

Low-Resolution Face Recognition via Coupled Locality Preserving Mappings

Bo Li, Hong Chang, *Member, IEEE*, Shiguang Shan, *Member, IEEE*, and Xilin Chen, *Member, IEEE*

Abstract—Practical face recognition systems are sometimes confronted with low-resolution face images. Traditional two-step methods solve this problem through employing super-resolution (SR). However, these methods usually have limited performance because the target of SR is not absolutely consistent with that of face recognition. Moreover, time-consuming sophisticated SR algorithms are not suitable for real-time applications. To avoid these limitations, we propose a novel approach for LR face recognition without any SR preprocessing. Our method based on coupled mappings (CMs), projects the face images with different resolutions into a unified feature space which favors the task of classification. These CMs are learned through optimizing the objective function to minimize the difference between the correspondences (i.e., low-resolution image and its high-resolution counterpart). Inspired by locality preserving methods for dimensionality reduction, we introduce a penalty weighting matrix into our objective function. Our method significantly improves the recognition performance. Finally, we conduct experiments on publicly available databases to verify the efficacy of our algorithm.

Index Terms—Coupled locality preserving mappings, face recognition, low-resolution.

I. INTRODUCTION

THE performance of a real-world face recognition system usually declines when the input face images are degraded seriously, such as low-resolution (LR) with size of only 12×12 pixels. This is a critical problem for surveillance circumstances. Compared with high-resolution (HR) images, these LR images lose some discriminative details across different persons.

Intuitively, recovering the lost information of LR face images first is a promising solution for achieving better performance. In fact, most existing “two-step” methods of LR face recognition employ a preprocessing of SR as the first step. Subsequently, the super-resolved face images are passed to the second step for recognition. During the past decade, many SR methods are proposed to predict the corresponding HR image from a single

Manuscript received March 26, 2009; revised July 27, 2009. First published September 04, 2009; current version published October 07, 2009. This work was performed at the Institute of Computing Technology (ICT) and supported by Natural Science Foundation of China under Contracts 60803084, 60872124, and U0835005, and also by the National Basic Research Program of China (973 Program) under Contract 2009CB320902. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Deniz Erdogmus.

B. Li is with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China. (e-mail: bli@jdl.ac.cn).

H. Chang, S. Shan and X. Chen are with the Key Laboratory of Intelligent Information Processing, and the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190, China (e-mail: hchang@jdl.ac.cn; sgshan@jdl.ac.cn; xlchen@jdl.ac.cn)

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2009.2031705

LR image [7] or multiple LR ones [6]. Most real-world systems select the rapid and simple interpolation methods, e.g., bilinear, cubic, and spline. Recently, learning-based super-resolution (LSR) methods [4], [5], [7], [12] attract more attention because of their outstanding performance. Freeman *et al.* [7] propose a patch-wise Markov network learned from the training set as the SR prediction model. Baker and Kanade [4] propose “face hallucination” to infer the HR face image from an input LR one based on face priors. Liu *et al.* [12] propose to integrate a holistic model and a local model for SR reconstruction. Chang *et al.* [5] propose a method based on LLE [16] which has fairly good performance. The following LSR methods are proposed based on the early works mentioned above.

Recently, some algorithms attempt to avoid explicit SR in the image domain. The approach performs SR reconstruction in the eigenface domain has been investigated in [8]. The researchers propose a video-based LR face recognition approach with implicit SR [1]. More recently, P. Hennings-Yeomans *et al.* [10] propose a joint objective function that integrates the aims of super-resolution and face recognition. Compared with two-step methods, this method improves the recognition rate. However, the speed of this algorithm is slow, even for the speed-up version, because the optimization procedure needs to be executed for each test image with regard to each enrollment. In this letter, we focus on the problem of single LR face recognition.

