

Research Article

Kernel Coupled Cross-Regression for Low-Resolution Face Recognition

Zhifei Wang,^{1,2} Zhenjiang Miao,^{1,2} Yanli Wan,^{1,2} and Zhen Tang^{1,2}

¹ Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China

² Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China

Correspondence should be addressed to Zhifei Wang; yichen0201@163.com

Received 25 March 2013; Accepted 28 May 2013

Academic Editor: Rafael Martinez-Guerra

Copyright © 2013 Zhifei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Low resolution (LR) in face recognition (FR) surveillance applications will cause the problem of dimensional mismatch between LR image and its high-resolution (HR) template. In this paper, a novel method called kernel coupled cross-regression (KCCR) is proposed to deal with this problem. Instead of processing in the original observing space directly, KCCR projects LR and HR face images into a unified nonlinear embedding feature space using kernel coupled mappings and graph embedding. Spectral regression is further employed to improve the generalization performance and reduce the time complexity. Meanwhile, cross-regression is developed to fully utilize the HR embedding to increase the information of the LR space, thus to improve the recognition performance. Experiments on the FERET and CMU PIE face database show that KCCR outperforms the existing structure-based methods in terms of recognition rate as well as time complexity.

1. Introduction

Face recognition (FR) has been widely studied for decades due to its great potential applications. Many technologies focused on dealing with complex conditions such as aging, occlusion, disguise, and variations in pose, illumination, and expression. Although the recognition accuracy of face recognition in controlled environments with cooperative subjects is satisfactory, the performance in real applications such as surveillance and mobile service is still an unsolved problem partially due to low-resolution (LR) image quality [1]. With the growing installation of embedded cameras in many places, there are increasing demands for face recognition in interaction and surveillance applications from interactive processing cameras in mobile communications, small-scale stand-alone cameras in banks and supermarkets, to large-scale multiple networked close-circuit televisions in public streets. In such cases, subjects are far from cameras without any restriction, and face regions tend to be small or poor quality. This issue is called low-resolution face recognition (LR FR).

Gallery images are generally assumed to be high resolution (HR) in practical face recognition applications.

Therefore, LR will obviously cause the problem of dimensional mismatch between HR gallery images and LR probe ones in the special applications such as criminals monitoring and mobile communication with their HR images in network databases. To deal with the problem, three general ways can be considered as follows. The first way is down-scaling HR gallery image, which seems to be a feasible solution for the mismatching problem. Unfortunately, it drastically reduces the amount of available information, especially the high-frequency information mainly for recognition when the resolution is very low such as 12×12 or even 8×8 . Secondly, upscaling LR probe image, such as interpolation, is conventionally adopted in most of the subspace-based face recognition methods. Some researchers tried to propose some resolution-robust feature representation methods, which is to extract the discriminative information from LR images (or upsampled versions) directly. The landmark works are RQCr color feature [2], local frequency texture descriptor [3], and novel kernel correlation feature analysis [4]. However, they are sensitive to resolutions more or less and fail to complete full resolution robustness. For further refined solution, the widely used super resolution (SR) or face hallucination [5] technique can be employed to estimate

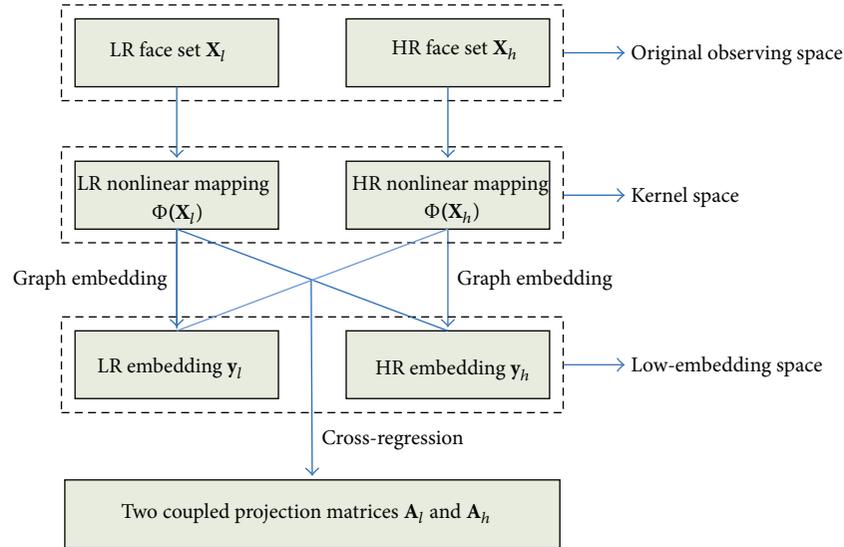


FIGURE 1: Overview of kernel coupled cross-regression.

