Robust Manifold Embedding for Face Recognition

ZHONGHUA LIU [1], **LINGYUN XIANG** [2], **(Member, IEEE), KAIMING SHI** [1],
KAIBING ZHANG [3], **AND QINGTAO WU** [1]

[1] School of Information Engineering, Henan University of Science and Technology, Luoyang, China
[2] Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, School of Computer and Communication Engineering, Hunan Key Laboratory of Smart Roadway and Cooperative Vehicle-Infrastructure Systems, Changsha University of Science and Technology, Changsha, China
[3] College of Electronics and Information, Xi'an Polytechnic University, Xi'an, China

Corresponding author: Lingyun Xiang (xiangly210@163.com)

ABSTRACT Flexible manifold embedding (FME) has been recognized as an effective method for face recognition by integrating both class label information from labeled data and manifold structure information of all data. In order to achieve better performance, this particular method usually requires sufficient samples to make manifold smooth. However, it is often hard to provide enough samples for FME in practice. In view of facial symmetry, we utilize left/right mirror face images to address the deficiency of samples in manifold embedding. These mirror images enable to reflect variations of illuminations, or poses or both them that the original face images cannot provide. Therefore, we propose a robust manifold embedding (RME) algorithm in this paper, which can fully use the class label information and correctly capture the underlying manifold structure. The proposed RME algorithm integrates two complementary characteristics of the label fitness and the manifold smoothness. Moreover, the original face images and its left/right mirror images are jointly used in the learning of RME, which shows better robustness against the variations of both illuminations and poses. Extensive experiments on several public face databases demonstrate that the proposed RME algorithm is promising for higher recognition accuracy than other compared methods in reference.

INDEX TERMS Manifold embedding, mirror image, robust manifold embedding (RME), face recognition.

I. INTRODUCTION

Dimension reduction is a hot topic for the recognition tasks of high dimensional image. In past decades, a large number of dimension reduction algorithms have been proposed [1]–[7]. Linear discriminant analysis (LDA) and principal component analysis (PCA) are the two most classical algorithms for dimension reduction. As a supervised method, LDA uses the class label information from training samples and solves the classification problem by simultaneously making the between-class scatter matrix (S_W) maximizing and the within-class scatter matrix (S_B) minimizing in the expected low-dimensional feature space. By contrast, PCA is an unsupervised method and it finds the directions of maximum scatter for optimal reconstruction without using any class label information. Owing to the simplicity and high efficiency of LDA and PCA, lots of variants based on the two algorithms have been proposed [8]–[14]. Due to the application of class label information, the supervised methods offer better recognition performance than the unsupervised methods when there are enough labeled training samples available. In addition, with the development of deep learning in recent years, the face recognition methods based on neural network has attracted more and more attention [15]–[17]. However, although deep learning has made a great breakthrough in image recognition due to its strong learning ability, data-driven and adaptability, it has a large amount of computation, high hardware requirements, complex model design and poor interpretability.

Structurally, human faces are non-rigid, and the corresponding face images essentially show a complicated nonlinear structure. Therefore, the essential nonlinear structure

of face images cannot be well revealed by the traditional linear dimension reduction methods. Due to the effectiveness of manifold learning in nonlinear dimension reduction, it is widely used by capturing low-dimensional essential structures embedded in high-dimensional space [18]–[22]. In general, the most classical manifold learning methods include locally linear embedding (LLE), ISOMAP and Laplacian Eigenmap (LE). Despite their effectiveness, the weakness of these manifold learning methods is so-called out-of-sample problem. That is to say, these methods cannot obtain a projection matrix to map a new sample to the desired low-dimensional embedding space. To solve this issue, many researchers focused on many improved manifold learning methods. He *et al.* [23] proposed a locality preserving projection (LPP) algorithm for the description of the essential face manifold structure. Vural and Guillemot [24] presented a semi-supervised manifold learning method, in which an interpolation function is built to provide an out-of-sample extension for general supervised manifold learning methods. In order to utilize the local structure information of the data, Lu *et al.* [25] proposed a novel manifold linear regression framework. Based on this framework, manifold discriminant regression learning (MDRL) and robust manifold discriminant regression learning are respectively given. By using the findings in sparse coding theory, Raducanu and Dornaika [26] proposed a generalized out-of-sample extension solution. Shi *et al.* [27] presented a novel supervised multi-manifold learning method, in which the multi-manifold features of images are extracted. In manifold learning algorithms, the local neighborhood size is very important and sensitive for the algorithm performance. In order to determine the neighborhood size, Zhang *et al.* [28] presented an adaptive manifold learning framework. Based on graph manifold learning in high dimensional feature space, Wang *et al.* [29] presented an unsupervised feature selection method, in which the importance of features in the original space was computed by means of L1-regularized least square and the spectral regression analysis. Fang *et al.* [30] proposed a locality and similarity preserving embedding (LSPE) for the preservation of locality, similarity and the sparse reconstruction relationship. In order to eliminate effects of occlusions and illumination variations, Wang *et al.* [31] presented a manifold regularized local sparse representation (MRLSR) method. Zhang *et al.* [32] proposed a patch alignment framework to unify spectral analysis based dimensionality reduction algorithms, in which discriminative locality alignment (DLA) was used. DLA consists of two major stages: the part optimization and the whole alignment. Nevertheless, DLA is too much sensitive to the values of the parameters such as neighborhood size and dimensional size. Liu and Jin [33] presented an enhanced discriminative locality alignment (EDLA) algorithm, which simultaneously uses the local structure information and class label information, leading to better performance than DLA.