To overcome the limitations of previous methods, we propose a new efficient method for LR face recognition without any SR preprocessing. According to the aim of recognition, we learn the coupled projections to map the face images with different resolutions into a unified feature space and carry out the recognition step in the new space. An overview of our method is shown in Fig. 1. Inspired by the works of [3], [11], we involve the objective of preserving the local structure of the original space into our optimization to achieve better performance. With proper constraints, the formulated optimization problem could be solved in an analytical close-form. These good characteristics make our procedure efficient and effective in both offline training phase and online testing phase. The efficacy of our method will be illustrated in the experiments on the FERET.

II. LOW-RESOLUTION FACE RECOGNITION

A. Problem Statement

Under many cases, the task of LR face recognition may reduce to find a proper distance measure between a LR face image and a HR one, i.e., $d_{ij} = D(\mathbf{l}_i, \mathbf{h}_j)$. Here, $\mathbf{l}_i \in \mathbb{R}^m$, $i = 1, 2, \dots, N_p$ and $\mathbf{h}_j \in \mathbb{R}^M$, $j = 1, 2, \dots, N_g$ represent the feature vectors of the LR query images in the probe set and the HR ones enrolled in the gallery set. Obviously, some common distances (e.g., Euclidean distance) can not be calculated directly since the dimensions of LR and HR features are not equal (i.e., $m < M$). To solve this problem, the traditional two-step

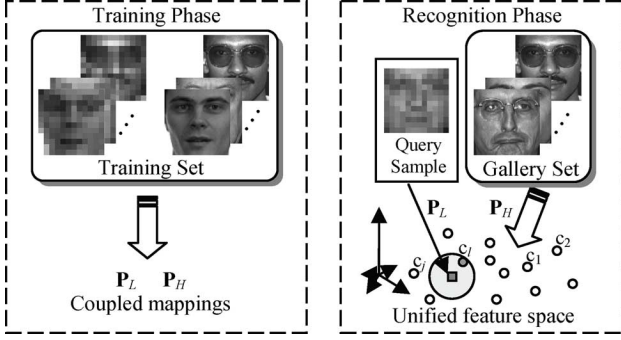


Fig. 1. Overview of our method via coupled locality preserving mappings.

methods employ super-resolution functions to project the LR image into the target HR space, and then calculate the distance in the HR space as:

$$d_{ij} = D(f_{sr}(\mathbf{l}_i), \mathbf{h}_j). \quad (1)$$

In (1), $\mathbf{y} = f_{sr}(\mathbf{x})$ denotes the SR functions, where $\mathbf{y} \in \mathbb{R}^M$ is the HR estimate for the LR image $\mathbf{x} \in \mathbb{R}^m$. Obviously, the performance of the first SR step is very important for the two-step method. Unfortunately, most existing SR methods are not in the most proper way to improve the performance of the subsequent recognition step.

Different from previous methods, our approach attempts to project the data points in the original HR and LR feature spaces into a unified feature space by coupled mappings: one for LR feature vectors, $f_L : \mathbb{R}^m \rightarrow \mathbb{R}^d$; the other for HR ones, $f_H : \mathbb{R}^M \rightarrow \mathbb{R}^d$. d represents the dimensionality of the new feature space. Then, we can measure the distance by

$$d_{ij} = D(f_L(\mathbf{l}_i), f_H(\mathbf{h}_j)). \quad (2)$$

Note that the two-step methods can be seen as a special case of our algorithm with settings of $f_L(\mathbf{x}) = f_{sr}(\mathbf{x})$, $f_H(\mathbf{x}) = \mathbf{x}$ and $d = M$.

Now the key problem of our method becomes to pursue the ideal unified feature space. For face recognition, we expect that the projection of each LR image and the corresponding HR projected image should be as close as possible in the new feature space. This principle could be formulated by the following objective function:

$$J(f_L, f_H) = \sum_{i=1}^{N_t} \|f_L(\mathbf{l}_i) - f_H(\mathbf{h}_i)\|^2. \quad (3)$$

N_t indicates the number of the training images. To minimize the above objective function, we can get coupled mappings.

