IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 15, 2016, accepted November 21, 2016, date of publication November 29, 2016,
date of current version January 4, 2017.

Digital Object Identifier 10.1109/ACCESS.2016.2633281

Locality-Constrained Double Low-Rank Representation for Effective Face Hallucination

GUANGWEI GAO[1,5], (Member IEEE), XIAO-YUAN JING[2], PU HUANG[3], QUAN ZHOU[4,5], SONGSONG WU[2], (Member IEEE), AND DONG YUE[1], (Senior Member, IEEE)

[1]Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[2]School of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[3]School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[4]Key Laboratory of Ministry of Education for Broad Band Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[5]School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China

Corresponding author: G. Gao (csggao@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61402240, Grant 61401228, Grant 61502245, Grant 61503195, Grant 61571236, Grant 61502081, Grant 61503188, and Grant 61671253, in part by the China Postdoctoral Science Foundation under Grant 2015M571786, and Grant 2016M600433, in part by the Postdoctoral Science Foundation of Jiangsu Province under Grant 1501019A, in part by the Natural Science Fund for Colleges and Universities in Jiangsu Province under Grant 15KJB520026, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20150849, in part by NUPTSF under Grant NY214204 and Grant NY214119, in part by the PAPD, in part by the CICAEET, in part by the Key Project of Chinese Ministry of Education under Grant 313030, in part by the 973 Program under Grant 2014CB349303, and in part by the Program for Changjiang Scholars and Innovative Research Team in University under Grant IRT13072.

ABSTRACT Recently, position-patch-based face hallucination methods have received much attention, and obtained promising progresses due to their effectiveness and efficiency. A locality-constrained double low-rank representation (LCDLRR) method is proposed for effective face hallucination in this paper. LCDLRR attempts to directly use the image-matrix based regression model to compute the representation coefficients to maintain the essential structural information. On the other hand, LCDLRR imposes a low-rank constraint on the representation coefficients to adaptively select the training samples that belong to the same subspace as the inputs. Moreover, a locality constraint is also enforced to preserve the locality and the sparsity simultaneously. Compared with previous methods, our proposed LCDLRR considers locality manifold structure, cluster constraints, and structure error simultaneously. Extensive experimental results on standard face hallucination databases indicate that our proposed method outperforms some state-of-the-art algorithms in terms of both visual quantity and objective metrics.

INDEX TERMS Face hallucination, low-rank representation, position-patch, nuclear norm.

I. INTRODUCTION

Object classification is an essential problem in pattern recognition community [1]–[9]. In the past twenty years, face recognition has been an active topic in various applications [10]–[22], such as surveillance, human-computer interface, etc. However, due to the restrictions of surveillance system, such as long distance to the interest object, it is sometimes difficult to gain high-definition face images. Face hallucination, or face super-resolution, is a technology to acquire high-resolution (HR) face images from low-resolution (LR) observations [23], thus providing more facial details for the following recognition process.

Various approaches have been presented in the past few decades for face hallucination, which can be roughly classified into two categories: interpolation and learning-based methods. Compared with the interpolation-based methods, learning-based methods have gained more attention since they can significantly improve the visual quality for super-resolution reconstruction with large magnification factor [24]. As we all know, Baker and Kanade [25] firstly proposed the term "face hallucination" under a Bayesian formulation, which estimates HR image with the assistance of HR-LR training sample pairs. Wang and Tang [26] represented the input LR image as a linear combination over the LR training images by principal component analysis (PCA). Nevertheless, the obtained HR images usually contain ghost artifacts duo to the usage of the PCA-based holistic appearance model. Subsequently, Hu *et al.* [27] devised a kernel-based extension of the eigentransformation method, which takes non-linear manifold structures of the image data into account. Liang *et al.* [28] formulated the face hallucination process as an image decomposition problem by morphological component analysis (MCA) and presented a three-step framework for this task. Huang *et al.* [29] and An and Bhanu [30] applied canonical correlation analysis (CCA) to find a coherent subspace which maximizes the correlation between the PCA coefficients of corresponding LR and HR images. Then the global face image is reconstructed in the coherent subspace. These methods can preserve the structure of global face variations, but ignore some facial details beyond principal components and fail to get well the individual details, particularly when the available training samples are limited.

In order to handle the problem of global face hallucination methods, researchers develop local patch based methods that divide the whole face into several overlapped patches. Following the well-known locally linear embedding (LLE) [31], Chang *et al.* [32] presented a neighbor embedding (NE) based super-resolution method, which first seeks K-nearest patches from LR training samples to learn optimal representation coefficients under the minimization of reconstruction error and then uses the corresponding HR training samples to recover hallucinated image with the same combination weights. Many following works of NE have been reported for face hallucination, e.g. locality-constraint iterative neighbor embedding [33], low-rank neighbor embedding [34], coupled-layer neighbor embedding [35], etc. Different from those methods applying a fixed number of neighbors for representation, Yang *et al.* [36] introduced sparse coding technique that can adaptively select the most related candidates to minimize the reconstruction error. Recognizing that face image is highly structured, the position prior can be fully incorporated into face hallucination procedure. Ma *et al.* [37] presented the position-patch based face hallucination technique, which utilizes the constrained least square formulation to get the representation vectors with the same position patches extracted from the training samples. However, if the number of the training manifolds is very large which exceeds the dimension of each patch, there is no optimal reconstruction weights because of lacking unique solution to least square representation. To tackle this problem, Jung *et al.* [38] employed sparse representation (SR) constraints prior to improve the reconstruction results by adaptively selecting dictionary atoms. Wang *et al.* [39], [40] further proposed a weighted adaptive sparse regularization (WASR) method to super-resolve face images. Jiang *et al.* [41]–[43] recently proposed locality-constrained representation (LcR) to generate super resolved images. This method selects the most relevant patches adaptively by penalizing the Euclid distances between input patch and each training patch and constraining the least square reconstruction, which is conductive to preserve the sparsity and locality simultaneously.