Recently, representation based classification methods have become a research hotspot in pattern recognition and computer vision. Wright *et al.* [34] introduced sparse representation to image classification and presented a sparse representation based classification (SRC) algorithm. In SRC, a given test sample is first represented as a linear combination of all the training samples, and then a minimum reconstruction error is used to achieve the final classification. In order to examine the classification capability of collaborative representation or L1-norm sparsity, Zhang *et al.* [35] elaborated on the mechanism of SRC and concluded that it is collaborative representation that contributes to the performance of SRC. Based on the above observation, a collaborative representation based classification (CRC) method was developed for image classification. An observed data includes lots of features. However, it is well known that each feature of the observed data has different contribution to the pattern representation and classification. Based on this fact, Yang *et al.* [36] proposed a relaxed collaborative representation (RCR) method. Xu *et al.* [37] presented a two-phase test sample sparse representation method (TPTSSR) which makes a coarse-to-fine classification decision for the face samples. Considering that all samples from an object lie in a linear subspace, a linear regression classification (LRC) algorithm was proposed in [38]. In order to effectively use label information and manifold structure of the observed data, Nie *et al.* [39] presented a unified manifold learning framework called flexible manifold embedding (FME) by employing a linear regression function to map a new sample to desired feature space.

In practical application scenario, the same face shows considerable changes in pose, illumination and expression, so the variations of pose, illumination and expression regarding the same person are almost always larger than image variations of face identity [40]. To perform feature extraction for better performance, one should prepare enough training samples varying in different poses, illuminations and expressions. However, in practice, it is difficult or burdensome to collect sufficient facial images with different poses, illuminations and expressions of each object as the training samples. Fortunately, both the facial structure and the facial expression are symmetrical [41]. This has motivated many studies on the application of symmetry to enhance the diversity of training samples. For example, Xu *et al.* [42] proposed to first generate a set of new samples based on the symmetry of the face and then both the original samples and the newly generated symmetrical training samples are used to perform face classification.

Inspired by the prior work proposed in [43] and the symmetry feature of human faces [44], we propose a novel robust manifold embedding (RME) algorithm. The key idea of the proposed RME algorithm is that both the manifold information and the symmetrical information of facial image are fully taken into consideration to boost up recognition performance. The advantages of RME are two-fold. Firstly, due to the use the mirror images of the original images, the proposed RME can get better recognition performance even though there are few training samples. Secondly, our method integrates two

complementary characteristics of the label fitness and the manifold smoothness.

The rest of this paper is organized as follows. Section 2 reviews the related works. Section 3 presents the proposed RME algorithm. Experimental results are presented in section 4. Finally, Section 5 concludes the paper.

II. BRIEF REVIEW OF RELATED WORKS

Suppose that there are n training samples for c classes and each class has n_1 training samples. In each class, the first n_2 samples are labeled and the rest are unlabeled. Let $X = [x_1, x_2, \cdots, x_u, x_{u+1}, \cdots, x_n] \in R^{m \times n}$ be the training set, where m is the dimensional size of the samples, $u = c \times n_2$ is the number of labeled data, and $n = c \times n_1$ is the total number of training samples. For any labeled sample x_i from i^{th} class, it should belong to y_i ($y_i = 1, 2, \cdots, c$) class. We denote the class label of all samples as a binary label matrix $Y \in R^{n \times c}$, and the ith row and j^{th} column element Y_{ij} in matrix Y is defined as

$$Y_{ij} = \begin{cases} 1, & if \ y_i = j \\ 0, & otherwise \end{cases} \tag{1}$$

Let $G = \{X, S\}$ be an undirected connected graph with nodes X and similarity matrix $S \in R^{n \times n}$. The similarity matrix S is a symmetric matrix, in which each element S_{ij} represents the similarity of a pair of samples [45]. The Laplacian matrix is denoted by $L = D - S$, and D is a diagonal matrix and its diagonal elements are denoted as $D_{ii} = \sum_j S_{ij}$.

A. GAUSSIAN FIELDS AND HARMONIC FUNCTIONS (GFHF)

Based on the label fitness and the manifold smoothness, GFHF is presented to predict class label matrix $F \in R^{n \times c}$. As shown in [43], the objective function $g_G(F)$ of GFHF is minimized as below:

$$g_G(F) = \frac{1}{2} \sum_{i,j=1}^{n} \|F_i - F_j\|_2^2 S_{ij} + \lambda_\infty \sum_{i=1}^{u} \|F_i - Y_i\|_2^2 \tag{2}$$

where Y_i and F_i are respectively the ith row of matrix Y and F, and Y_i is comprised of the class label of the ith labeled sample, F_i is comprised of the predicted class label of the ith unlabeled sample. In order to make $\sum_{i=1}^{u} \|F_i - Y_i\|_2^2 = 0$, or $F_i = Y_i$, λ_∞ is defined to a very large number [46]. Equation (2) can be rewritten as

$$trace(F^T L F) + trace((F - Y)^T U(F - Y)) \tag{3}$$

where L is a graph Laplacian matrix, and it is defined as $L = D - S$, D is a diagonal matrix and its diagonal elements are $D_{ii} = \sum_j S_{ij}$, U is a diagonal matrix, in which the first u diagonal elements are λ_∞, and the rest diagonal elements are 0, respectively.

B. FLEXIBLE MANIFOLD EMBEDDING (FME)

On the basis of keeping the label fitness and the manifold smoothness, GFHF predicts the class labels of unlabeled

samples. However, it can not get a projection matrix in GFHF algorithm. That is to say, GFHF can not map new samples to the desired subspace. By contrast to GFHF, the method proposed in [47] enables to convert new samples to feature subspace by linear Laplacian regularized least squares (LapRLS/L) method through the linear regression function as follows.

$$F_i = h(x_i) = W^T x_i + t \tag{4}$$

where $W \in R^{m \times c}$ is a transformation matrix and $t \in R^{c \times 1}$ is a translation vector. The objective function of LapRLS/L is to minimize the ridge regression errors and preserve the manifold smoothness simultaneously, which is defined as follows.

$$g_M(W, t) = \lambda \|W\|_2^2 + \mu trace(W^T X L X^T W) + \frac{1}{u} \sum_{i=1}^{u} \|W^T x_i + t - Y_i^T\|_2^2 \tag{5}$$

where λ and μ are two balance factors.

Equation (4) can be rewritten as

$$F = h(X) = X^T W + e_n t^T \tag{6}$$

where $e_n = \underbrace{[1, 1, \cdots, 1]}_{n}^T$ is a n-dimensional vector of all ones.