HR faces from LR ones. However, they usually require a lot of images which belong to the same scene with precise alignment and also need large time consumption. In addition, most of these methods aimed to improve face appearance but failed to optimize face images from recognition perspective. Recently, a few attempts were made to achieve these two criteria under very LR case [6]. Also some researchers studied simultaneous super resolution and recognition for providing a promising way to solve this problem [7, 8], though they also need high-computational cost.

The third way is unified feature space, also called inter-resolution space, which is used to project HR gallery images and LR probe ones into a common space [9]. Here we call it as structure-based method, which is an essential way to solve the mismatch problem. It aims to build the holistic structure to facilitate direct matching between them from classification perspective. For example, Li et al. [10] proposed a general framework called coupled mappings (CMs) to project LR and HR face images into a unified feature space. The problem with this method is that it offers poor classification ability. Thus, they further introduced locality weight relationships [11] into CMs and proposed coupled locality preserving mappings (CLPMs), which significantly improved the performance. However, how to efficiently solve the eigenvalue decomposition in the projection matrix computation remains unsolved in CLPMs, which is the common problem in the structure-based method. Furthermore, CLPMs is sensitive to the parameters of the penalty weighting matrix such as scale and neighbor. Subsequently, other methods are developed to improve the classification performance of CMs. Deng et al. [12] considered two other relationships in LR classification (LR versus LR and HR versus HR) besides LR versus HR used in CMs. Zhou et al. [13] proposed simultaneous discriminant analysis method by introducing linear discriminant relationships [14] between intraclass scattering and interclass scattering into CMs. Ren et al. [15] integrated nonlinear

kernel trick into CMs/CLPMs, which transferred LR/HR pairs from the input space to the kernel-embedding space. Although these methods are proven to be discriminative and efficient, they still require much more computational costs for system implementation.

Inspired by the works in [15, 16], we propose a novel structure-based method to deal with LR FR, called kernel coupled cross-regression (KCCR), which adopts two mapping matrices to project LR and HR face images represented by nonlinear kernel into a unified feature space for the final feature matching, shown in Figure 1. KCCR firstly transforms LR and HR images from the original observing space into the low-embedding space obtained by linear graph embedding technique [17]. Meanwhile, the features of images are represented by nonlinear kernel tricks to increase the computational power of the linear embeddings. Spectral regression (SR) [18] is then introduced to generalize the model and reduce the time complexity. Furthermore, cross-regression is developed to fully utilize the HR embedding to increase the information of the LR space, by which two more discriminant coupled projection matrices are obtained. The transformations can be learned offline from the training images. During testing, the input images are transformed using the two coupled projection matrices and then matching. Therefore, KCCR that combines cross-regression with kernel is more efficient and effective and can achieve better performance than the existing structure-based methods as well as some subspace learning methods or kernel-based methods. It will be shown in the experiments with the FERET and CMU PIE face database.

There are three main contributions in this paper.

- (1) We build the transformations in the low-embedding space, which is totally different from other existing structure-based methods performing in the original observing space, and the time complexity is reduced from cubic to linear by spectral regression.

- (2) We creatively propose cross-regression technique to improve the discriminant ability, which helps to extract more high frequency or local information among HR samples for LR matching.
- (3) Since nonlinear kernel functions have been successfully used to build more efficient coupled mapping models in [15, 19], we also introduce the kernel trick into our cross-regression model and further increase its recognition performance.

The rest of this paper is organized as follows. Section 2 presents our proposed KCCR method. In Section 3, we give an analysis of generality about KCCR and their connections to linear graph embedding and coupled mappings. Experimental results on the publicly used databases are shown in Section 4. Section 5 concludes the paper.

2. Kernel Coupled Cross-Regression

LR results in the problem of dimensional mismatch between HR gallery images and LR probe ones, which will directly lose a lot of useful information for optimal classification. Thus, it is feasible to project LR/HR pairs into a unified feature space. Moreover, such mappings are required to be most discriminative for feature matching, and the computation should be efficient. Our KCCR method is just designed to achieve these goals based on kernel coupled mappings and spectral regression.