Although the aforementioned methods take linear subspace properties into account, they neglect that patches from different manifolds may lie in independent linear subspaces. Low-rank structured constraints should be imposed on these linear subspaces. On the other hand, the above methods (such as [32], and [36]–[38]) are all vector-based ones. In other words, before calculating the representation coefficients, we must convert the patch matrices into vectors beforehand. Nevertheless, in the converting process, some structural information (such as the rank of matrix) might be ignored. To overcome these shortcomings, a novel face hallucination method called locality-constraint double low-rank representation (LCDLRR) is presented in this paper. LCDLRR aims to cluster the input to its potential low-rank subspace and then apply the matrix based regression model to straightforward calculate the representation coefficients (without the matrix-to-vector conversion). Low-rank constraint on representations utilizes the training samples from the same subspace with more discriminative ability to reconstruct the input patches. We also desire to utilize the minimal rank of representation residual matrix as a criterion to compute the representation coefficients. Compared with SR and LcR (SR only favor the sparsity, while LcR only stress the closest atoms), LCDLRR utilizes locality manifold structure, subspace constraints and structure error simultaneously, with atoms nearest to the input for the reconstruction. Experimental results on standard face hallucination database demonstrate that our method can gain better performance than some other state-of-the-art approaches.

The remainder of the paper is organized as follows. Section 2 briefly reviews related works on position-patch based face hallucination. In Section 3, we present our proposed locality-constrained double low-rank representation method. Section 4 evaluates the performance of the proposed methods on several commonly used face hallucination databases. Section 5 concludes this paper.

II. RELATED WORKS

Let L^m and H^m $(m = 1, \ldots, N)$ represent the LR and HR training samples, where N denotes the size of the training samples. Each face image is divided into M overlapped patches and denoted as $\{L^m(i,j), H^m(i,j) | 1 \leq i \leq R, 1 \leq j \leq C\}$, $M = R \times C$, R and C are the patch number in every row and column respectively, term (i,j) denotes the position information.

For the LR input testing face image y, its divided patches are also denoted as $\{y(i,j) | 1 \leq i \leq R, 1 \leq j \leq C\}$, whose reconstruction weight vector over the training sample is $w(i,j) = [w_1(i,j), w_2(i,j), \ldots, w_N(i,j)]^T$, and $w_m(i,j)$ represents the contribution of each training image patch to the reconstruction of the input image patch located at position (i,j). Different methods transform each input patch into an N-dimensional coefficient vector to generate the final patch representation. In this section, we briefly introduce three existing patch based methods.

A. LEAST SQUARE REPRESENTATION

By introducing the position prior information, Ma *et al.* [37] used all training patches at the same position to collaboratively represent input patch with

$$y(i,j) = \sum_{m=1}^{N} L^m(i,j)w_m(i,j) + e \quad (1)$$

here e denotes the reconstruction error vector.

The optimal representation coefficients of the input image patch $y(i,j)$ can be obtained via solving the following constrained least square representation formulation:

$$w^*(i,j) = \arg\min_{x(i,j)} \left\| y(i,j) - \sum_{m=1}^{N} L^m(i,j)w_m(i,j) \right\|_2^2 \quad s.t. \sum_{m=1}^{N} w_m(i,j) = 1 \quad (2)$$

It is a constrained least square problem, whose solution can be obtained via computing a Gram matrix.

B. SPARSE REPRESENTATION

In fact, equation (2) may have unstable solution. One technical substitution is to enforce regularization terms on the objective function. Jung *et al.* [38] incorporated the sparse coding theory that can adaptively choose the most related components to minimize the reconstruction error. Equation (2) can be converted into a standard sparse coding problem:

$$\min_{w} \|w(i,j)\|_0 \quad s.t. \left\| y(i,j) - \sum_{m=1}^{N} L^m(i,j)w_m(i,j) \right\|_2^2 \le \varepsilon \quad (3)$$

where the l_0 norm counts the number of nonzero entries in a vector. Since the l_0 norm minimization is an NP-hard problem, latest studies from sparse representation [44] suggest that if the solution is sparse enough, we can recover the sparsest solution via the following l_1-norm minimization:

$$\min_{w} \|w(i,j)\|_1 \quad s.t. \left\| y(i,j) - \sum_{m=1}^{N} L^m(i,j)w_m(i,j) \right\|_2^2 \le \varepsilon \quad (4)$$

where the l_1 norm sums up the absolute value of all entries in a vector. It should be noted that the sparsity constraint enables the learned representation to capture salient properties, yielding minimized reconstruction error.

C. LOCALITY-CONSTRAINED REPRESENTATION

In sparse representation method [38], it imposes great importance onto sparsity of the reconstruction weights, while neglects the locality characteristic which can maintain the essential geometry of training samples. Locality-constrained representation (LcR) [41] method is proposed to represent the input LR patch to gain sparsity and locality simultaneously, which is formulated via minimizing the following objective:

$$w^*(i,j) = \arg\min_{w(i,j)} \left\| y(i,j) - \sum_{m=1}^{N} L^m(i,j)w_m(i,j) \right\|_2^2 + \lambda \sum_{m=1}^{N} [d_m(i,j)w_m(i,j)]^2 \quad s.t. \sum_{m=1}^{N} w_m(i,j) = 1 \quad (5)$$

Here λ is a variable to balance the contribution of the reconstruction error and the solution locality, and $d_m(i,j)$ is obtained by the Euclidean distance:

$$d_m(i,j) = \left\| y(i,j) - L^m(i,j) \right\|_2, \quad m = 1, \ldots, M \quad (6)$$

Given the observed LR input $y(i,j)$, the goal of face hallucination is to obtain its corresponding HR output $x(i,j)$ and then integrate all position patches to constitute its HR counterpart X. Similar to previous methods, we compute the HR patch via $x(i,j) = H(i,j)w(i,j)$.

III. LOCALITY-CONSTRAINED DOUBLE LOW-RANK REPRESENTATION FOR FACE HALLUCINATION

A. PROBLEM FORMULATION

In previous methods, the image patches are all denoted in vector form, and l_2 norm (or l_1 norm) are usually applied to describe the reconstruction residual. However, these norms are based on pixel values, thus the structural information (such as the rank of matrix) of the error image is ignored. In addition, from the distribution point of view, we can observe that l_2 norm provides an optimal description for errors following the Gaussian distribution, while l_1 norm is optimal for Laplacian distribution characterization. Thus, l_2 and l_1 norm cannot characterize this kind of structure error effectively. Fortunately, as it can be seen from the experiments that the singular values of reconstruction residual fit the Laplacian distribution well. We know that nuclear norm is the sum of all singular values of a matrix, which can also be considered as l_1 norm of the singular value vector. To this end, we propose to apply the nuclear norm based matrix regression to characterize the reconstruction residual.