More specifically, we consider two matrices \mathbf{P}_L and \mathbf{P}_H with sizes of $m \times d$ and $M \times d$ to specify the mapping functions

as $f_L(\mathbf{x}) = \mathbf{P}_L^T \mathbf{x}$ and $f_H(\mathbf{x}) = \mathbf{P}_H^T \mathbf{x}$, respectively. In this context, (3) could be reformulated as

$$J(\mathbf{P}_L, \mathbf{P}_H) = \sum_{i=1}^{N_t} \|\mathbf{P}_L^T \mathbf{l}_i - \mathbf{P}_H^T \mathbf{h}_i\|^2. \quad (4)$$

In the following Section II-B, we derive the analytical solution for this minimization problem.

B. Optimization Solution

Using some deductions of linear algebra [2], we can rewrite the objective function (4) into a new form as (5), shown at the bottom of the page. Here, $\text{Tr}(\cdot)$ is the matrix trace operator. The training set in the original LR and HR feature spaces is denoted as $\mathbf{L} = [\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_{N_t}]$ and $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_t}]$, respectively.

Furthermore, the objective function is rewritten as

$$\begin{aligned} J(\mathbf{P}_L, \mathbf{P}_H) &= \text{Tr} \left(\begin{bmatrix} \mathbf{P}_L \\ \mathbf{P}_H \end{bmatrix}^T \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}^T \begin{bmatrix} \mathbf{P}_L \\ \mathbf{P}_H \end{bmatrix} \right). \end{aligned} \quad (6)$$

Let $\mathbf{P} = \begin{bmatrix} \mathbf{P}_L \\ \mathbf{P}_H \end{bmatrix}$, $\mathbf{Z} = \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}$ and $\mathbf{A} = \begin{bmatrix} \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{bmatrix}$, where \mathbf{I} is the identity matrix. Consequently, we obtain a concise form as

$$J(\mathbf{P}_L, \mathbf{P}_H) = \text{Tr}(\mathbf{P}^T \mathbf{Z} \mathbf{A} \mathbf{Z}^T \mathbf{P}). \quad (7)$$

Finally, we add the constraints to achieve scaling and translation invariance, and solve the optimization problem by minimizing

$$J(\mathbf{P}_L, \mathbf{P}_H) \text{ s.t. } \mathbf{P}^T \mathbf{Z} \mathbf{Z}^T \mathbf{P} = \mathbf{I} \text{ and } \mathbf{P}^T \mathbf{Z} \mathbf{1} = \mathbf{0} \quad (8)$$

where $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T$ is the vector of ones with $2N_t$ entries. Let $\mathbf{E} = \mathbf{Z} \mathbf{A} \mathbf{Z}^T$ and $\mathbf{F} = \mathbf{Z} \mathbf{Z}^T$; the solution to the optimization problem with respect to \mathbf{P} could be given by the second to $(d+1)$ st smallest generalized eigenvectors \mathbf{p} of $\mathbf{E} \mathbf{p} = \lambda \mathbf{F} \mathbf{p}$, where both sizes of matrices \mathbf{E} and \mathbf{F} are $(m+M) \times (m+M)$.

A more efficient solution for this eigendecomposition problem is proposed in the following. Let us expand $\mathbf{E} \mathbf{p} = \lambda \mathbf{F} \mathbf{p}$ to obtain two linear equations as follows:

$$\mathbf{L} \mathbf{L}^T \mathbf{p}_L - \mathbf{L} \mathbf{H}^T \mathbf{p}_H = \lambda \mathbf{L} \mathbf{L}^T \mathbf{p}_L \quad (9)$$

$$\text{and } \mathbf{H} \mathbf{H}^T \mathbf{p}_H - \mathbf{H} \mathbf{L}^T \mathbf{p}_L = \lambda \mathbf{H} \mathbf{H}^T \mathbf{p}_H. \quad (10)$$