2.1. Problem Description. The idea of coupled mappings aims to obtain two projection matrices \mathbf{A}_h and \mathbf{A}_l with sizes of $m \times d$ and $M \times d$ to realize the minimization problem as shown in (1), where x_{h_i} and x_{l_i} represent HR and LR face images, respectively and N is the total number of samples.

$$J(\mathbf{A}_h, \mathbf{A}_l) = \sum_{i=1}^N \|x_{h_i}^T \mathbf{A}_h - x_{l_i}^T \mathbf{A}_l\|^2. \quad (1)$$

CMs model [10] aims to obtain a unified feature space at the pixel or feature level. Therefore, the performance of CMs essentially depends on the contents or features of face images. In other words, CMs model itself does not bring any new information for LR FR. Here, we introduce a low-dimension embedding of \mathbf{X} , denoted as \mathbf{y} , to transform LR and HR into a low-embedding space based on spectral regression instead of an observing pixel space. Then it will obtain more information about LR samples by exploring the relationships among HR samples and improve the efficiency of CMs model drastically:

$$\mathbf{X}_h^T \mathbf{A}_h \longrightarrow \mathbf{y} \longleftarrow \mathbf{X}_l^T \mathbf{A}_l. \quad (2)$$

2.2. Graph Embedding and Regression Framework. Given a graph $G = \{\mathbf{X}, \mathbf{W}\}$ with vertex set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ and similarity matrix $\mathbf{W} \in \mathfrak{R}^{N \times N}$; each vertex \mathbf{x}_i represents a face image and W_{ij} represents the weight of the edge joining vertices \mathbf{x}_i and \mathbf{x}_j . The graph G can then be used for characterizing various statistical or geometric properties of

the data set. Therefore, graph embedding [17] is to represent each vertex of the graph as a low-dimensional vector that preserves similarities between the data point pairs, and regression is used for solving the optimization problem of graph embedding, which reduces the time complexity. Based on these points, spectral regression [18] is introduced into CMs, which makes the CMs model more efficient for LR FR.

Let $\mathbf{y} = \{y_1, y_2, \dots, y_N\}^T$ be the low-dimensional embeddings of the vertex set \mathbf{X} , which means if vertices \mathbf{x}_i and \mathbf{x}_j are close then y_i and y_j are also close. The optimal \mathbf{y} is given by minimizing

$$\mathbf{y} = \arg \min_{\mathbf{y}} \sum_{i,j} \|y_i - y_j\|^2 W_{ij} = \arg \min_{\mathbf{y}} 2\mathbf{y}^T \mathbf{L} \mathbf{y}, \quad (3)$$

where $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is the graph Laplacian matrix and \mathbf{D} is a diagonal matrix with $D_{ii} = \sum_j W_{ij}$. A constraint $\mathbf{y}^T \mathbf{D} \mathbf{y} = 1$ is then imposed for removing the arbitrary scaling factor in the embedding:

$$\mathbf{y} = \arg \min_{\mathbf{y}^T \mathbf{D} \mathbf{y} = 1} \mathbf{y}^T \mathbf{L} \mathbf{y} = \arg \min_{\mathbf{y}^T \mathbf{D} \mathbf{y} = 1} \frac{\mathbf{y}^T \mathbf{L} \mathbf{y}}{\mathbf{y}^T \mathbf{D} \mathbf{y}} = \arg \max_{\mathbf{y}^T \mathbf{D} \mathbf{y} = 1} \frac{\mathbf{y}^T \mathbf{W} \mathbf{y}}{\mathbf{y}^T \mathbf{D} \mathbf{y}}. \quad (4)$$

The optimal \mathbf{y} can be obtained by solving the maximum eigenvalue problem:

$$\mathbf{W} \mathbf{y} = \lambda \mathbf{D} \mathbf{y}. \quad (5)$$

From the classification perspective, a relationship f between y_i and \mathbf{x}_i should be introduced, that is, $y_i = f(\mathbf{x}_i) = \mathbf{x}_i^T \mathbf{A}$ via a linear function. Equation (4) can be written as

$$\mathbf{A} = \arg \max_{\mathbf{A}} \frac{\mathbf{y}^T \mathbf{W} \mathbf{y}}{\mathbf{y}^T \mathbf{D} \mathbf{y}} = \arg \max_{\mathbf{A}} \frac{\mathbf{A}^T \mathbf{X} \mathbf{W} \mathbf{X}^T \mathbf{A}}{\mathbf{A}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{A}}. \quad (6)$$

The optimal \mathbf{A} can be obtained by solving the maximum eigenvalue problem:

$$\mathbf{X} \mathbf{W} \mathbf{X}^T \mathbf{A} = \lambda \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{A}. \quad (7)$$

This framework is called linearization of graph embedding (LGE). With different choices of similarity matrix \mathbf{W} , LGE will lead to many popular linear subspace learning methods such as linear discriminant analysis (LDA) [14] and locality preserving projection (LPP) [11]. For LDA and LPP, \mathbf{W} is defined as follows.