In this subsection, all the face image patches are represented in the matrix form. In other words, the patches of the input and training face images can be denoted as $y(i,j) \in \Re^{d\times d}$ and $L^m(i,j) \in \Re^{d\times d}$ $(m = 1, \ldots, N)$, respectively. For the convenience of expression, we omit the position index (i,j) in the following text. Then, the problem can be formulated as follows:

$$\min_{w} \|L(w) - y\|_* + \beta \|d \otimes w\|_2^2 \quad (7)$$

where $L(w) = w_1L^1 + w_2L^2 + \ldots + w_NL^N$, $\otimes$ denotes the element wise product.

In [45], low-rank representation was used to characterize the data structure via clustering the input into its most relevant independent subspace, gaining promising performance in recognition goals. By low-rank structure, the training samples and input patch in the similar subspace can be clustered. In this subspace, the original input can be linearly represented by the neighborhood from the similar subspace. Mathematically, our objective function can be formulated as follows:

$$\min_{w} \|L(w) - y\|_* + \alpha \|Hdiag(w)\|_* + \beta \|d \otimes w\|_2^2 \quad (8)$$

Here, $H = [Vec(L^1), Vec(L^2), \ldots, Vec(L^N)]$ and $Vec(\cdot)$ denotes the vectorization operator, α and β are variables to balance the low-rank property and locality constraints, $||\cdot||_*$ is the nuclear norm (i.e. the sum of the singular values) of a matrix. d (as defined in (6)) is the locality adaptor used to measure the distance between the input and dictionary atoms. The nuclear norm in (8) is a convex substitution for rank operation and it has been shown that nuclear norm based models can achieve the optimal low rank solution in many scenarios [46]. Matrix $Hdiag(w)$ represents the atoms that are exploited to reconstruct the observation y. Minimizing the rank of this matrix means reconstructing y applying only those patches belong to a low-rank subspace.

B. OPTIMIZATION PROCEDURE

For the convenience of expression, objective (8) can be converted into the following equivalent formulation:

$$\min_{w,E,Z} \|E\|_* + \alpha \|Z\|_* + \beta \|d \otimes w\|_2^2$$
$$s.t.\ E = L(w) - y, \quad Z = Hdiag(w) \quad (9)$$

The above problem can be solved via the alternating direction method of multipliers (ADMM) [47]–[51], which minimizes the following augmented Lagrangian function:

$$\begin{aligned}\min_{w,E,Z} &\|E\|_* + \alpha \|Z\|_* + \beta \|d \otimes w\|_2^2 \\ &+ \langle Y_1, L(w) - y - E\rangle + \langle Y_2, Hdiag(w) - Z\rangle \\ &+ \frac{\mu}{2}\left(\|L(w) - y - E\|_F^2 + \|Hdiag(w) - Z\|_F^2\right) \quad (10)\end{aligned}$$

where $\langle A, B\rangle = trace(A^T B)$, $|| \cdot ||_F$ denotes the Frobenius norm, Y_1 and Y_2 are the Lagrange multipliers, $\mu > 0$ is a penalty parameter. The above objective is unconstrained, thus it can be solved as regards to w, E and Z, respectively, by fixing other variables, and then updating the Lagrange multipliers Y_1 and Y_2.

For convenience, we rewrite the augmented Lagrangian function (10) as follows:

$$\begin{aligned}\min_{w,E,Z} &\|E\|_* + \alpha \|Z\|_* + \beta \|d \otimes w\|_2^2 \\ &+ \frac{\mu}{2}\left(\left\|L(w) - y - E + \frac{1}{\mu}Y_1\right\|_F^2 \right. \\ &\left. + \left\|Hdiag(w) - Z + \frac{1}{\mu}Y_2\right\|_F^2\right) \\ &- \frac{1}{2\mu}\|Y_1\|_F^2 - \frac{1}{2\mu}\|Y_2\|_F^2 \quad (11)\end{aligned}$$

Updating Z

Given x and E, the objective function (11) can be reformulated as

$$\min_{Z} \frac{\alpha}{\mu}\|Z\|_* + \frac{1}{2}\left\|Z - \left(Hdiag(w) + \frac{1}{\mu}Y_2\right)\right\|_F^2 \quad (12)$$

The optimal solution can be obtained by the singular value thresholding operator [52]. Specifically, given the matrix $Q \in \Re^{p\times q}$ with rank r, its singular value decomposition (SVD) is

$$Q = \mathbf{U}_{p\times r}\mathbf{\Sigma}\mathbf{V}_{q\times r}^T, \quad \mathbf{\Sigma} = diag(\sigma_1, \ldots, \sigma_r) \quad (13)$$

where $\sigma_1, \ldots, \sigma_r$ are singular values, U and V are corresponding orthogonal matrices. For a given $\tau > 0$, the singular value thresholding operator $T_\tau(\cdot)$ is defined as

$$T_\tau(Q) = \mathbf{U}_{p\times r} diag\left(\left\{\max\left(0, \sigma_j - \tau\right)\right\}_{1\le j\le r}\right)\mathbf{V}_{q\times r}^T \quad (14)$$

Theorem 1 [52]: For matrix $Q \in \Re^{p\times q}$ and $\tau > 0$, the singular value thresholding operator in (14) submits to

$$T_\tau(Q) = \arg\min_{J}\left(\tau\|J\|_* + \frac{1}{2}\|J - Q\|_F^2\right) \quad (15)$$

From Theorem 1, the optimal solution of (12) is

$$Z = T_{\frac{\alpha}{\mu}}\left(Hdiag(w) + \frac{1}{\mu}Y_2\right) \quad (16)$$

Updating E

Given x and Z, the objective function (11) can be rewritten as

$$\min_{E} \frac{1}{\mu}\|E\|_* + \frac{1}{2}\left\|E - \left(L(w) - y + \frac{1}{\mu}Y_1\right)\right\|_F^2 \quad (17)$$

