldsymbol{\psi}_3, \ldots, \boldsymbol{\psi}_k] \in \Re^{d \times n}$ are the training samples, where $\boldsymbol{\psi}_i \in \Re^{d \times n_i}$ are the training samples with class i. Each column of $\boldsymbol{\psi}_i$ represents a training sample with class i. That is to say, $\boldsymbol{\psi}_i$ contains n_i training samples.

Let $\mathbf{B} = [\mathbf{B}_1 - c_1 e_1, \ldots, \mathbf{B}_l - c_l e_l] \in \Re^{d \times m}$ represents the intra-class variant bases, where $e_i = [1, \ldots, 1] \in \Re^{1 \times m_i}$, $c_i \in \Re^{d \times 1}$ is the class centroid of class i, $\mathbf{B}_i \in \Re^{d \times m_i}$, $i = 1, 2, \ldots, l$, $\sum_{i=1}^{l} m_i = m$. $\mathbf{B}_i$ are randomly selected from $\boldsymbol{\psi}_i$, and they are used to obtain the intra-class variant bases. Each column of $\mathbf{B}_i$ represents a training sample with class i. That is to say, $\mathbf{B}_i$ contain m_i training samples.

The main idea of ESWRC can be illustrated by formula (13).

$$\eta = (\mathbf{W}_{\boldsymbol{\psi}} \otimes \boldsymbol{\psi}) x + (\mathbf{W}_{\mathbf{B}} \otimes \mathbf{B}) \beta + z \tag{13}$$

where η is a test sample. $\mathbf{W}_{\boldsymbol{\psi}}$ and $\mathbf{W}_{\mathbf{B}}$ are the weight coefficients. They correspond to $\boldsymbol{\psi}$ and $\mathbf{B}$ respectively. x and β are the sparse vectors. z is a noise term with bounded energy $\|z\|_2 < \varepsilon$, and $z \in \Re^d$. '$\otimes$' is the Hadamard product operator. Figure 2 is the diagram of ESWRC.

$$\mathbf{W}_{\boldsymbol{\psi}} = [\mathbf{W}_{\boldsymbol{\psi}_1}, \mathbf{W}_{\boldsymbol{\psi}_2}, \mathbf{W}_{\boldsymbol{\psi}_3}, \ldots, \mathbf{W}_{\boldsymbol{\psi}_k}] \tag{14}$$

$$\mathbf{W}_{\mathbf{B}} = [\mathbf{W}_{\mathbf{B}_1}, \mathbf{W}_{\mathbf{B}_2}, \mathbf{W}_{\mathbf{B}_3}, \ldots, \mathbf{W}_{\mathbf{B}_l}] \tag{15}$$
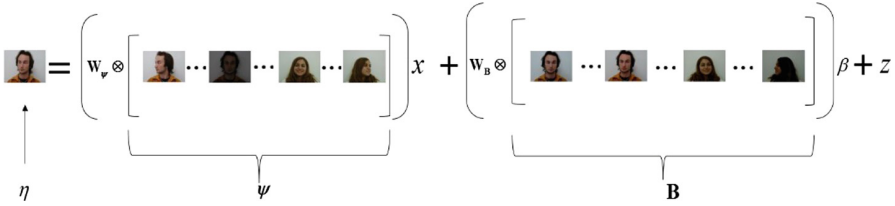
Fig. 2. The diagram of ESWRC.

where $\mathbf{W}_{\psi_i} \in \Re^{1 \times n_i}$ corresponds to ψ_i, $i = 1, 2, \ldots, k$. $\mathbf{W}_{\mathbf{B}_j} \in \Re^{1 \times m_j}$ corresponds to $\mathbf{B}_j - c_j e_j$, $j = 1, 2, \ldots, l$.