LDA. Suppose the t th class has m_t samples:

$$W_{ij} = \begin{cases} \frac{1}{m_t} & \text{if } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ both belong to the } t\text{th class,} \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

LPP. Let $N_k(\mathbf{x}_i)$ denote the set of k nearest neighbors of \mathbf{x}_i .

$$W_{ij} = \begin{cases} \exp^{-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2} & \text{if } \mathbf{x}_i \in N_k(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in N_k(\mathbf{x}_i), \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

2.3. *The Proposed KCCR Method.* Inspired by the work of Lei and Li [16] adopting two embeddings to describe heterogeneous face data, we introduce \mathbf{y}_l and \mathbf{y}_h to represent the embeddings of LR and HR face data, respectively. In the embedding solution, LDA aims to preserve the global relationships between data points in favor of LR face space mainly containing low-frequency information, while LPP aims to preserve the local relationships between data points in favor of HR face space mostly depending on high-frequency features. Therefore, compared with the representing LR and HR space both by LPP in CLPMs [10] and by LDA in SDA [13], \mathbf{y}_l and \mathbf{y}_h are obtained by LDA and LPP, respectively, in our CCR method for making full use of the combination of global and local information in data. Furthermore, in order to utilize the HR embedding by LPP to increase the information of the LR space by LDA, coupled cross-regression is designed for exploring the relationships between LR/HR data sets $\{\mathbf{X}_l, \mathbf{X}_h\}$ and their crossed low-embedding $\{\mathbf{y}_h, \mathbf{y}_l\}$. With linear assumption, the problem is simplified to find the two projection matrices \mathbf{A}_l and \mathbf{A}_h for LR/HR data sets as follows:

$$\begin{aligned} \mathbf{y}_h &\leftarrow \mathbf{X}_l^T \mathbf{A}_l, \\ \mathbf{y}_l &\leftarrow \mathbf{X}_h^T \mathbf{A}_h. \end{aligned} \quad (10)$$

To solve the solution \mathbf{A}_l and \mathbf{A}_h , the classical least square method is used for minimizing the cost. Besides, due to limited training samples, the variance of the estimation may be large, and thus the estimation is not reliable. To overcome this problem, penalties are imposed on the norm of \mathbf{A}_l and \mathbf{A}_h as in ridge regression. The regularization parameter $\alpha, \beta \geq 0$ controls the trade-off between the bias and variance of the estimation. Based on these discussions, the objective function is formulated as follows:

$$\begin{aligned} J(\mathbf{A}_l, \mathbf{A}_h) &= \frac{1}{N} \|\mathbf{y}_h - \mathbf{X}_l^T \mathbf{A}_l\|^2 + \frac{1}{N} \|\mathbf{y}_l - \mathbf{X}_h^T \mathbf{A}_h\|^2 \\ &+ \alpha \|\mathbf{A}_l\|^2 + \beta \|\mathbf{A}_h\|^2. \end{aligned} \quad (11)$$

By requiring the derivatives of the objective function with respect to \mathbf{A}_l and \mathbf{A}_h vanish, we get

$$\begin{aligned} &\begin{cases} (\mathbf{X}_l \mathbf{X}_l^T + N\alpha \mathbf{I}_l) \mathbf{A}_l = \mathbf{X}_l \mathbf{y}_h, \\ (\mathbf{X}_h \mathbf{X}_h^T + N\beta \mathbf{I}_h) \mathbf{A}_h = \mathbf{X}_h \mathbf{y}_l, \end{cases} \\ \Rightarrow &\begin{cases} \mathbf{A}_l = (\mathbf{X}_l \mathbf{X}_l^T + N\alpha \mathbf{I}_l)^{-1} \mathbf{X}_l \mathbf{y}_h, \\ \mathbf{A}_h = (\mathbf{X}_h \mathbf{X}_h^T + N\beta \mathbf{I}_h)^{-1} \mathbf{X}_h \mathbf{y}_l. \end{cases} \end{aligned} \quad (12)$$