From Theorem 1, the optimal solution of (17) is

$$E = T_{\frac{1}{\mu}}\left(L(w) - y + \frac{1}{\mu}Y_1\right) \quad (18)$$

Updating x

Given E and Z, the minimization problem (11) can be reformulated as

$$\begin{aligned}\min_{w} \beta\|d \otimes w\|_2^2 &+ \frac{\mu}{2}\left(\left\|L(w) - y - E + \frac{1}{\mu}Y_1\right\|_F^2\right. \\ &\left. + \left\|Hdiag(w) - Z + \frac{1}{\mu}Y_2\right\|_F^2\right) \quad (19)\end{aligned}$$

This problem is a quadratic form in variable w. Differentiating the objective function in regard to x, and let it be zero, we can obtain the optimal solution as follows:

$$\begin{aligned} w &= (P + diag(p_1)) \backslash p_2 \\ P &= \mu H^T H + 2\beta diag(d) \otimes diag(d) \\ p_1 &= \mu (H \otimes H)^T \mathbf{1} \\ p_2 &= \mu H^T p_3 + \mu (Z \otimes H)^T \mathbf{1} - (Y_2 \otimes H)^T \mathbf{1} \\ p_3 &= Vec(y + E - Y_1/\mu) \quad (20)\end{aligned}$$

Algorithm 1 ADMM Algorithm for Solving LCDLRR

Input: Training patch matrices $L^1, \ldots, L^N$ and test patch matrix y, parameters α and β, the termination condition parameter ε.
Initialize: $w = 0$, $Y_1 = 0$, $Y_2 = 0$, $\mu = 10^{-6}$, $max_\mu = 10^{10}$, $\rho = 1.1$, $\varepsilon = 10^{-8}$

1: Fix the others and update Z by

$$\min_Z (\alpha/\mu) \|Z\|_* + 0.5 \|Z - (Hdiag(w) + Y_2/\mu)\|_F^2 ;$$

2: Fix the others and update E by

$$\min_E (1/\mu) \|E\|_* + 0.5 \|E - (L(w) - y + Y_1/\mu)\|_F^2 ;$$

3: Fix the others and update w by formulation (20);
4: Update the multiplies

$$Y_1 \leftarrow Y_1 + \mu (L(w) - y - E)$$
$$Y_2 \leftarrow Y_2 + \mu (Hdiag(w) - Z)$$
$$\mu \leftarrow \min (\rho\mu, max_\mu)$$

5. If termination condition (21) is satisfied, go to 6; otherwise go to 1.

6. **Output:** Optimal coding vector w

where **1** denotes a $N \times 1$ column vector of ones, the operate "\" denotes the left matrix division operation.

Here, we apply the following termination conditions:

$$\|Hdiag(w) - Z\|_\infty \leq \varepsilon \quad and \quad \left\| L(w) - y - E + \frac{1}{\mu} Y_2 \right\|_\infty \leq \varepsilon \tag{21}$$

where ε is a given tolerance.

In summary, the detailed algorithm via ADMM to solve problem (9) is summarized in Algorithm 1.

C. FACE HALLUCINATION VIA LCDLRR

With regard to face hallucination, the training set consists of LR and HR face image pairs. H^m denote the HR face images, while L^m $(m = 1, \ldots, N)$ denote their LR counterparts. The face hallucination task aims to acquire the HR face image X from its LR observation y.

First of all, each training face image and the LR input are divided into overlapped patch matrices using the same dividing strategy as in [41] and denote them as $H^m(i, j)$, $L^m(i, j)$, $y(i, j)$. For each LR input image patch matrix $y(i, j)$, it is represented as a linear combination over the LR training patch matrices $L^m(i, j)$ $(m = 1, \ldots, N)$ using LCDLRR. By keeping the combination coefficients and replacing the LR training patch matrices via the corresponding HR counterparts, the desired HR patch matrix at the same position can be synthesized. By concatenating all the HR patch matrices to their relevant positions and averaging values in the overlapping regions, an estimated HR target face image can be obtained. The whole face hallucination algorithm is summarized in **Algorithm 2**.

Algorithm 2 Face Hallucination via NRC

Input: LR training images $L^1, \ldots, L^N$, corresponding HR training images $H^1, \ldots, H^N$, input LR images y.

1: Each of the training images and the LR input image is divided into M overlapped patch matrices respectively;
2: **For** each input patch matrix in y:

a) Compute the Euclidean distance between the LR input $y(i, j)$ and each of the LR training patch matrices $L^m(i, j)$ $(m = 1, \ldots, N)$:

$$d_m(i, j) = \left\| y(i, j) - L^m(i, j) \right\|_2, \; m = 1, \ldots, N$$

b) Calculate the optimal weights $w^*(i, j)$ with regard to the LR input $y(i, j)$ over the LR training patch matrices $L^m(i, j)$ $(m = 1, \ldots, N)$:

$$w^*(i, j) = \arg\min_{w(i,j)} \{ \left\| \sum_{m=1}^N L^m(i, j) w_m(i, j) - y(i, j) \right\|_* + \alpha \|H(i, j) diag(w(i, j))\| + \beta \|d(i, j) \otimes w(i, j)\|_{l_p} \}$$

c) Construct the desired HR patch by

$$x(i, j) = \sum_{m=1}^N H^m(i, j) w_m^*(i, j)$$

3: **End for**
4: The target HR image X can be obtained by integrating all the reconstructed HR patch matrices. Pixels in the overlapping regions can be obtained by averaging the pixel values in the overlapping regions between two adjacent patches.

Output: The hallucinated HR face image X.

IV. EXPERIMENTS AND DISCUSSIONS

A. DATASETS DESCRIPTION

The frontal face hallucination experiments are conducted on the CAS-PEAL database [53] and FEI database [54] to verify the effectiveness of the proposed method. All face images are manually aligned according to the locations of three points: centers of left and right eyeballs and center of the mouth (some examples are shown in Fig. 1). We clip the whole face regions and resize the HR images to 128×112 (120×100 for FEI database). The LR counterparts are obtained via smoothing (an averaging filter with size 4×4) and down-sampling (the down-sampling factor is 4) the homologous HR ones, so the LR images have size 32×28 (30×25 for FEI database). For CAS-PEAL database, a subset that consists of 1040 frontal view face images with normal lighting and normal expression is selected. Each subject has only one single image. For FEI database, 400 images from 200 subjects are selected and each person has two frontal faces, one has a neutral expression and the other has a smiling facial expression.