For all samples with class i, the expression of $\mathbf{W}_{\psi_i}$ is as follows.

$$\mathbf{W}_{\psi_i} = [\Gamma_{i1}, \Gamma_{i2}, \Gamma_{i3}, \ldots, \Gamma_{in_i}] \tag{16}$$

where Γ_{ij} represents the contribution factor of jth image in class i.

$$\Gamma_{ij} = \frac{H_{ij}}{\sum_{j=1}^{n_i} H_{ij}} \tag{17}$$

where H_{ij} represents the information entropy of jth image in class i.

Let $(\mathbf{W}_\psi \otimes \psi) = \psi^\bullet = [\psi_1^\bullet, \psi_2^\bullet, \psi_3^\bullet, \ldots, \psi_k^\bullet] \in \Re^{d \times n}$, where $\psi_i^\bullet$ represent those new samples with class i. Each column of $\psi_i^\bullet$ represents a new sample. $(\mathbf{W}_\mathbf{B} \otimes \mathbf{B}) = \mathbf{B}^\bullet = [\mathbf{B}_1^\bullet - c_1^\bullet e_1, \mathbf{B}_2^\bullet - c_2^\bullet e_2, \ldots, \mathbf{B}_l^\bullet - c_l^\bullet e_l]$, where $e_i = [1, \ldots 1] \in \Re^{1 \times m_i}$, $c_i^\bullet \in \Re^{1 \times m_i}$ is the class centroid of class i, $\mathbf{B}_i^\bullet \in \Re^{1 \times m_i}$, $i = 1, 2, \ldots, l$, $\sum_{i=1}^{l} m_i = m$. $\mathbf{B}_i^\bullet$ are the new training samples, and they are used to obtain the intra-class variant bases. Each column of $\mathbf{B}_i^\bullet$ represents a new training sample with class i. That is to say, $\mathbf{B}_i^\bullet$ contain m_i new training samples.

Firstly, the dimensions of $\psi^\bullet, \mathbf{B}^\bullet$ and η are reduced by PCA. Then $\tilde{\psi} = [\tilde{\psi}_1, \tilde{\psi}_2, \tilde{\psi}_3, \ldots, \tilde{\psi}_k] \in \Re^{\tilde{d} \times n}$, $\tilde{\mathbf{B}} = [\tilde{\mathbf{B}}_1 - \tilde{c}_1 e_1, \ldots, \tilde{\mathbf{B}}_l - \tilde{c}_l e_l] \in \Re^{\tilde{d} \times m}$ and $\tilde{b} \in \Re^{\tilde{d} \times 1}$ are obtained, where $\tilde{\psi}_i \in \Re^{\tilde{d} \times n_i}$, $\tilde{\mathbf{B}}_i \in \Re^{\tilde{d} \times m_i}$ and $\tilde{c}_i \in \Re^{\tilde{d} \times 1}$ correspond to ψ_i, $\mathbf{B}_i$ and c_i respectively.

Let $\tilde{A} = [\tilde{\psi}, \tilde{\mathbf{B}}]$, $\tilde{x} = \begin{bmatrix} x \\ \beta \end{bmatrix}$, $G(\tilde{x}) = \|\tilde{x}\|_1$, $H(\tilde{x}) = \tilde{b} - \tilde{A}\tilde{x}$.

We have the following sparse representation problem.

$$\min_{\tilde{x}} \|\tilde{x}\|_1 \quad \text{sub.to } \tilde{b} = \tilde{A}\tilde{x} \tag{18}$$

where $\tilde{b}$ is the test sample that needs to be represented, $\tilde{A}$ is the sparse dictionary matrix.

We use the augmented Lagrangian method to solve (18), then obtain the solution of (18), and denoted by $\tilde{x}^* = \begin{bmatrix} \breve{x} \\ \breve{\beta} \end{bmatrix}$. After that, the residual is computed by

$$R_i = (\tilde{b}) = \left\| \tilde{b} - \tilde{A} \begin{bmatrix} \delta_i(\breve{x}) \\ \breve{\beta} \end{bmatrix} \right\|_2 \tag{19}$$

where $i = 1, 2, \ldots, k$, $\delta_i(\breve{x})$ is a new vector whose only nonzero entries are the entries in $\breve{x}$ those are associated with class i.