From the above definition, the obtained prediction labels F in LapRLS/L are restricted in the space spanned by all the training samples X. For a given new sample, its class label can be obtained by the projection matrix W. Nie *et al.* [39] suggested that Equation (6) may be overstrict to fit the data samples from a non-linear manifold. In order to solve this problem, flexible manifold embedding (FME) is presented. In FME, the constraint F_0 is added to represent the regression residue in Equation (6), i.e.,

$$F = X^T W + e_n t^T + F_0 = h(X) + F_0 \tag{7}$$

where regression residue $F_0 \in R^{m \times c}$ is used to evaluate the mismatch between F and $h(X)$. In order to obtain the optimal prediction label matrix F, the objective function of FME is to minimize the following function.

$$g_{FME}(W, F) = \frac{1}{2} \sum_{i,j=1}^{n} \|F_i - F_j\|_2^2 S_{ij} + \lambda_\infty \sum_{i=1}^{u} \|F_i - Y_i\|_2^2 + \lambda_1 \|X^T W - F\|_F^2 + \lambda_2 \|W\|_F^2 \tag{8}$$

where λ_1 and λ_2 are balance factors. Equation 8 can further be converted as

$$trace(F^T L F) + trace((F - Y)^T U(F - Y)) + \lambda_1 \|X^T W - F\|_F^2 + \lambda_2 \|W\|_F^2 \tag{9}$$

To seek the optimal solution to Equation (9), we can respectively fix W and F, and make the derivatives of the objective function in Equation (9) with respect to W and F

Algorithm 1 FME

Step 1. Set each element of matrix $F \in R^{n \times c}$ to 1

Step 2. Compute the optimal projection matrix W using Equation (10).

Step 3. Use Equation (11) to compute the optimal label matrix F.

---

equal to zero. The solutions of W and F can be represented as follows.

$$W = (\lambda_1 X X^T + \lambda_2 I)^{-1} \lambda_1 X F \qquad (10)$$

$$F = (L + U + \lambda_1 I)^{-1} (UY + \lambda_1 X^T W) \qquad (11)$$

Because the solution to W and F in Equation (9) is contained each other, its solution cannot directly be obtained. An iterative way is to iteratively find the solution of the objective function. The FME algorithm is outlined as follows.

III. ROBUST MANIFOLD EMBEDDING (RME)

The advantages of the FME are two-fold. First, it can use a linear regression function $h(X) = X^T W + e_n t^T$ to map the feature space of new samples to the expected class label space. Second, both the label information and manifold structure are integrated in FME for the improvement of recognition performance. However, the performance of this method heavily depends on enough training samples for each object when learning. We usually have not so many training samples for each object to apply to FME in reality. Especially for face recognition, there are often few training samples for each object, so it may encounter the problem of small sample size. In addition, face appearance can change drastically due to the variations in expression, illumination and pose. In particular, the variations between the images of the same face are nearly greater than the variations of face identity [48]–[51]. These above factors can bring adverse effect on the recognition performance of FME.

It is widely recognized that both facial structure and the facial expression are symmetric [41], [42]. Based upon this observation, we can use the mirror image of the face image to represent more possible illumination and pose changes which the original face image cannot provide. Inspired by this idea, we can generate the left-mirror-image and right-mirror-image for each face image by using the left/right half face image. By this way, we can obtain triple training samples for each face image. Although the method of generating mirror images is simple, the generated mirror images are beneficial to reflect more possible variations of the original face image in both illumination and pose. As a result, the proposed RME algorithm is robust to the variations of illumination and pose.

The objective function of the proposed RME algorithm is minimized as follows.

$$obj(W, F) = \frac{1}{2} \sum_{i,j=1}^{n} \|F_i - F_j\|_2^2 S_{ij} + \lambda_\infty \sum_{i=1}^{u} \|F_i - Y_i\|_2^2$$
$$+ \lambda_1 \left\| X^T W - F \right\|_F^2 + \lambda_2 \|W\|_F^2$$
$$+ \lambda_3 \left(\left\| Z_1^T W - F \right\|_F^2 + \left\| Z_2^T W - F \right\|_F^2 \right) \quad (12)$$

where Z_1 and Z_2 are the data matrices consisting of the left-mirror-images and the right-mirror-images corresponding to the original training samples from training set X, respectively. $\lambda_\infty, \lambda_1, \lambda_2$ and λ_3 are the balance factors. Equation (12) can further be reformulated as follows.

$$obj(W, F) = trace(F^T L F)$$
$$+ trace((F - Y)^T U(F - Y))$$
$$+ \lambda_1 \left\| X^T W - F \right\|_F^2 + \lambda_2 \|W\|_F^2$$
$$+ \lambda_3 \left\| Z_1^T W - F \right\|_F^2 + \lambda_3 \left\| Z_2^T W - F \right\|_F^2 \quad (13)$$

In order to get the optimal solution to Equation (13), we respectively make the derivatives of the objective function in Equation (13) with respect to W and F, and set them to be zero. The analytic solutions of W and F can be represented as follows.

$$W = (\lambda_1 X X^T + \lambda_2 I + \lambda_3 Z_1 Z_1^T + \lambda_3 Z_2 Z_2^T)^{-1}$$
$$\times (\lambda_1 X F + \lambda_3 Z_1 F + \lambda_3 Z_2 F) \quad (14)$$
$$F = (L + U + \lambda_1 I + 2\lambda_3 I)^{-1}$$
$$\times (UY + \lambda_1 X^T W + \lambda_3 Z_1^T W + \lambda_3 Z_2^T W) \quad (15)$$

In our algorithm, we assume that there are n samples for all objects, which includes u labeled samples (called the training samples) and $l = n - u$ unlabeled samples (called the test samples). The proposed RME algorithm is outlined as follows.

IV. EXPERIMENTAL RESULTS

In this section, we carry out extensive experiments to evaluate the performance of the proposed RME algorithm by comparing with other five state-of-art algorithms: MSEC [52], CRC [35], FME [39], SOSI [24], MDRL [25], the method presented in [42], GFHF [43], and NLDLSR [53]. In our experiments, six public face databases including the PIE, YaleB, FERET and GT face databases under large variations of illuminations and poses are used as the benchmark to show the promise of the proposed method.