Consequently, from (10), we get

$$\mathbf{p}_H = \frac{(\mathbf{H} \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{L}^T \mathbf{p}_L}{(1 - \lambda)}. \quad (11)$$

Finally, substituting (11) into (9) we obtain

$$J(\mathbf{P}_L, \mathbf{P}_H) = \text{Tr} \left(\mathbf{P}_L^T \mathbf{L} \mathbf{L}^T \mathbf{P}_L + \mathbf{P}_H^T \mathbf{H} \mathbf{H}^T \mathbf{P}_H - \mathbf{P}_L^T \mathbf{L} \mathbf{H}^T \mathbf{P}_H - \mathbf{P}_H^T \mathbf{H} \mathbf{L}^T \mathbf{P}_L \right). \quad (5)$$

$$\mathbf{LH}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}\mathbf{L}^T\mathbf{p}_L = (1 - \lambda)^2\mathbf{L}\mathbf{L}^T\mathbf{p}_L. \quad (12)$$

Let $\mathbf{U} = \mathbf{LH}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}\mathbf{L}^T$, $\mathbf{V} = \mathbf{L}\mathbf{L}^T$; (12) is still a generalized eigendecomposition problem with eigenvalue $(1 - \lambda)^2$ and the corresponding eigenvector \mathbf{p}_L . The sizes of matrices \mathbf{U} and \mathbf{V} are $m \times m$. Obviously the computational complexity of $\mathbf{U}\mathbf{p}_L = (1 - \lambda)^2\mathbf{V}\mathbf{p}_L$ is lower than $\mathbf{E}\mathbf{p} = \lambda\mathbf{F}\mathbf{p}$. After \mathbf{p}_L is calculated, \mathbf{p}_H could be computed by (11) directly. From the formulations above, we can see that our objective agrees with canonical correlation analysis (CCA) [9], which is a powerful tool in multivariate analysis.

Note that, since $\mathbf{H}\mathbf{H}^T$ is not always invertible (e.g., $M > N_t$), we need to perform a regularization operation, say $\mathbf{H}\mathbf{H}^T + \tau\mathbf{I}$. τ is set to a small positive value that is smaller than the smallest nonzero eigenvalue (such as $\tau = 10^{-6}$). The same regularization is adopted for $\mathbf{L}\mathbf{L}^T$.

C. Coupled Locality Preserving Mappings

Inspired by the works of [3], [11], we introduce the locality preserving objective into our minimization criterion as follows:

$$J(\mathbf{P}_L, \mathbf{P}_H) = \sum_{i=1}^{N_t} \sum_{j=1}^{N_t} \|\mathbf{P}_L^T \mathbf{l}_i - \mathbf{P}_H^T \mathbf{h}_j\|^2 \mathbf{W}_{ij}. \quad (13)$$

Similar with the work of [15], the penalty weighting matrix \mathbf{W} (with size of $N_t \times N_t$) serves to preserve the local relationship between data points in the original feature spaces. \mathbf{W} is defined on the neighborhoods of the data points as follows:

$$\mathbf{W}_{ij} = \begin{cases} \exp\left(-\frac{\|\mathbf{h}_i - \mathbf{h}_j\|^2}{\sigma^2}\right) & j \in N(i) \\ 0 & \text{otherwise} \end{cases}. \quad (14)$$

Here, σ is the parameter specifying the width of Gaussian function, which is set to $\alpha \sum_{i,j} \|\mathbf{h}_i - \mathbf{h}_j\|^2 / N_t^2$ where α is a scale parameter. $N(i)$ contains the indices of k nearest neighbors of \mathbf{h}_i . In this letter, we select the neighborhood system defined in the HR feature space, because HR feature is considered to have more discriminative information [10].