The category of probe sample $\tilde{b}$ can be obtained by

$$\text{Identity } (\tilde{b}) = \arg \min_i R_i(\tilde{b}) \tag{20}$$

Hence, the label of η is Identity $(\tilde{b})$.
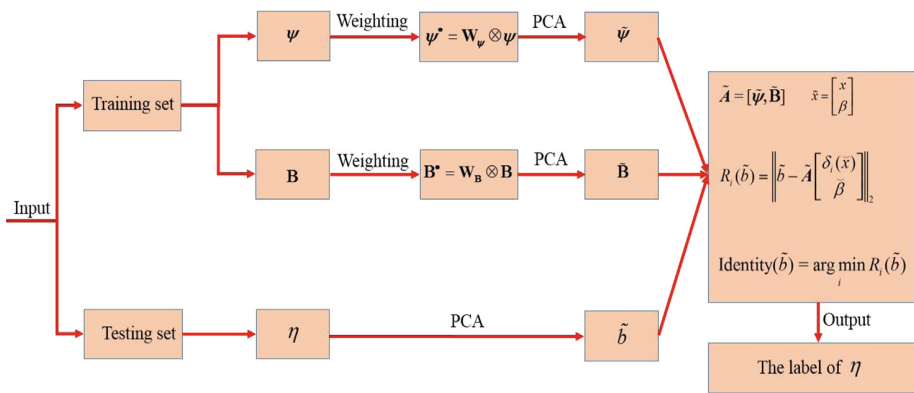Figure 3 is the flow chart of EWSRC.

Fig. 3. The flow chart of EWSRC.

5   Improved Extended Sparse Representation Classifier and Feature Descriptor Method

On the one hand, the defined LWS feature descriptor is robust to illumination variations and pose variations. On the other hand, EWSRC has a strong robustness to occlusion and non-occlusion. In order to combine the advantages of LWS feature descriptor and EWSRC, we propose an improved extended sparse representation classifier and feature descriptor (IESRCFD) method. IESRCFD is as follows. First of all, LWS feature descriptor of each image is obtained by formula (9). Then the obtained LWS feature descriptors are used as the input of ESWRC. Finally, the labels of testing samples are obtained. The time complexity of IESRCFD is about $O(n^2)$. Figure 4 is the flow chart of IESRCFD.
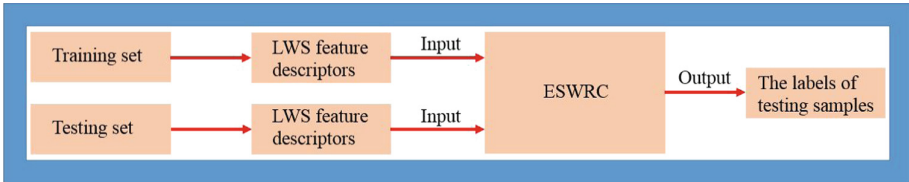

Fig. 4. The flow chart of IESRCFD.

6   Experimental Results and Analysis

Our IESRCFD is compared with other algorithms on FEI face database and FERET database respectively.

6.1   FEI Face Database

The FEI face database is a Brazilian face database, and it contains 200 different persons. Each person has 14 different images. All images are colorful and taken against a white homogenous background in an upright frontal position with profile rotation of up to about 180°. Scale might vary about 10% and the original size of each image is 640×480 pixels. In our experiments, the whole dataset is used for our experiments. The training set and test set both account for half of the whole dataset. That is to say, for each person, we randomly select seven images as the training set, and the rest seven images are used as the test set. The size of each image is down sampled to 64×64. Figure 5 shows some examples of image variations from the FEI face database.

Fig. 5. Some examples of image variations from the FEI face database.

Table 1 lists the recognition rates of different algorithms on FEI face database.