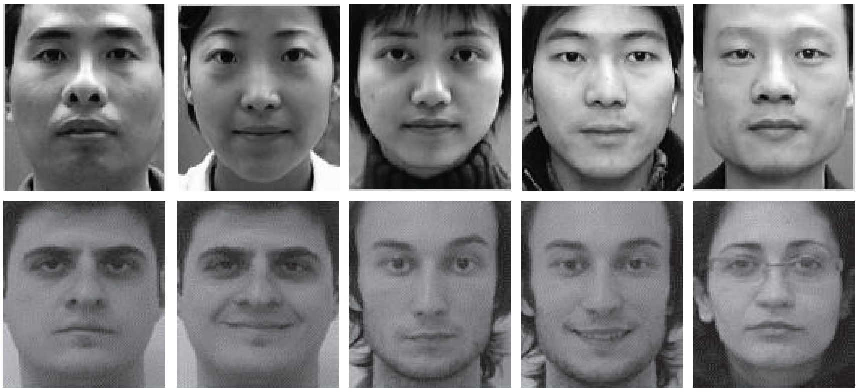


FIGURE 1. Some sample images from two databases: CAS-PEAL database (top) and FEI database (bottom).

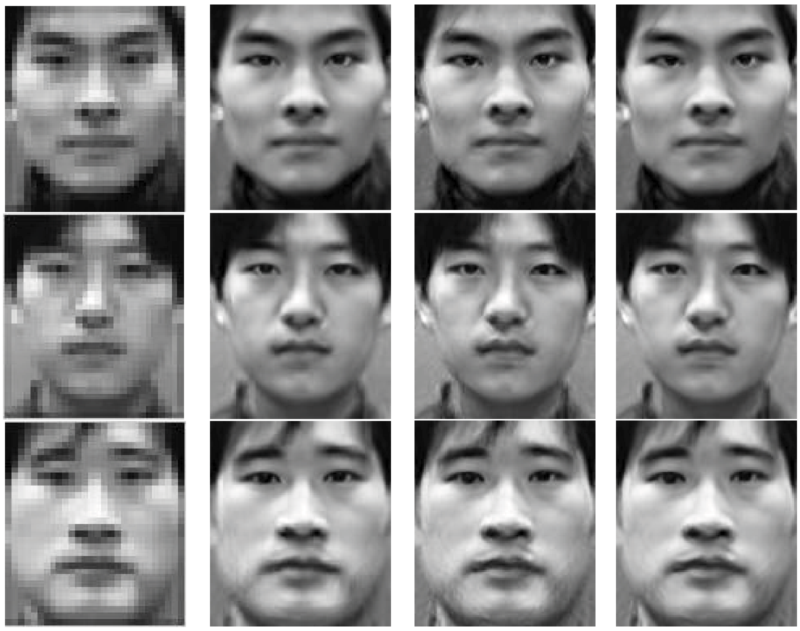


FIGURE 2. Comparisons of hallucinated results with different methods. From left to right are the input LR image, the results of Jiang's LcR method, LCMR and LCDLRR.

B. EFFECTIVENESS OF NUCLEAR NORM AND LOW-RANK REGULARIZATION

In this subsection, experiments on the CAS-PEAL database are conducted to demonstrate the effectiveness of LRR in LCDLRR method. LRR is used to choose suitable training samples to represent the input image. For fair comparison, we also visualize the hallucinated results of Jiang's locality-constrained representation (LcR) method and our method without LRR constraint (model (7), denoted as LCMR). The hallucinated results with different schemes are shown in Fig. 2. The peak signal-to-noise ration (PSNR) and Structural SIMilarity (SSIM) [55] values of the hallucinated results with different schemes are tabulated in Table 1.

From the above explorations, we can learn that: (i) by charactering the representation error with matrix regression form, the proposed LCMR and LCDLRR can gain better performance than LcR; (ii) LRR plays an important role on clustering training samples for coding, thus LCDLRR can provide complementary details for the final hallucinated results.

TABLE 1. PSNR (dB) and SSIM indexes for different methods.

Images index	1	2	3
LcR	30.6932	28.0988	27.9726
	0.9023	0.8728	0.8667
LCMR	31.1987	28.5735	28.5333
	0.9134	0.8897	0.8751
LCDLRR	31.5390	28.8647	28.7944
	0.9219	0.9013	0.8866

FIGURE 3. Comparisons of hallucinated results with different methods on the FEI database. From top to bottom are the Input LR image, the results of PCA [26], NE [32], LSR [37], SR [38], LcR [41], LCDLRR and the original HR image.

C. HALLUCINATION COMPARISONS WITH THE STATE-OF-THE-ART METHODS

In this subsection, we compare LCDLRR method with Wang's PCA approach [26], and four other patch-based approaches, e.g., Chang's neighbor embedding (NE) approach [32], Ma's least square representation (LSR) approach [37], Jung's sparse representation (SR) approach [38] and Jiang's locality-constrained representation (LcR)

FIGURE 4. Comparisons of hallucinated results with different methods on the CAS-PEAL database. From top to bottom are the Input LR image, the results of PCA [26], NE [32], LSR [37], SR [38], LcR [41], LCDLRR, and the original HR image.

approach [41]. For both FEI and CAS-PEAL databases, we randomly select 250 images for training, and another 40 images for testing. For all comparative methods, we tune their parameters to achieve the best possible results. The neighbors number in Chang's method is 50. As for these patch-based methods, we suggest using the size