Then (13) could be written in matrix form as

$$J(\mathbf{P}_L, \mathbf{P}_H) = \text{Tr} \left(\mathbf{P}_L^T \mathbf{L} \mathbf{D}^L \mathbf{L}^T \mathbf{P}_L + \mathbf{P}_H^T \mathbf{H} \mathbf{D}^H \mathbf{H}^T \mathbf{P}_H \right) - \text{Tr} \left(\mathbf{P}_L^T \mathbf{L} \mathbf{W} \mathbf{H}^T \mathbf{P}_H + \mathbf{P}_H^T \mathbf{H} \mathbf{W}^T \mathbf{L}^T \mathbf{P}_L \right). \quad (15)$$

Here, the matrices \mathbf{D}^L and \mathbf{D}^H are defined based on the weight matrix \mathbf{W} as $\mathbf{D}_{ii}^L = \sum_j \mathbf{W}_{ij}$ and $\mathbf{D}_{jj}^H = \sum_i \mathbf{W}_{ij}$, respectively. The off diagonal entries in the two matrices are all zeroes.

Similar deduction with (5) to (7), we obtain

$$J(\mathbf{P}_L, \mathbf{P}_H) = \text{Tr} \left(\mathbf{P}^T \mathbf{Z} \mathbf{A} \mathbf{Z}^T \mathbf{P} \right), \quad (16)$$

where $\mathbf{P} = \begin{bmatrix} \mathbf{P}_L \\ \mathbf{P}_H \end{bmatrix}$, $\mathbf{Z} = \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}$ and $\mathbf{A} = \begin{bmatrix} \mathbf{D}^L & -\mathbf{W} \\ -\mathbf{W}^T & \mathbf{D}^H \end{bmatrix}$. Finally, we solve the optimization problem by minimizing

$$J(\mathbf{P}_L, \mathbf{P}_H) \text{ s.t. } \mathbf{P}^T \mathbf{Z} \mathbf{Z}^T \mathbf{P} = \mathbf{I} \text{ and } \mathbf{P}^T \mathbf{Z} \mathbf{1} = \mathbf{0} \quad (17)$$

TABLE I
LOW-RESOLUTION FACE RECOGNITION BASED ON CMS/CLPMs

Input:

LR face images, corresponding HR ones in training set: $\mathbf{L} \in \mathfrak{R}^{m \times N_t}$, $\mathbf{H} \in \mathfrak{R}^{M \times N_t}$.

Enrolled HR face images and the corresponding class labels in Gallery set:

$\mathbf{H}^g \in \mathfrak{R}^{M \times N_g}$ and $\mathbf{c}^g \in \mathfrak{R}^{N_g}$.

A query LR face image $\mathbf{l}_i^q \in \mathfrak{R}^m$.

Output:

The class label c_i^q for the query image.

Offline phase:

Learn coupled mapping matrices \mathbf{P}_L and \mathbf{P}_H by the algorithms proposed in Section II B or C.

Transform the enrolled HR face images \mathbf{H}^g into the new unified feature space by $\tilde{\mathbf{H}}^g = \mathbf{P}_H^T \mathbf{H}^g$.

Online phase:

Transform the query LR face image \mathbf{l}_i^q into the new feature space by $\tilde{\mathbf{l}}_i^q = \mathbf{P}_L \mathbf{l}_i^q$.

Infer the class label c_i^q of $\tilde{\mathbf{l}}_i^q$ from $\tilde{\mathbf{H}}^g$ and \mathbf{c}^g by the nearest neighbor classifier.

where \mathbf{I} is the identity matrix with size of $d \times d$, and $\mathbf{1}$ is the vector of ones with $2N_t$ entries. Similarly, let $\mathbf{E} = \mathbf{Z}\mathbf{A}\mathbf{Z}^T$ and $\mathbf{F} = \mathbf{Z}\mathbf{Z}^T$, the solution to the optimization problem with respect to \mathbf{P} could be given by the 2-nd to $(d+1)$ st smallest generalized eigenvectors \mathbf{p} of $\mathbf{E}\mathbf{p} = \lambda\mathbf{F}\mathbf{p}$. Note that, this formulation can not be solved in the efficient way as described from (9) to (12).