A. PARAMETERS SELECTION

In our RME algorithm, there are four parameters: the balance factors λ_1, λ_2 and λ_3, and the number of nearest neighbors k. The average recognition rates of RME respectively vary with the parameters $\lambda_1, \lambda_2, \lambda_3$ and k in Fig. 1. From Fig. 1, we can find that the value of λ_1 should be more than 10, the value of λ_2 should be between 1 and 10, the value of λ_3 should be less than 4, and the value of k should be between 3 and 5.

B. EXPERIMENTS ON CMU PIE FACE DATABASE

The CMU PIE face database contains 41368 face images from 68 subjects as a whole. The face images were under variations in poses, illuminations, and expressions. In this paper, we fixed the pose and expression, and we got 21 images under different lighting conditions for each subject. The size
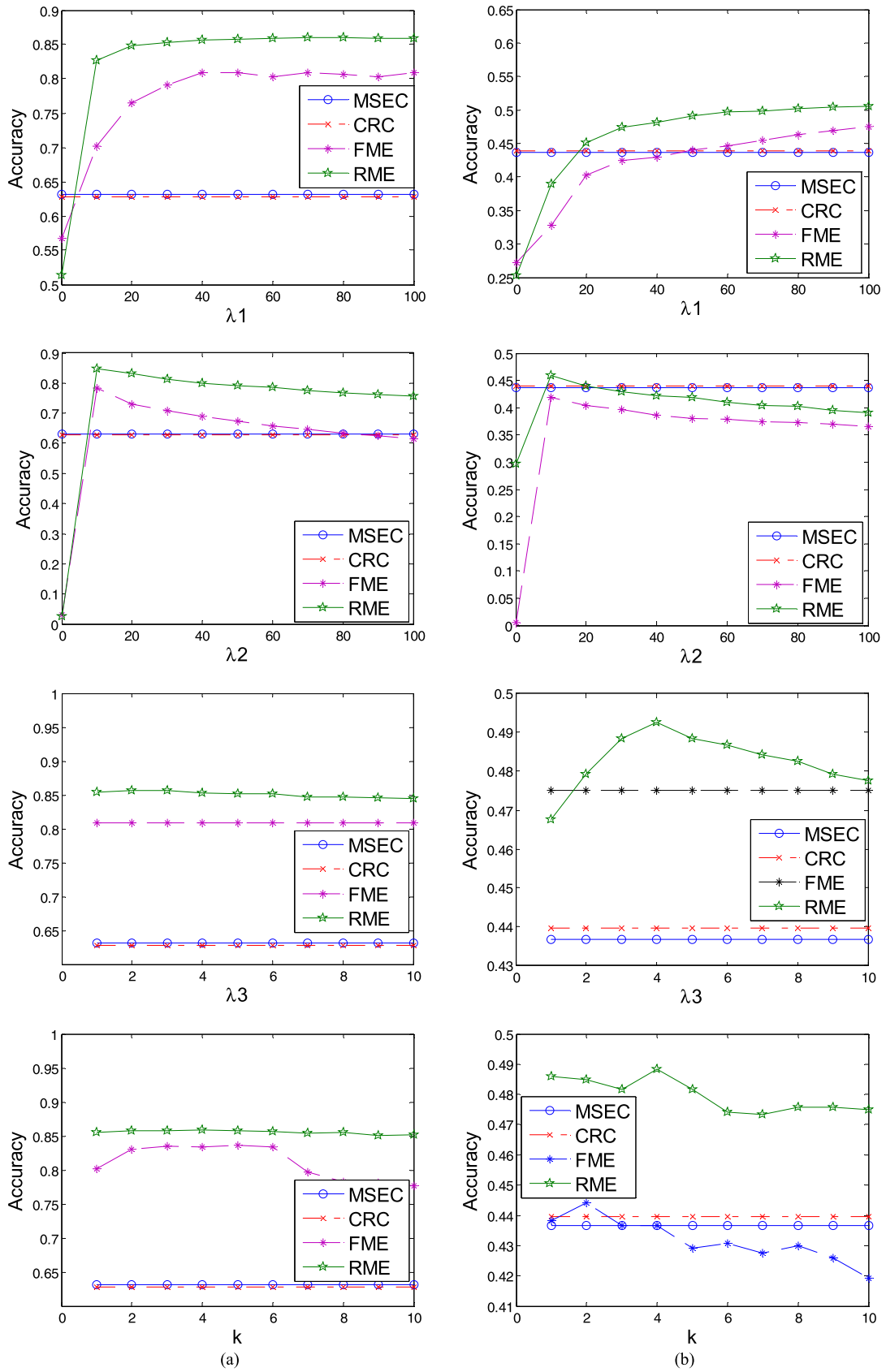
FIGURE 1. The performance of RME versus parameters λ_1, λ_2 λ_3, and k. (a) YaleB, (b) Feret.

Algorithm 2 RME Algorithm

Step 1. Generate the left-mirror and right-mirror images for each face image. Let $z \in R^{p \times q}$ be a given face image, and z_1 and z_2 denote the corresponding left and right mirror face images, respectively. Along the center of the face image, the left-half image z_L and right-right image z_R are firstly produced from the image z. Then the obtained two half face images are respectively mirrored to generate two other half images z'_L and z'_R. The obtained mirror images are represented as.

$$z_1 = [z_L, z'_L]$$
$$z_2 = [z'_R, z_R] \quad (16)$$

where $z_L = z(:, 1 : q/2)$ and $z_R = z(:, q/2 + 1 : end)$. The corresponding left mirror image z'_L and right mirror image z'_R can be gotten respectively using the following equations.

$$z'_L(:, (1 : q/2)) = z_L(:, q/2 + 1 - (1 : q/2)) \quad (17)$$
$$z'_R(:, (1 : q/2)) = z_R(:, q/2 + 1 - (1 : q/2)) \quad (18)$$

Step 2. Set $t = 0$ and all the elements in the matrix $F_0 \in R^{n \times c}$ are set to 1.