Suppose that the Euclidean space \mathbf{R}^n is mapped to a Hilbert space \mathbf{H} through a nonlinear mapping function $\Phi: \mathbf{R}^n \rightarrow \mathbf{H}$. Let $\Phi(X)$ denote the data matrix in the Hilbert space. Here, we replace LR and HR images with their nonlinear feature vectors; that is, $x_{l_i} \rightarrow \Phi_{l_i} = \Phi(x_{l_i})$ and $x_{h_i} \rightarrow \Phi_{h_i} = \Phi(x_{h_i})$, which are induced by a kernel function $k(\cdot, \cdot)$; that is, $k(x_{l_i}, x_{h_i}) = \Phi(x_{l_i})^T \Phi(x_{h_i})$. Here $k(\cdot, \cdot)$ is the kernel function such as the typical Gaussian kernel

$k(x_{l_i}, x_{h_i}) = \exp(-\|x_{l_i} - x_{h_i}\|^2 / 2\sigma^2)$. Thus, the objective function $J(\mathbf{A}_l, \mathbf{A}_h)$ can be replaced as follows:

$$\begin{aligned} J(\mathbf{A}_l, \mathbf{A}_h) &= \frac{1}{N} \|\mathbf{y}_h - \Phi(\mathbf{X}_l)^T \mathbf{A}_l\|^2 + \frac{1}{N} \|\mathbf{y}_l - \Phi(\mathbf{X}_h)^T \mathbf{A}_h\|^2 \\ &+ \alpha \|\mathbf{A}_l\|^2 + \beta \|\mathbf{A}_h\|^2. \end{aligned} \quad (13)$$

Equation (12), therefore, can be changed into

$$\begin{aligned} &\begin{cases} (\Phi(\mathbf{X}_l) \Phi(\mathbf{X}_l)^T + N\alpha \mathbf{I}_l) \mathbf{A}_l = \Phi(\mathbf{X}_l) \mathbf{y}_h, \\ (\Phi(\mathbf{X}_h) \Phi(\mathbf{X}_h)^T + N\beta \mathbf{I}_h) \mathbf{A}_h = \Phi(\mathbf{X}_h) \mathbf{y}_l, \end{cases} \\ \Rightarrow &\begin{cases} \mathbf{A}_l = (\Phi(\mathbf{X}_l) \Phi(\mathbf{X}_l)^T + N\alpha \mathbf{I}_l)^{-1} \Phi(\mathbf{X}_l) \mathbf{y}_h, \\ \mathbf{A}_h = (\Phi(\mathbf{X}_h) \Phi(\mathbf{X}_h)^T + N\beta \mathbf{I}_h)^{-1} \Phi(\mathbf{X}_h) \mathbf{y}_l, \end{cases} \end{aligned} \quad (14)$$

where \mathbf{I}_l and \mathbf{I}_h are the identity matrices with the size of $m \times m$ and $M \times M$, respectively, $\Phi(\mathbf{X}_l) = [\Phi(x_{l_1}), \dots, \Phi(x_{l_n})]$ and $\Phi(\mathbf{X}_h) = [\Phi(x_{h_1}), \dots, \Phi(x_{h_n})]$.

By using the matrix manipulation in kernel ridge regression [20], we have

$$\begin{aligned} &(\Phi(\mathbf{X}_l) \Phi(\mathbf{X}_l)^T + N\alpha \mathbf{I}_l)^{-1} \Phi(\mathbf{X}_l) \\ &= \Phi(\mathbf{X}_l) (\Phi(\mathbf{X}_l)^T \Phi(\mathbf{X}_l) + N\alpha \mathbf{I}_l)^{-1}, \\ &(\Phi(\mathbf{X}_h) \Phi(\mathbf{X}_h)^T + N\beta \mathbf{I}_h)^{-1} \Phi(\mathbf{X}_h) \\ &= \Phi(\mathbf{X}_h) (\Phi(\mathbf{X}_h)^T \Phi(\mathbf{X}_h) + N\beta \mathbf{I}_h)^{-1}, \end{aligned} \quad (15)$$