Table 1. Recognition rates of different algorithms on FEI face database ($\tilde{\lambda} = 0.1$)

Algorithm	Recognition rate
RRC [8]	56.4%
ESRC [9]	79.1%
SRC [7]	79.1%
KNN [15]	66.1%
NN [16]	66.1%
ELM [6]	61.5%
ELMSRC [17]	80.7%
RSC [18]	71.3%
KSRC [19]	77.9%
IESRCFD	**91.6%**

From Table 1 we can see that the recognition rate of our IESRCFD is 91.6%, which achieves the highest recognition rate, with about 35%, 12%, 12%, 25%, 30%, 11%, 20% and 14% improvements over RRC, ESRC, SRC, KNN, ELM, ELMSRC, RSC and KSRC respectively.

6.2   FERET Database

As in [20, 21], the "b" subset of FERET database is used to verify the recognition performance of various algorithms. The "b" subset of FERET consists of 198 subjects, each subject contains 7 different images. The training set images contain frontal face images with neutral expression "ba", smiling expression "bj", and illumination changes "bk". While the test set images involve face images of varying pose angle: "bd"-$+25°$, "be"-$+15°$, "bf"-$-15°$, and "bg"-$-25°$. As in [20], all images are down sampled to 64×64. Some samples of face images in "b" subset of FERET database are shown in Fig. 6.

| bd | be | bf | bg | ba | bj | bk |

Fig. 6. Some samples of face images in "b" subset of FERET database.

Table 2 lists the recognition rates of different algorithms on the "b" subset of FERET database.

In Table 2 one see that when the test set are "bd", "be", "bf" and "bg" subsets, the recognition rates of our IESRCFD are 84.5%, 99.0%, 99.5% and 90.5% respectively, which are higher than those of other algorithms. The average recognition rate of IESRCFD is 93.4%, which not only achieves the highest recognition rate, but also shows a 7%, 30%, 37%, 42%, 33%, 39%, 31% and 35% better performance compared with that of GSRC, LogE-SR, TSC, ESRC, KNN, ELM, ELMSRC and KSRC respectively.

Table 2. Recognition rates of different algorithms on "b" subset of FERET database ($\tilde{\lambda} = 0.1$)

Algorithm\Dataset	bd	be	bf	bg	Average
RSR [20]	79.5%	96.5%	97.5%	86.0%	89.9%
GSRC [22]	77.0%	93.5%	97.0%	79.0%	86.6%
LogE-SR [23]	34.5%	81.0%	91.0%	46.5%	63.3%
TSC [21]	36.0%	73.0%	73.5%	44.5%	56.8%
ESRC [9]	38.5%	59.0%	68.0%	41.0%	51.6%
KNN [15]	56.5%	60.5%	74.0%	48.0%	59.8%
NN [16]	56.5%	60.5%	74.0%	48.0%	59.8%
KSRC [19]	50.5%	67.0%	60.0%	55.5%	58.3%
ELM [6]	45.5%	58.5%	61.0%	53.5%	54.6%
ELMSRC [17]	56.5%	63.0%	65.0%	65.5%	62.5%
IESRCFD	**84.5%**	**99.0%**	**99.5%**	**90.5%**	**93.4%**

7   Conclusion

In this paper, we propose an improved extended sparse representation classifier and feature descriptor (IESRCFD) method by using proposed LWS feature descriptor and ESWRC simultaneously. Experimental results showed that our IESRCFD performs better than many algorithms. The recognition rate of IESRCFD is about 20%, 14%, 12% and 11% higher than those of RSC, KSRC, ESRC and ELMSRC respectively on FEI face database. The recognition rate of IESRCFD is about 31%, 7% and 4% higher than those of ELMSRC, GSRC and RSR respectively on FERET face database. Meanwhile, experimental results illustrated that IESRCFD could convergence and meet the real time requirement.