Step 3. Compute matrix W_t using Equation 14.
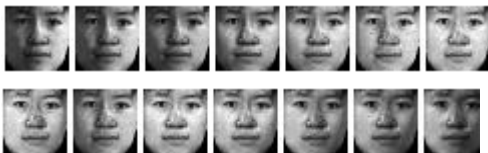
Step 4. Compute matrix F_t using Equation 15.

Step 5. Set $t = t + 1$.

Step 6. Checking convergence: if $|obj(W_{t+1}, F_{t+1}) - obj(W_t, F_t)| < 0.01$, stop the iteration procedure and output the optimal $F^* = F_t$ and $W^* = W_t$. Otherwise, go to Step 3.

Step 7. Obtain the class labels by $test\_Label = F(u + 1 : end, :)$.

Step 8. Use the nearest neighbor classifier for classification.

FIGURE 2. Some sample face images from the PIE database.

of each image in CMU PIE was 32×32 pixels. Fig. 2 shows some sample images of one person.

In the experiment, the first $l(l = 1, 2, 3)$ face images of each object are respectively selected for the training set, and the rest for the test set. In order to truly reflect the performance of each algorithm, we seek the optimal parameters of each algorithm by manual experiments. In the FME algorithm, we set $k = 10, \lambda_1 = 40$, and $\lambda_2 = 5$. In the RME algorithm, we set $k = 10, \lambda_1 = 40, \lambda_2 = 5$, and $\lambda_3 = 3$, respectively. The recognition performance of MSEC, CRC, FME, MDRL, SOSI, and the proposed RME is shown in Table 1.

From Table 1, we can see that the recognition performance of the proposed RME is significantly superior to that of

MSEC, CRC MDRL, and SOSI, especially when the number of training samples is 1 or 2. In addition, the recognition rates of both the FME and the proposed RME are 100% irrespective of the variation of the training sample sizes.

C. EXPERIMENTS ON THE EXTENDED YALEB FACE DATABASE

The Extended YaleB database contains images of 38 distinct persons, and each person has his/her frontal images taken under 45 different lighting directions. In our experiments, each image is resized to 48×42 pixels. Fig. 3 shows some sample images of one person.

For this experiment, we respectively select the first $l(l = 1, 2)$ images of each object as the training set, and the rest is taken as the test set. In FME algorithm, the optimal neighbor number k is set to 6, and the optimal balance coefficients λ_1 and λ_2 are set to 40 and 5, respectively. The optimal neighbor number k and balance coefficients λ_1, λ_2 and λ_3 are respectively set to 6, 40, 5, and 3 in the proposed RME algorithm. The recognition rates of the compared methods are shown in Table 2.

It can be seen from Table 2 that the proposed RME achieves the best recognition rate among the compared approaches. Its recognition rate is far higher than MSEC, CRC, FME, MDRL and SOSI. Note that FME and RME can obtain the same recognition performance on CMU PIE face database, and this is because that they are not sensitive to the variations of illumination. However, when the two methods are applied to Extended YaleB face database with large illumination variation, the recognition rates of our method is superior to that of the FME algorithm, which means that the proposed RME is more robust to lighting variations.

D. EXPERIMENTS ON FERET FACE DATABASE

The FERET database [54] contains a total of 13,539 face images of 1,565 subjects. The images vary in size, pose, illumination, facial expression, and age. We selected 1400 images of 200 individuals (each one has 7 images). Each image was resized to 40×40 pixels. Fig. 4 illustrates the sample images of one individual.

In this experiment, the first $l(l = 1, 2, 3, 4)$ face images are respectively taken as training set, and the rest as test set. The optimal neighbor number k and balance coefficients λ_1 and λ_2 are respectively set to 10, 100 and 5 for FME algorithm. For the proposed RME algorithm, the optimal neighbor number k and balance coefficients λ_1, λ_2 and λ_3 are respectively set to 10, 100, 5 and 3. Table 3 tabulates the recognition rates of each method.

Based on the results in Table 3, we can see that the proposed RME algorithm perform better than MSEC, CRC, FME, MDRL, and SOSI algorithms irrespective of the variation of training sample size. With one training sample for each object, the recognition rates of six compared algorithms are 43.67%, 43.97%, 47.50%, 47.12%, 46.53%, and 50.58%, respectively.

TABLE 1. Recognition rates of different methods on the PIE database (%).

Method \ Training sample size	1	2	3
MSEC	85.22	92.41	97.02
CRC	80.74	91.11	97.84
FME	100	100	100
MDRL	93.67	94.51	96.73
SOSI	95.31	96.71	98.50
GFHF	92.87	94.74	96.90
The method presented in [35]	98.57	99.14	99.49
NLDLSR	94.53	95.12	97.68
The proposed RME	**100**	**100**	**100**

FIGURE 3. Some sample face images from the YaleB database.

FIGURE 4. Sample face images from the FERET database.

E. EXPERIMENTS ON GEORGIA TECH FACE DATABASE

Georgia Tech face database contains images of 50 people taken in two or three sessions, which were built at Georgia Institute of Technology. For each individual in this database, 15 color JPEG images with cluttered background were taken at the resolution of 640×480 pixels. Faces illustrated in these images may be frontal and/or tilted with different expressions, illuminations, and scales. Each image was resized to 60×50 pixels. The images are converted to grayscale. Fig. 5 shows a set of sample images of one person.

In this experiment, we respectively select the first one to six images per individual as training set and the rest as test set. In FME algorithm, the optimal neighbor number k and balance coefficients λ_1, and λ_2 are respectively set to 5, 40 and 5 for FME algorithm. For the proposed RME algorithm, the optimal neighbor number k and balance coefficients λ_1, λ_2 and λ_3 are respectively set to 5, 40, 5 and 3. Table 4 presents the recognition results of compared methods.

From Table 4, we can see that the recognition rates of the proposed RME are much higher than that of MSEC, CRC,

TABLE 2. Recognition rates of different methods on the Extended YaleB database (%).