and, therefore,

$$\begin{aligned} \mathbf{A}_l &= \Phi(\mathbf{X}_l) (\Phi(\mathbf{X}_l)^T \Phi(\mathbf{X}_l) + N\alpha \mathbf{I}_l)^{-1} \mathbf{y}_h \\ &= \Phi(\mathbf{X}_l) (K_l + N\alpha \mathbf{I}_l)^{-1} \mathbf{y}_h, \\ \mathbf{A}_h &= \Phi(\mathbf{X}_h) (\Phi(\mathbf{X}_h)^T \Phi(\mathbf{X}_h) + N\beta \mathbf{I}_h)^{-1} \mathbf{y}_l \\ &= \Phi(\mathbf{X}_h) (K_h + N\beta \mathbf{I}_h)^{-1} \mathbf{y}_l, \end{aligned} \quad (16)$$

where $K_{l,j} = \Phi(x_{l_i})^T \Phi(x_{l_j}) = k(x_{l_i}, x_{l_j})$ and $K_{h,j} = \Phi(x_{h_i})^T \Phi(x_{h_j}) = k(x_{h_i}, x_{h_j})$. Naturally, corresponding to d eigenvectors $\{y_{l_i}\}_{i=1}^d$ and $\{y_{h_i}\}_{i=1}^d$ to (5), we can obtain d coupled cross mapping matrices $\mathbf{A}_l = \{\mathbf{A}_{l_1}, \mathbf{A}_{l_2}, \dots, \mathbf{A}_{l_d}\}$ and $\mathbf{A}_h = \{\mathbf{A}_{h_1}, \mathbf{A}_{h_2}, \dots, \mathbf{A}_{h_d}\}$ to project LR/HR pairs into a unified discriminative space and thus for classification. It is worth noting that KCCR and CCR in the current version are specifically designed for single resolution case. For practical face recognition applications where the resolution of probe faces is likely to be changed, the two projection matrices for multiple resolutions case, for example, 16×16 , 12×12 and 8×8 , are needed for handling this problem. However, how to determine variable m in \mathbf{A}_l is a very difficult problem in this context. One possible solution is to select different $\mathbf{A}_l/\mathbf{A}_h$

TABLE 1: The KCCR method used for LR FR.

LR/HR training images: $\mathbf{X}_l^t \in \mathfrak{R}^{m \times N}$ and $\mathbf{X}_h^t \in \mathfrak{R}^{M \times N}$
HR gallery images: $\mathbf{X}_h^g \in \mathfrak{R}^{M \times N^g}$ with class labels: $\mathbf{c}^g \in \mathfrak{R}^C$
LR probe images: $\mathbf{X}_l^p \in \mathfrak{R}^{m \times N^p}$ with class labels: $\mathbf{c}^p \in \mathfrak{R}^C$
Offline training phase:
(1) Learning low embeddings: $\mathbf{y}_l \in \mathfrak{R}^{N \times d}$ and $\mathbf{y}_h \in \mathfrak{R}^{N \times d}$
(2) Computing kernel coupled cross projection matrices: $\mathbf{A}_l \in \mathfrak{R}^{m \times d}$ and $\mathbf{A}_h \in \mathfrak{R}^{M \times d}$
(3) Transforming HR gallery images: $\mathbf{y}_h^g = \Phi(\mathbf{X}_h^g)^T \mathbf{A}_h$ $= \Phi(\mathbf{X}_h^g)^T \Phi(\mathbf{X}_h)(K_h + N\beta \mathbf{I}_h)^{-1} \mathbf{y}_l$ $= k_h(K_h + N\beta \mathbf{I}_h)^{-1} \mathbf{y}_l,$ where $k_{h_{i,j}} = \Phi(x_{h_i}^g)^T \Phi(x_{h_j}) = k(x_{h_i}^g, x_{h_j})$
Online testing phase:
(1) Transforming LR probe images: $\mathbf{y}_l^p = \Phi(\mathbf{X}_l^p)^T \mathbf{A}_l$ $= \Phi(\mathbf{X}_l^p)^T \Phi(\mathbf{X}_l)(K_l + N\alpha \mathbf{I}_l)^{-1} \mathbf{y}_h$ $= k_l(K_l + N\alpha \mathbf{I}_l)^{-1} \mathbf{y}_h,$ where $k_{l_{i,j}} = \Phi(x_{l_i}^p)^T \Phi(x_{l_j}) = k(x_{l_i}^p, x_{l_j})$
(2) Performing feature matching: comparing \mathbf{y}_l^p and \mathbf{y}_h^g by nearest neighbour classifier and obtaining \mathbf{c}^p from \mathbf{c}^g

pairs according to the nearest resolution. For example, when the probe image has the resolution of 14×14 , 16×16 or 12×12 can be used to determine m for constructing coupled mappings. Similarly, 12×12 or 8×8 can be used when the probe resolution is between 12×12 and 8×8 .