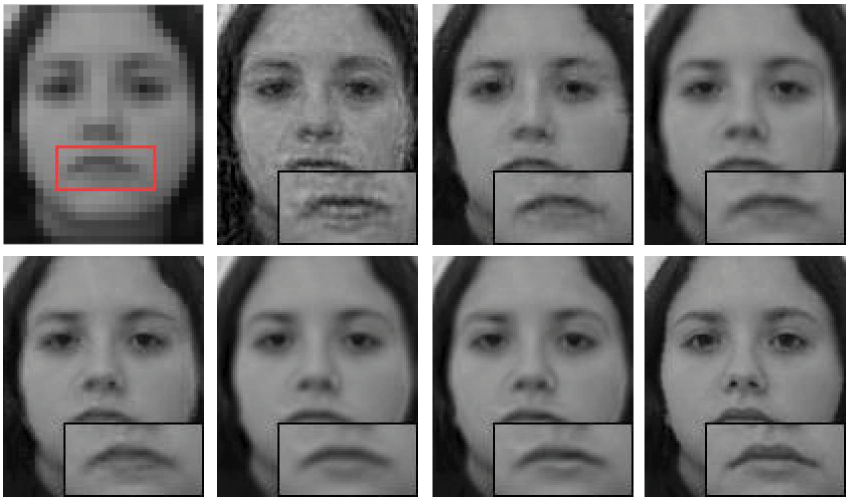


FIGURE 5. Visual differences of hallucinated results with different methods on the FEI database. From left to right and top to bottom are the Input LR image, the results of PCA [26], NE [32], LSR [37], SR [38], LcR [41], LCDLRR and the original HR image.

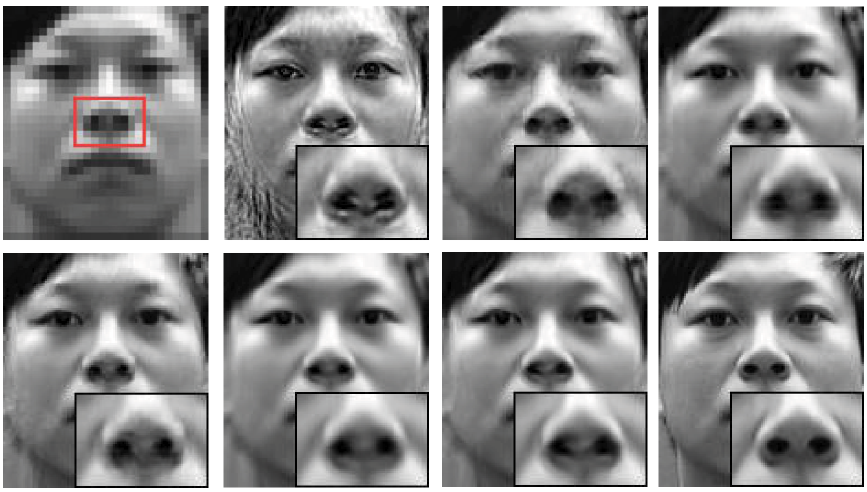


FIGURE 6. Visual differences of hallucinated results with different methods on the CAS-PEAL database. From left to right and top to bottom are the Input LR image, the results of PCA [26], NE [32], LSR [37], SR [38], LcR [41], LCDLRR and the original HR image.

3×3 pixels for LR patch and the overlaps between its neighbor patches are 3×1 pixels, and the homologous HR patch is 12×12 pixels with an overlap of 12×4 pixels.

Fig. 3 and Fig. 4 show some representative hallucinated results obtained by different approaches. The first row shows the input LR faces, the last row shows their original HR ones, while the second to the seventh rows show the hallucinated HR counterparts using different approaches. From Fig. 3 and Fig. 4 we can see that Wang's PCA method leads to "ghosting" artifacts around eye, mouth and face contours. Furthermore, the hallucinated faces seem not be similar to the original HR ones to some extent. Chang's NE approach tends to produce noises on the hallucinated results especially on locations around eyes. Ma's LSR approach generates some smoothness effects around the eyes and mouth. Jung's SR approach enhances the edges around the eyes and mouth to some extent, while the ringing effects also occur around the face contour. In contrast, Jiang's LcR method and our LCDLRR method hallucinate competitive face images with more details in the eye, mouth and face contour than other methods and meanwhile the results of LCDLRR are much more similar to the original HR ones. Fig. 5 and Fig. 6 mark some visual differences on the hallucinated HR faces generated by different approaches, from which we can observe that our LCDLRR can obtain better super-resolved individual details.

The average PSNR and SSIM values of different methods are tabulated in Tables 2 and 3. It can be seen that the proposed LCDLRR obtains the best performance among all competitive methods. From the above analysis, we can have the

TABLE 2. The average PSNR (dB) and SSIM indexes on the FEI database.

Methods	PCA[26]	NE[32]	LSR[37]	SR[38]	LcR[41]	LCDLRR
PSNR(dB)	26.1360	31.1680	31.8206	31.8860	32.3029	**32.8313**
SSIM	0.6831	0.8920	0.9013	0.9087	0.9108	**0.9162**

TABLE 3. The average PSNR (dB) and SSIM indexes on the CAS-PEAL database.

Methods	PCA[26]	NE[32]	LSR[37]	SR[38]	LcR[41]	LCDLRR
PSNR(dB)	25.3450	27.3842	28.2078	28.0306	28.5735	**28.9978**
SSIM	0.7560	0.8333	0.8910	0.8827	0.9005	**0.9034**

following observations: (i) patch-based methods substantially outperform global-based ones; (ii) by taking the face position semantics priors into account, position-patch based methods (LSR, SR and our method) can outperform NE method; (iii) by incorporating the matrix regression and LRR constraint, our LCDLRR can obtain the best performances in both of visual quality and objective assessment.

V. CONCLUSIONS

A locality-constrained double low-rank representation method is proposed for face hallucination problem in this paper. The proposed method can compute representation coefficients of input patch matrix directly without matrix-to-vector conversion. In addition, we take consideration of the low-rank and locality constraints on representation coefficients, enforcing to adaptively select the training samples from the same subspace with input and taking into account of the locality manifold distance. For each input low-resolution (LR) patch matrix, its representation coefficients over the training patch matrices at the same position can be obtained. Then its corresponding high-resolution (HR) patch matrix can be computed with the LR training patches replaced by the corresponding HR ones. Experiments on standard face hallucination databases demonstrate that our method could yield both better subjective and objective performance than some state-of-the-art approaches.