Method \ Training sample size	1	2
MSEC	63.14	66.26
CRC	62.78	74.04
FME	80.85	82.24
MDRL	80.97	83.62
SOSI	81.14	82.49
GFHF	80.62	83.46
The method presented in [35]	81.74	84.35
NLDLSR	82.51	84.89
The proposed RME	**85.64**	**87.94**

TABLE 3. Recognition rates of different methods on the FERET database (%).

Method \ Training sample size	1	2	3	4
MSEC	43.67	57.60	45.12	55.33
CRC	43.97	58.50	50.88	56.33
FME	47.50	57.33	54.25	68.33
MDRL	47.12	59.68	58.09	70.96
SOSI	46.53	58.49	56.35	69.36
GFHF	45.61	57.83	54.52	63.79
The method presented in [35]	47.12	59.64	61.18	70.42
NLDLSR	48.31	60.73	62.94	70.89
The proposed RME	**50.58**	**61.25**	**63.13**	**72.58**

TABLE 4. Recognition rates of different methods on the Georgia Tech database (%).

method \ Training sample size	1	2	3	4	5	6
MSEC	33.00	42.77	46.00	48.36	51.60	57.33
CRC	33.00	43.23	45.50	49.09	53.20	58.89
FME	38.00	48.15	51.67	54.18	57.6	65.78
MDRL	37.53	47.61	49.61	53.71	56.27	63.94
SOSI	36.28	45.39	48.83	52.19	54.83	62.95
GFHF	34.87	44.86	48.12	51.29	54.43	60.61
The method presented in [35]	39.32	46.94	49.68	51.69	54.89	62.11
NLDLSR	36.14	45.38	49.00	53.64	56.80	62.87
The proposed RME	**43.43**	**50.00**	**52.17**	**56.18**	**60.40**	**66.00**

FME, MDRL, and SOSI algorithms. Especially when there is only one training sample for each object, the proposed RME can obtain better recognition performance than other competitors. The results further indicate that our RME algorithm

TABLE 5. Recognition rates of different methods on UMIST database (%).

method ╲ Training sample size	1	2	3	4	5
MSEC	45.75	60.00	73.79	79.31	85.29
CRC	45.75	59.77	73.33	78.85	87.13
FME	47.16	62.99	73.33	78.62	83.91
MDRL	46.54	61.39	72.88	77.93	85.46
SOSI	45.71	60.46	72.11	76.84	84.68
GFHF	46.09	60.74	72.72	76.69	85.64
The method presented in [35]	46.12	61.58	73.83	79.31	86.23
NLDLSR	45.81	60.69	73.79	79.51	87.01
The proposed RME	**47.44**	**63.30**	**74.01**	**80.36**	**88.25**

TABLE 6. Recognition rates of different methods on Yale database (%).

method ╲ Training sample size	2	3	4	5
MSEC	92.59	95.83	96.61	97.12
CRC	59.77	73.33	78.85	87.13
FME	94.07	96.50	97.59	97.89
MDRL	93.17	95.12	97.62	96.12
SOSI	94.08	96.16	97.03	97.64
GFHF	84.44	82.50	89.52	91.01
The method presented in [35]	93.41	95.91	97.46	98.13
NLDLSR	88.80	93.48	95.33	97.76
The proposed RME	**94.13**	**97.18**	**98.23**	**99.01**

is not sensitive to the variation of illumination, expression, and pose.

F. EXPERIMENTS ON UMIST FACE DATABASE

The UMIST database consists of a total 575 face images of 20 people. The individuals are a mix of race, sex, and appearance and are photographed in a range of poses from profile to frontal views. The number of different views per subject varies from 19 to 48. The size of each face image is 56×46 pixels.

In this experiment, the $l(l = 1, 2, 3, 4, 5)$ face images of each person are respectively used to take as training set and the rest to be test set. All algorithms are repeated for 10 times. In FME algorithm, the optimal neighbor number k and balance coefficients λ_1, and λ_2 are respectively set to 5, 35 and 6 for FME algorithm. For the proposed RME algorithm, the optimal neighbor number and balance coefficients λ_1, λ_2 and λ_3 are respectively set to 5, 35,5 and 2. The recognition results are shown as Table 5. It can be known from Table 5 that our LRE achieves the best recognition performance irrespective of the variation of training sample size.
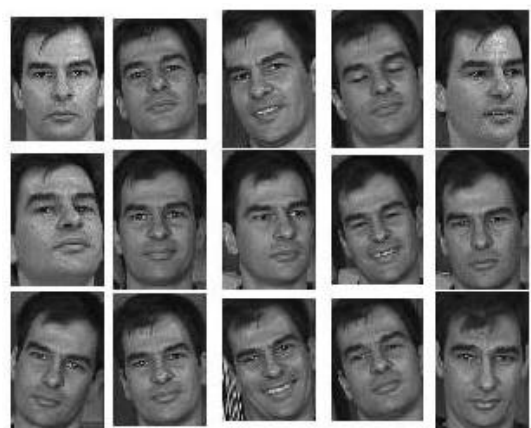
FIGURE 5. Sample face images from Georgia Tech face database.

G. EXPERIMENTS ON YALE FACE DATABASE

The Yale face database includes 165 images from 15 objects, and each object has 11 images under various lighting conditions and facial expressions. In this experiment, we manually crop each image and make them to 50×40 pixels.

In this experiment, 2, 3, and up to 5 images per individual are respectively selected for training set and the rest images are used for testing images. All algorithms are repeated for 10 times. In FME algorithm, the optimal neighbor number k and balance coefficients λ_1, and λ_2 are respectively set to 5, 20 and 4 for FME algorithm. For the proposed RME algorithm, the optimal neighbor number k and balance coefficients λ_1, λ_2 and λ_3 are respectively set to 5, 25,5 and 3. Table 6 shows the recognition results. From Table 6, we can know that our method has better performance than the compared algorithms.