2.4. Algorithm Summarization. Similar to CMs, the proposed KCCR also consists of two main phases, which is summarized in Table 1. One is offline training phase, including three parts: learning the two low-embeddings \mathbf{y}_l and \mathbf{y}_h , computing the two coupled cross projection matrices \mathbf{A}_l and \mathbf{A}_h and transforming HR gallery images into the unified space. The other is online testing phase, including two parts: transforming LR probe images into the unified space and performing feature matching.

3. Generality Analysis about KCCR

In this section, we present an analysis of generality about KCCR and CCR and their connections to linear graph embedding (LGE) methods and coupled mappings (CMs) models. We divide the baseline methods into four groups: original, kernel, original CMs, and kernel CMs, which are shown in Table 2. In the original methods, LDA [14] and LPP [11] are two typical LGE methods, and spectral regression (SR) [18] introduces regression into LGE framework to solve the optimization problem. Obviously, the kernel tricks are induced into LGE and SR methods, that is kernel-based methods, such as kernel LDA [21], kernel LPP [22], and kernel SR [23], respectively.

The former two groups are correspondingly used to describe the relationships between samples in CMs models,

TABLE 2: Different methods for evaluations.

	Original	Kernel	Original CMs	Kernel CMs
LDA	LDA [14]	KDA [21]	SDA [13]	/
LPP	LPP [11]	KLPP [22]	CLPMs [10]	CKE [15]
SR	SR [18]	SRKDA [23]	CCR	KCCR

and then some representative structure-based LR FR methods are proposed, that is, original CMs and kernel CMs. In the original CMs group, coupled locality preserving mappings (CLPMs) [10] and simultaneous discriminant analysis (SDA) [13] improve CMs based on LPP and LDA, respectively. In this paper, we use SR technique instead of LPP or LDA to describe the relationships and propose coupled cross-regression (CCR) method. The recently proposed coupled kernel embedding (CKE) method [15] and our kernel CCR method introduced the kernel trick into CLPMs and CCR model, respectively, and expanded the ideas of nonlinear dimensionality reduction for LR FR problem in the fourth group. Naturally, SDA method can also be improved by kernel, which is the future work.

From the above analyses, we can know that the original version and the kernel one of LGE methods can be successfully embedded into CMs models to discover the relationships between LR and HR feature space. It is worth noting that our KCCR method based on cross-regression and kernel technique shows more efficiency than other baseline methods.

4. Experimental Results and Discussions

To validate the effectiveness and efficiency of the proposed method for LR FR, 12×12 and 8×8 probe images in FERET [24] and CMU PIE [25] face database are adopted for experiments to compare with the original and kernel linear graph embedding (LGE) methods and the original and kernel coupled mappings (CMs) methods, which are shown in Table 2.

4.1. Databases and Settings. The training set of FERET contains 1002 frontal face images from 429 subjects. The experiments are based on the standard gallery (1196 images) and the probe set ‘‘fafb’’ (1195 images). For CMU PIE, 1428 frontal view face images with neutral expression and illumination variations in 68 subjects (21 images per subject) were selected in our experiment. For each subject, 3 images are randomly selected for training and the remaining 18 images for testing. Here, only one image with no flash illumination is taken as gallery, while the probe set contains the images with different illuminations (17 images per subject).

One sample in FERET and CMU PIE with the HR size 72×72 and 64×64 , respectively, and the corresponding LR size 12×12 and 8×8 are shown in Figure 2. In all experiments, HR face images are aligned by the positions of two eyes and normalized to zero mean and unit variance. Then the preprocessed HR face images are synthetically translated by random subpixel amounts, blurred with a Gaussian filter, and down-sampled to generate LR face images, as [5] does.