D. Our Algorithm

After the coupled mapping matrices are computed, our method projects the LR query images and the HR enrollments into a new unified feature space and performs face recognition. The proposed method of the LR face recognition has two phases. One is the offline phase which includes learning the coupled mappings and performing transformations on enrolled HR images. The other is online phase which consists of query LR image transformation and feature matching. In Table I, we summarize our method of LR face recognition.

From the algorithm, we can see that our approach does not involve any SR or optimization process in the online recognition phase. Therefore, our method is more suitable for real-time systems.

III. EXPERIMENTS

We carry out our experiments on the FERET [14] face database. The training set has 1002 frontal face images from 429 persons. We evaluate our method based on the standard gallery (1196 images) and the probe set "fabf" (1195 images). In all experiments, the HR face images with size of 72×72 pixels are aligned with the positions of two eyes. The LR images with size of 12×12 pixels are generated by the operation of smoothing and downsampling.

In our experiments, we compare our methods with S²R² (with PCA features) [10] and several two-step algorithms, which include HR-PCA/LDA (using the HR query images), Spline-PCA/LDA (first restoring the LR query images by spline interpolation), and LSR-PCA/LDA (first restoring the LR query images by LSR [5] with five neighbors and patch size of 8×8

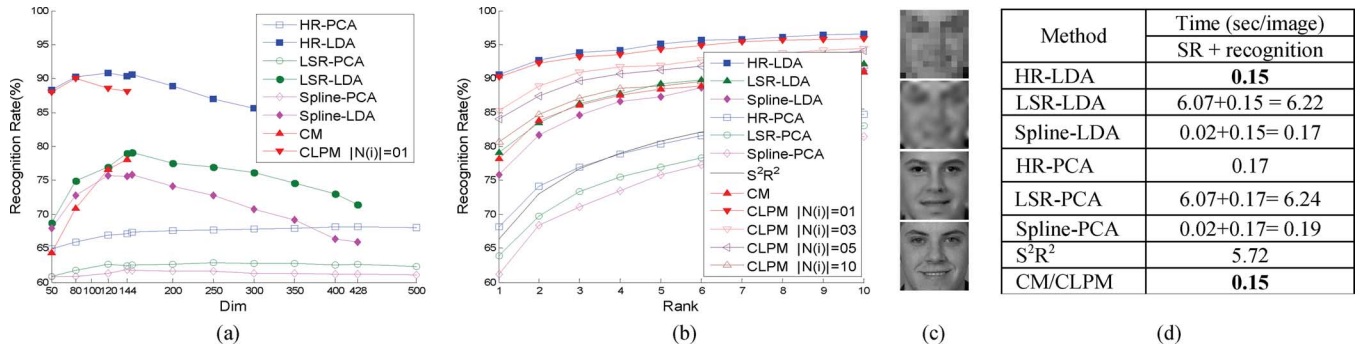


Fig. 2. (a) Recognition results with different dimensionalities. (b) Cumulative recognition results. (c) Super-resolution examples. From top to bottom: LR image, spline interpolation result, LSR [5] result, and original HR face image. (d) Time requirements.

pixels). There are some super-resolved face examples shown in Fig. 2(c). In the experiments, the distribution of each pixel is adjusted to mean zero, variance one. In our algorithm based on CLPMs, we set $\alpha = 11$.

In Fig. 2(a), the recognition rate with different feature dimensionalities are drawn. Our methods based on coupled mappings (CMs) with 144-D features and coupled locality preserving mappings (CLPMs) with 80-D features achieve the recognition rate of 78.2% and 90.1%, respectively. On the other hand, the rates achieved by Spline-PCA and LSR-PCA with 400-D features are 61.8% and 62.5%. The rates of Spline-LDA and LSR-LDA with 150-D features are only 75.8% and 79.1% and still lower than the proposed CLPMs.