V. CONCLUSION

In this paper, we propose a novel robust manifold embedding algorithm (RMF). The label information and manifold smoothness from the training samples and test samples are both used in the proposed RME algorithm. In order to accurately capture the manifold structure of the data, plenty of samples are needed. However, there are not enough samples in face recognition applications. Based on the symmetry of the face, the left/right mirror images of all face image samples are generated and applied to the proposed RMF. The advantages of the proposed RMF are as follows. Firstly, the label information and manifold structure are fully utilized in our RMF. Secondly, the RMR algorithm is more robust to variations of illuminations and poses. The experiments on six public face databases prove that the proposed RMF algorithm has a good performance.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

[1] Y. Luo, J. Qin, X. Xiang, Y. Tan, Q. Liu, and L. Xiang, "Coverless real-time image information hiding based on image block matching and dense convolutional network," *J. Real-Time Image Process.*, vol. 17, no. 1, pp. 125–135, Feb. 2020, doi: 10.1007/s11554-019-00917-3.

[2] L. Xiang, G. Guo, J. Yu, V. S. Sheng, and P. Yang, "A convolutional neural network-based linguistic steganalysis for synonym substitution steganography," *Math. Biosci. Eng.*, vol. 17, no. 2, pp. 1041–1058, 2020.

[3] Y. Chen, W. Xu, J. Zuo, and K. Yang, "The fire recognition algorithm using dynamic feature fusion and IV-SVM classifier," *Cluster Comput.*, vol. 22, no. S3, pp. 7665–7675, May 2019.

[4] Y. Zhang, W. Lu, W. Ou, G. Zhang, X. Zhang, J. Cheng, and W. Zhang, "Chinese medical question answer selection via hybrid models based on CNN and GRU," *Multimedia Tools Appl.*, Mar. 2019.

[5] F. Yu, Z. Zhang, L. Liu, H. Shen, Y. Huang, C. Shi, S. Cai, Y. Song, S. Du, and Q. Xu, "Secure communication scheme based on a new 5D multistable four-wing memristive hyperchaotic system with disturbance inputs," *Complexity*, vol. 2020, pp. 1–16, Jan. 2020.

[6] F. Wu, X.-Y.Jing, S. Wu, G. Gao, Q. Ge, and R. Wang, "'Like charges repulsion and opposite charges attraction' law based multilinear subspace analysis for face recognition," *Knowledge Based Syst.*, vol. 149, pp. 76–87, Jun. 2018.

[7] J. Wen, Y. Xu, and H. Liu, "Incomplete multiview spectral clustering with adaptive graph learning," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1418–1429, Apr. 2020.

[8] G. Gao, Y. Yu, M. Yang, H. Chang, P. Huang, and D. Yue, "Cross-resolution face recognition with pose variations via multilayer locality-constrained structural orthogonal procrustes regression," *Inf. Sci.*, vol. 506, pp. 19–36, Jan. 2020.

[9] G.-H. Liu and J.-Y. Yang, "Exploiting color volume and color difference for salient region detection," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 6–16, Jan. 2019.

[10] Z. Liu, J. Wang, G. Liu, and L. Zhang, "Discriminative low-rank preserving projection for dimensionality reduction," *Appl. Soft Comput.*, vol. 85, Dec. 2019, Art. no. 105768.

[11] F. Nie, S. Xiang, Y. Jia, and C. Zhang, "Semi-supervised orthogonal discriminant analysis via label propagation," *Pattern Recognit.*, vol. 42, no. 11, pp. 2615–2627, Nov. 2009.

[12] W. Lu, X. Zhang, H. Lu, and F. Li, "Deep hierarchical encoding model for sentence semantic matching," *J. Vis. Commun. Image Represent.*, Mar. 2020, doi: 10.1016/j.jvcir.2020.102794.

[13] J. Wen, Z. Zhang, Z. Zhang, L. Fei, and M. Wang, "Generalized incomplete multiview clustering with flexible locality structure diffusion," *IEEE Trans. Cybern.*, early access, May 7, 2020, doi: 10.1109/TCYB.2020.2987164.

[14] X. Dong, F. Wu, and X. Y. Jing, "Generic training set based multimanifold discriminant learning for single sample face recognition," *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 1, pp. 368–391, 2018.

[15] F. Wu, X.-Y. Jing, X. Dong, R. Hu, D. Yue, L. Wang, Y.-M. Ji, R. Wang, and G. Chen, "Intraspectrum discrimination and interspectrum correlation analysis deep network for multispectral face recognition," *IEEE Trans. Cybern.*, vol. 50, no. 3, pp. 1009–1022, Mar. 2020.

[16] K. Wu, Y. Guo, and C. Zhang, "Compressing deep neural networks with sparse matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Nov. 14, 2019, doi: 10.1109/TNNLS.2019.2946636.

[17] Y. Feng, F. Wu, Q. Huang, X.-Y. Jing, Y. Ji, J. Yu, F. Chen, and L. Han, "Cross-modality multi-task deep metric learning for sketch face recognition," in *Proc. Chin. Autom. Congr. (CAC)*, Hangzhou, China, Nov. 2019, pp. 2277–2281.

[18] G. Liu, Z. Liu, S. Liu, J. Ma, and F. Wang, "Registration of infrared and visible light image based on visual saliency and scale invariant feature transform," *EURASIP J. Image Video Process.*, vol. 2018, no. 1, p. 45, Dec. 2018.

[19] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," in *Proc. Int. Conf. Knowl. Discovery Data Mining*, 2014, pp. 977–986.

[20] Y. Chen, J. Wang, X. Chen, M. Zhu, K. Yang, Z. Wang, and R. Xia, "Single-image super-resolution algorithm based on structural self-similarity and deformation block features," *IEEE Access*, vol. 7, pp. 58791–58801, 2019.

[21] L. Xiang, X. Shen, J. Qin, and W. Hao, "Discrete multi-graph hashing for large-scale visual search," *Neural Process. Lett.*, vol. 49, no. 3, pp. 1055–1069, Jun. 2019.