REFERENCES

[1] X. Wen, L. Shao, Y. Xue, and W. Fang, "A rapid learning algorithm for vehicle classification," *Inf. Sci.*, vol. 295, pp. 395–406, Feb. 2015.

[2] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 7, pp. 1403–1416, Jul. 2015.

[3] B. Gu, V. S. Sheng, Z. Wang, D. Ho, S. Osman, and S. Li, "Incremental learning for ν-support vector regression," *Neural Netw.*, vol. 67, pp. 140–150, Jul. 2015.

[4] B. Gu and V. S. Sheng, "A robust regularization path algorithm for ν-support vector classification," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.

[5] B. Gu, X. Sun, and V. S. Sheng, "Structural minimax probability machine," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.

[6] Z. Lai, Y. Xu, Z. Jin, and D. Zhang, "Human gait recognition via sparse discriminant projection learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 10, pp. 1651–1662, Oct. 2014.

[7] S. Banerji, A. Sinha, and C. Liu, "New image descriptors based on color, texture, shape, and wavelets for object and scene image classification," *Neurocomputing*, vol. 117, pp. 173–185, Oct. 2013.

[8] G. Gao, J. Yang, J. Qian, and L. Zhang, "Integration of multiple orientation and texture information for finger-knuckle-print verification," *Neurocomputing*, vol. 135, pp. 180–191, Jul. 2014.

[9] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 11, pp. 2594–2608, Nov. 2016.

[10] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[11] Z. Lai, Y. Xu, Q. Chen, J. Yang, and D. Zhang, "Multilinear sparse principal component analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 10, pp. 1942–1950, Oct. 2014.

[12] Z. Lai, W. K. Wong, Y. Xu, C. Zhao, and M. Sun, "Sparse alignment for robust tensor learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 10, pp. 1779–1792, Oct. 2014.

[13] W. K. Wong, Z. Lai, Y. Xu, J. Wen, and C. P. Ho, "Joint tensor feature analysis for visual object recognition," *IEEE Trans. Cybern.*, vol. 45, no. 11, pp. 2425–2436, Nov. 2015.

[14] M. Yang, P. Zhu, F. Liu, and L. Shen, "Joint representation and pattern learning for robust face recognition," *Neurocomputing*, vol. 168, pp. 70–80, Nov. 2015.

[15] Z. Lai, W. K. Wong, Y. Xu, J. Yang, and D. Zhang, "Approximate orthogonal sparse embedding for dimensionality reduction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 4, pp. 723–735, Apr. 2016.

[16] G. Gao, J. Yang, S. Wu, X. Jing, and D. Yue, "Bayesian sample steered discriminative regression for biometric image classification," *Appl. Soft Comput.*, vol. 37, pp. 48–59, Dec. 2015.

[17] P. Huang and G. Gao, "Parameterless reconstructive discriminant analysis for feature extraction," *Neurocomputing*, vol. 190, pp. 50–59, May 2016.

[18] B. Chen, H. Shu, G. Coatrieux, G. Chen, X. Sun, and J. L. Coatrieux, "Color image analysis by quaternion-type moments," *J. Math. Imag. Vis.*, vol. 51, no. 1, pp. 124–144, 2015.

[19] P. Huang, C. Chen, Z. Tang, and Z. Yang, "Feature extraction using local structure preserving discriminant analysis," *Neurocomputing*, vol. 140, pp. 104–113, Sep. 2014.

[20] H. Nguyen, W. Yang, F. Shen, and C. Sun, "Kernel Low-Rank Representation for face recognition," *Neurocomputing*, vol. 155, pp. 32–42, May 2015.

[21] W. Yang, C. Sun, and W. Zheng, "A regularized least square based discriminative projections for feature extraction," *Neurocomputing*, vol. 175, pp. 198–205, Jan. 2016.

[22] 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.

[23] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[24] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1167–1183, Sep. 2002.

[25] S. Baker and T. Kanade, "Hallucinating faces," in *Proc. 4th IEEE Int. Conf. Autom. Face Gesture Recognit.*, Mar. 2000, pp. 83–88.

[26] X. Wang and X. Tang, "Hallucinating face by eigentransformation," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 35, no. 3, pp. 425–434, Aug. 2005.

[27] Y. Hu, K. M. Lam, T. Shen, and W. Wang, "A novel kernel-based framework for facial-image hallucination," *Image Vis. Comput.*, vol. 29, no. 4, pp. 219–229, Mar. 2011.

[28] Y. Liang, X. Xie, and J.-H. Lai, "Face hallucination based on morphological component analysis," *Signal Process.*, vol. 93, no. 2, pp. 445–458, Feb. 2013.

[29] H. Huang, H. He, X. Fan, and J. Zhang, "Super-resolution of human face image using canonical correlation analysis," *Pattern Recognit.*, vol. 43, no. 7, pp. 2532–2543, 2010.

[30] L. An and B. Bhanu, "Face image super-resolution using 2D CCA," *Signal Process.*, vol. 103, pp. 184–194, Oct. 2014.

[31] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[32] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun./Jul. 2004, pp. 1275–1282.

[33] J. Jiang, R. Hu, Z. Wang, and Z. Han, "Face super-resolution via multilayer locality-constrained iterative neighbor embedding and intermediate dictionary learning," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4220–4231, Oct. 2014.

[34] X. Chen and C. Qi, "Low-rank neighbor embedding for single image super-resolution," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 79–82, Jan. 2014.

[35] J. Jiang, R. Hu, Z. Wang, Z. Han, and J. Ma, "Facial image hallucination through coupled-layer neighbor embedding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1674–1684, Sep. 2016.

[36] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[37] X. Ma, J. Zhang, and C. Qi, "Hallucinating face by position-patch," *Pattern Recognit.*, vol. 43, no. 6, pp. 2224–2236, 2010.

[38] C. Jung, L. Jiao, B. Liu, and M. Gong, "Position-patch based face hallucination using convex optimization," *IEEE Signal Process. Lett.*, vol. 18, no. 6, pp. 367–370, Jun. 2011.

[39] Z. Wang, R. Hu, S. Wang, and J. Jiang, "Face hallucination via weighted adaptive sparse regularization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 5, pp. 802–813, May 2014.