References

1. Huang, Z.W., Shan, S.G., Wang, R.P., Zhang, H.H., Lao, S.H., Kuerban, A., Chen, X.L.: A benchmark and comparative study of video-based face recognition on COX face database. IEEE Trans. Image Process. **24**, 5967–5981 (2015)
2. Adbi, H., Williams, L.J.: Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat. **2**, 433–459 (2010)
3. Etemad, K., Chellapa, R.: Discriminant analysis for recognition of human face images. J. Opt. Soc. Am. A **14**, 125–142 (1997)
4. He, X.F., Niyogi, P.: Locality preserving projections. In: Advances in Neural Information Processing Systems, Canada (2003)
5. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cognit. Neurosci. **3**, 71–86 (1991)
6. Huang, G.B., Ding, X., Zhou, H.: Optimization method based extreme learning machine for classification. Neurocomputing **74**, 155–163 (2010)
7. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 210–227 (2009)
8. Yang, M., Zhang, L., Yang, J., Zhang, D.: Regularized robust coding for face recognition. IEEE Trans. Image Process. **22**, 1753–1766 (2013)
9. Deng, W.H., Hu, J.N., Guo, J.: Extended SRC: undersampled face recognition via intraclass variant dictionary. IEEE Trans. Pattern Anal. Mach. Intell. **34**, 1864–1870 (2012)
10. Jourlin, M., Pinoli, J.C.: Logarithmic image processing: the mathematical and physical framework for the representation and processing of transmitted images. Adv. Imaging Electron. Phys. **115**, 130–196 (2001)
11. Mandal, D.: Human visual system based object detection and recognition and introduction of logarithmic local binary patterns for face recognition. M.S. thesis, TUFTS Univ (2012)
12. Drabycz, S., Stockwell, R.G., Mitchell, J.R.: Image texture characterization using the discrete orthonormal s-transform. J. Digit. Imaging. **22**, 696–708 (2009)
13. Gibson, P.C., Lamoureux, M.P., Margrave, G.F.: Letter to the editor: stokwell and wavelet transforms. J. Fourier Anal. Appl. **12**, 713–721 (2006)
14. Stockwell, R.G.: A basis for efficient representation of the S-transform. Digit. Signal Proc. **17**, 371–393 (2007)
15. Weinberger, K.Q., Saul, L.K.: Distance metric learning for large margin nearest neighbor classification. J. Mach. Learn. Res. **10**, 207–244 (2009)
16. Boiman, O., Shechtman, Irani, M.: In defense of nearest-neighbor based image recognition. In: Proceedings of IEEE Conference Computer Vision Pattern Recognition, pp. 1–8 (2008)
17. Luo, M.X., Zhang, K.: A hybrid approach combining extreme learning machine and sparse representation for image recognition. Eng. Appl. Artif. Intell. **27**, 228–235 (2014)
18. Yang, M., Zhang, L., Yang, J., Zhang, D.: Robust sparse coding for face recognition. In: Proceedings of IEEE Conference Compute Vision Pattern Recognition, pp. 625–632 (2011)
19. Gao, S.H., Tsang, I.W.H., Chia, L.T.: Sparse representation with kernels. IEEE Trans. Image Process. **22**, 423–434 (2012)
20. Harandi, M.T., Sanderson, C., Hartley, R., Lovell, B.C.: Sparse coding and dictionary learning for symmetric positive definite matrices: a kernel approach. In: Proceedings of IEEE European Conference on Computer Vision, pp. 216–229 (2012)

21. Sivalingam, R., Boley, D., Morellas, V., Papanikolopoulos, N.: Tensor sparse coding for region covariance. In: Proceedings of IEEE European Conference on Computer Vision, pp. 722–735 (2010)
22. Yang, M., Zhang, L.: Gabor feature based sparse representation for face recognition with Gabor occlusion dictionary. In: Proceedings of IEEE European Conference on Computer Vision, pp. 448–461 (2010)
23. Guo, K., Ishwar, P., Konrad, J.: Action recognition using sparse representation on covariance manifolds of optical flow. In: Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 188–195 (2010)