[22] S. Xiang, F. Nie, C. Zhang, and C. Zhang, "Nonlinear dimensionality reduction with local spline embedding," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1285–1298, Sep. 2009.

[23] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[24] E. Vural and C. Guillemot, "Out-of-Sample generalizations for supervised manifold learning for classification," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1410–1424, Mar. 2016.

[25] Y. Lu, Z. Lai, Z. Fan, J. Cui, and Q. Zhu, "Manifold discriminant regression learning for image classification," *Neurocomputing*, vol. 166, pp. 475–486, Oct. 2015.

[26] B. Raducanu and F. Dornaika, "Embedding new observations via sparse-coding for non-linear manifold learning," *Pattern Recognit.*, vol. 47, no. 1, pp. 480–492, Jan. 2014.

[27] L. Shi, J. Hao, and X. Zhang, "Image recognition method based on supervised multi-manifold learning," *J. Intell. Fuzzy Syst.*, vol. 32, no. 3, pp. 2221–2232, Feb. 2017.

[28] Z. Zhang, J. Wang, and H. Zha, "Adaptive manifold learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 253–265, Feb. 2012.

[29] L. Wang, K. Wang, and R. Li, "Unsupervised feature selection based on spectral regression from manifold learning for facial expression recognition," *IET Comput. Vis.*, vol. 9, no. 5, pp. 655–662, Oct. 2015.

[30] X. Fang, Y. Xu, X. Li, Z. Fan, H. Liu, and Y. Chen, "Locality and similarity preserving embedding for feature selection," *Neurocomputing*, vol. 128, pp. 304–315, Mar. 2014.

[31] L. Wang, H. Wu, and C. Pan, "Manifold regularized local sparse representation for face recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 4, pp. 651–659, Apr. 2015.

[32] T. Zhang, D. Tao, X. Li, and J. Yang, "Patch alignment for dimensionality reduction," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1299–1313, Sep. 2009.

[33] Z. Jin, "Enhanced discriminative locality alignment and its kernel extension," *Opt. Eng.*, vol. 50, no. 8, Aug. 2011, Art. no. 087002.

[34] J. Wright, A. Y. Yang, A. Ganesh, S. Shankar Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[35] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 471–478.

[36] M. Yang, L. Zhang, D. Zhang, and S. Wang, "Relaxed collaborative representation for pattern classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2224–2231.

[37] Y. Xu, D. Zhang, J. Yang, and J.-Y. Yang, "A two-phase test sample sparse representation method for use with face recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 9, pp. 1255–1262, Sep. 2011.

[38] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2106–2112, Nov. 2010.

[39] F. Nie, D. Xu, I. W.-H. Tsang, and C. Zhang, "Flexible manifold embedding: A framework for semi-supervised and unsupervised dimension reduction," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1921–1932, Jul. 2010.

[40] G.-H. Liu and J.-Y. Yang, "Exploiting color volume and color difference for salient region detection," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 6–16, Jan. 2019.

[41] G.-H. Liu, J.-Y. Yang, and Z. Li, "Content-based image retrieval using computational visual attention model," *Pattern Recognit.*, vol. 48, no. 8, pp. 2554–2566, Aug. 2015.

[42] Y. Xu, X. Zhu, Z. Li, G. Liu, Y. Lu, and H. Liu, "Using the original and 'symmetrical face' training samples to perform representation based two-step face recognition," *Pattern Recognit.*, vol. 46, no. 4, pp. 1151–1158, Apr. 2013.

[43] X. Zhu, Z. Ghahramani, and J. D. Lafferty, "Semi-supervised Learning using Gaussian fields and harmonic functions," in *Proc. 20th Int. Conf. Mach. Learn. (ICML)*, 2003, pp. 912–919.

[44] Y.-J. Song, Y.-G. Kim, U.-D. Chang, and H. B. Kwon, "Face recognition robust to left/right shadows; facial symmetry," *Pattern Recognit.*, vol. 39, no. 8, pp. 1542–1545, Aug. 2006.

[45] F. Nie, X. Wang, M. I. Jordan, and H. Huang, "The constrained Laplacian rank algorithm for graph-based clustering," in *Proc. 13th AAAI Conf. Artif. Intell.*, 2016, pp. 1969–1976.

[46] X. J. Zhu, "Semi-supervised learning literature survey," *Comput. Sci.*, vol. 37, no. 1, pp. 63–77, 2008.

[47] V. Sindhwani, P. Niyogi, M. Belkin, and S. Keerthi, "Linear manifold regularization for large scale semi-supervised learning," in *Proc. 22nd ICML Workshop Learn. Partially Classified Training Data*, Aug. 2005, pp. 80–83.

[48] W. Yang, Z. Wang, and C. Sun, "A collaborative representation based projections method for feature extraction," *Pattern Recognit.*, vol. 48, no. 1, pp. 20–27, Jan. 2015.

[49] F. Nie, "Orthogonal locality minimizing globality maximizing projections for feature extraction," *Opt. Eng.*, vol. 48, no. 1, Jan. 2009, Art. no. 017202.

[50] Y. Dong, H. Wu, X. Li, C. Zhou, and Q. Wu, "Multiscale symmetric dense micro-block difference for texture classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 12, pp. 3583–3594, Dec. 2019.

[51] Z. Liu, Z. Lai, W. Ou, K. Zhang, and R. Zheng, "Structured optimal graph based sparse feature extraction for semi-supervised learning," *Signal Process.*, vol. 170, May 2020, Art. no. 107456.

[52] S. A. Billings and K. L. Lee, "Nonlinear Fisher discriminant analysis using a minimum squared error cost function and the orthogonal least squares algorithm," *Neural Netw.*, vol. 15, no. 2, pp. 263–270, Mar. 2002.

[53] Z. Liu, G. Liu, J. Pu, and S. Liu, "Noisy label based discriminative least squares regression and its kernel extension for object identification," *KSII Trans. Internet Inf. Syst.*, vol. 11, no. 5, pp. 2523–2538, 2017.

[54] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image Vis. Comput.*, vol. 16, no. 5, pp. 295–306, Apr. 1998.

● ● ●