[40] Z.-Y. Wang, Z. Han, R.-M. Hu, and J.-J. Jiang, "Noise robust face hallucination employing Gaussian–Laplacian mixture model," *Neurocomputing*, vol. 133, pp. 153–160, Jun. 2014.

[41] J. Jiang, R. Hu, Z. Wang, and Z. Han, "Noise robust face hallucination via locality-constrained representation," *IEEE Trans. Multimedia*, vol. 16, no. 5, pp. 1268–1281, Aug. 2014.

[42] J. Jiang, C. Chen, J. Ma, Z. Wang, Z. Wang, and R. Hu, "SRLSP: A face image super-resolution algorithm using smooth regression with local structure prior," *IEEE Trans. Multimedia*, to be published.

[43] J. Jiang, J. Ma, C. Chen, X. Jiang, and Z. Wang, "Noise robust face image super-resolution through smooth sparse representation," *IEEE Trans. Cybern.*, to be published.

[44] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal $\iota 1$-norm solution is also the sparsest solution," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, Mar. 2006.

[45] D. Arpit, G. Srivastava, and Y. Fu, "Locality-constrained low rank coding for face recognition," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 1687–1690.

[46] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, May 2011.

[47] J. Yang, J. Qian, L. Luo, F. Zhang, and Y. Gao. [2014]. "Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes." [Online]. Available: https://arxiv.org/abs/1405.1207

[48] L. Luo, J. Yang, J. Qian, and J. Yang, "Nuclear norm regularized sparse coding," in *Proc. Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 1834–1839.

[49] J. Chen, J. Yang, L. Luo, J. Qian, and W. Xu, "Matrix variate distribution-induced sparse representation for robust image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2291–2300, Oct. 2015.

[50] J. Qian, L. Luo, J. Yang, F. Zhang, and Z. Lin, "Robust nuclear norm regularized regression for face recognition with occlusion," *Pattern Recognit.*, vol. 48, no. 10, pp. 3145–3159, Oct. 2015.

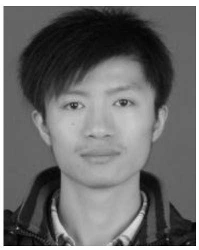
[51] F. Zhang, J. Yang, Y. Tai, and J. Tang, "Double nuclear norm-based matrix decomposition for occluded image recovery and background modeling," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1956–1966, Jun. 2015.

[52] J.-F. Cai, E. J. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.

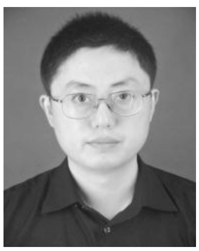
[53] W. Gao *et al.*, "The CAS-PEAL large-scale Chinese face database and baseline evaluations," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 1, pp. 149–161, Jan. 2008.

[54] C. E. Thomaz and G. A. Giraldi, "A new ranking method for principal components analysis and its application to face image analysis," *Image Vis. Comput.*, vol. 28, no. 6, pp. 902–913, 2010.

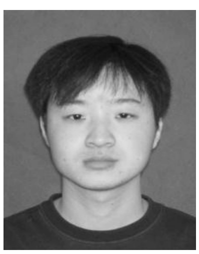
[55] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.



GUANGWEI GAO (M'15) received the B.S. degree in information and computation science from Nanjing Normal University, Nanjing, China, in 2009, and the Ph.D. degree in pattern recognition and intelligence systems from the Nanjing University of Science and Technology, Nanjing, in 2014. In 2011 and 2013, he was an Exchange Student with the Department of Computing, Hong Kong Polytechnic University. He is currently an Assistant Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications. His research interests include face recognition, face hallucination, and biometrics.

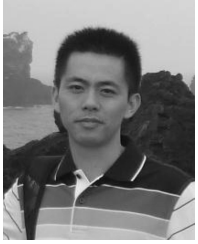


XIAO-YUAN JING received the Ph.D. degree in pattern recognition and artificial intelligence from the Nanjing University of Science and Technology in 1998. He was a Professor with the Department of Computer, Shenzhen Research Student School, Harbin Institute of Technology, in 2005, and a Research Fellow with the Department of Computing, Hong Kong Polytechnic University. He is currently a Professor with the State Key Laboratory of Software Engineering, School of Computer, Wuhan University, and the College of Automation, Nanjing University of Posts and Telecommunications, China. His research interests include pattern recognition, machine learning, software engineering, and artificial intelligence. He has authored over 70 papers in the international conferences and journals, such as CVPR, IJCAI, AAAI, ICSE, FSE, ICME, IEEE TIP, TCSVT, TSMC-B, and PR.



PU HUANG received the B.S. and M.S. degrees in computer applications from Yangzhou University, China, in 2007 and 2010, respectively, and the Ph.D. degree in pattern recognition and intelligent systems from the Nanjing University of Science and Technology, China, in 2014. He is currently a Lecturer with the School of Computer Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. His current research interests include pattern recognition, computer vision, and machine learning.

QUAN ZHOU received the Ph.D. degree in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2013. He is currently an Associate Professor with the College of Telecommunications and Information engineering, Nanjing University of Posts and Telecommunications. His research interests include computer vision and pattern recognition.



SONGSONG WU (M'16) received the B.S. degree in information and computing sciences from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2004, the M.S. degree in measuring and testing technologies and instruments from Nanjing Forestry University, Nanjing, in 2007, and the Ph.D. degree in pattern recognition and intelligence systems from the Nanjing University of Science and Technology, Nanjing, in 2012. He is currently a Lecturer with the Nanjing University of Posts and Telecommunications. His research interests include pattern recognition, machine learning, image processing and analysis, and data compression and retrieval.

DONG YUE (SM'08) received the Ph.D. degree from the South China University of Technology, Guangzhou, China, in 1995. He is currently a Professor and the Dean with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China, and a Changjiang Professor with the Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan, China. His current research interests include analysis and synthesis of networked control systems, multiagent systems, optimal control of power systems, and Internet of Things. He has authored over 100 papers in international journals, domestic journals, and international conferences. He is currently an Associate Editor of the IEEE Control Systems Society Conference Editorial Board and the *International Journal of Systems Science*.