According to the dimensionality selected above, we plot the cumulative recognition results of the different methods in Fig. 2(b). From the results, we see that our method outperform the methods of LR face recognition involved in comparison, several two-step methods and S^2R^2 . Specially, the result of the proposed CLPMs ($|N(i)| = 1$) is even close to that of HR-LDA.

Besides, we show the results of CLPMs using different numbers of the nearest neighbors in Fig. 2(b). We can see that the proposed algorithm achieves the best performance when $|N(i)| = 1$, and the recognition rate descends as $|N(i)|$ ascends. Because there are only two or four images in most classes in the training set of FERET, when $|N(i)| > 1$, our semi-supervised learning in CPLMs must involve the wrong neighbor information. In this case, the performance declines.

In Fig. 2(d), we list the mean time requirements for recognizing per query face image on a 3.4 GHz CPU. We can see that our method runs as fast as face recognition using HR face images directly. Meanwhile, the two-step methods are slower than the proposed method because of the expensive preprocessing SR. S^2R^2 is also time-consuming because of its optimization scheme.

IV. CONCLUSIONS

In this letter, we propose a novel approach to solve the problem of the LR face recognition. Different from the traditional two-step methods, our algorithm projects the face images with different resolutions into a unified feature space through coupled mappings. Under the aim of recognition, we propose a minimization objective function w.r.t. the coupled mappings to guarantee the discriminative ability in the new space. Our time-saving method, without any SR, is more suitable for real-time applications. The experimental results on FERET show that

our methods can achieve satisfactory performance. Besides, it is worth noting that it is straightforward to extend our method to incorporate supervised information (such as class labels) and to employ nonlinear mappings by kernel techniques.

The performance of real-world LR face recognition will be degraded by the factors of noise, geometric distortion and so on. Considering these factors with LR simultaneously and using more stable features (i.e., Gabor wavelets) instead of original intensity features in our method will be pursued in future work.

REFERENCES

- [1] O. Arandjelovic and R. Cipolla, "A manifold approach to face recognition from low quality video across illumination and pose using implicit super-resolution," in *Proc. IEEE Int. Conf. Computer Vision*, 2007.
- [2] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
- [3] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, pp. 1373–1396, 2003.
- [4] S. Baker and T. Kanade, "Hallucinating faces," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, 2000, pp. 83–88.
- [5] H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2004, pp. 275–282.
- [6] M. Elad and A. Feuer, "Super-resolution reconstruction of image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, pp. 817–834, 1999.
- [7] W. Freeman, E. Pasztor, and O. Carmichael, "Learning low-level vision," *Int. J. Comput. Vis.*, vol. 40, pp. 25–47, 2000.
- [8] B. Gunturk, A. Batur, Y. Altunbasak, M. Hayes, and R. Mersereau, "Eigenface-domain super-resolution for face recognition," *IEEE Trans. Image Process.*, vol. 12, pp. 597–606, 2003.
- [9] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 312–377, 1936.
- [10] P. Hennings-Yeomans, S. Baker, and B. Kumar, "Simultaneous super-resolution and feature extraction for recognition of low-resolution faces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008, pp. 1–8.
- [11] X. He and P. Niyogi, "Locality preserving projections," in *Advances of Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2003.
- [12] C. Liu, H. Shum, and C. Zhang, "A two-step approach to hallucinating faces: Global parametric model and local nonparametric model," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2001, pp. 192–198.
- [13] Z. Lin, J. He, X. Tang, and C. Tang, "Limits of learning-based super-resolution algorithms," in *Proc. IEEE Int. Conf. Computer Vision*, 2007.
- [14] P. Philips, H. Moon, P. Pauss, and S. Rivzvi, "The feret evaluation methodology for face-recognition algorithms," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1997, pp. 137–143.
- [15] T. Sun and S. Chen, "Locality preserving CCA with applications to data visualization and pose estimation," *Image Vis. Comput.*, vol. 25, no. 5, pp. 531–543, 2007.
- [16] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, 2000.