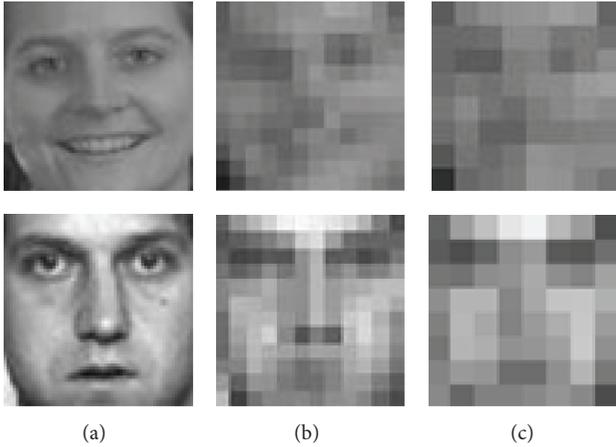


FIGURE 2: Face images in FERET (first row) and CMU PIE (second row). (a) HR with 72×72 (FERET) and 64×64 (CMU PIE). (b) LR with 12×12 . (c) LR with 8×8 .

4.2. Results and Analyses. We evaluate the performance of our KCCR and CCR with some representative LGE methods, specifically, NN-LDA/LPP/SRLDA (LDA [14]/LPP [11]/SRLDA [18] with first restoring the LR probe images by nearest neighbour (NN) interpolation) and HR-LDA (LDA with the HR probe images, denoted as HR in the following sections). Furthermore, the kernel versions of these LGE methods are used for comparison, such as KDA [21], KLPP [22], and SRKDA [23]. Meanwhile, other representative state-of-the-art structure-based methods for LR FR are taken for evaluation, that is, CLPMs [10], SDA [13], CKE [15].

From the recognition performances in FERET with the probe size 12×12 and 8×8 shown in Figures 3 and 4, respectively, the recognition results of different methods are obviously classified into four levels. That is, HR method and the kernel versions of CMs methods such as KCCR and CKE rank the first level, and the original CMs methods such as CLPMs, SDA, and CCR follow subsequently. Then the kernel versions of LGE methods with LR probe images rank the third. Finally, LGE methods with LR probe images obtain the worst results. This phenomenon is more obvious for 8×8 .

For 12×12 , although all CMs methods obtain relative close performance, our KCCR method achieves the Rank-1 recognition rate of 90.8%, which still outperforms the other kernel CMs method, that is, CKE (90.6%), and CCR (90.5%) is better than other original CMs methods such as CLPMs (89.7%) and SDA (90.2%). In particular, the result of KCCR is even close to HR method (90.9%). On the other hand, the recognition rates of kernel and original LGE methods are shown as follows: KDA (78.6%), KLPP (74.1%), SRKDA (80.3%), LDA (71.7%), LPP (65.2%), and SRLDA (74.6%), respectively. In general, the kernel LGE methods are relatively better than the original ones but inferior to the proposed KCCR and CCR.

For 8×8 , although the performances of all methods more or less decline, our KCCR and CCR methods still achieve 80.2% and 75.9%, which outperform CKE (78.8%), CLPMs (68.7%), and SDA (73.5%), respectively, and the performance

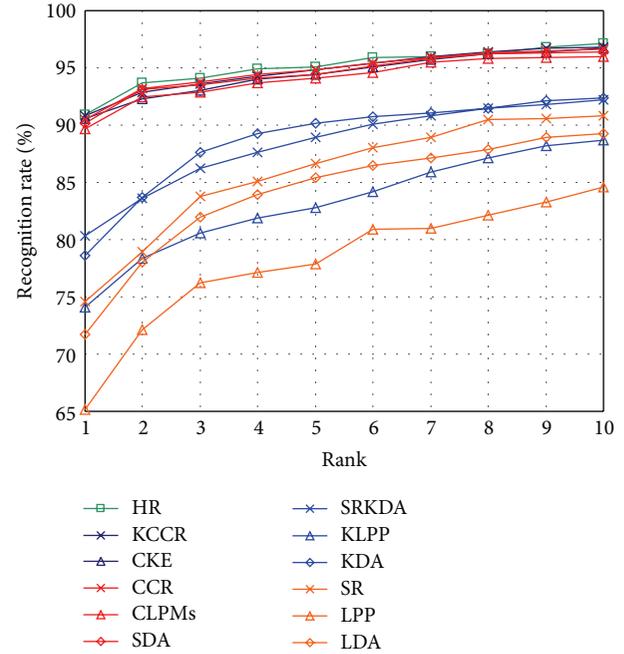


FIGURE 3: Cumulative recognition results for FERET database with the probe size 12×12 .

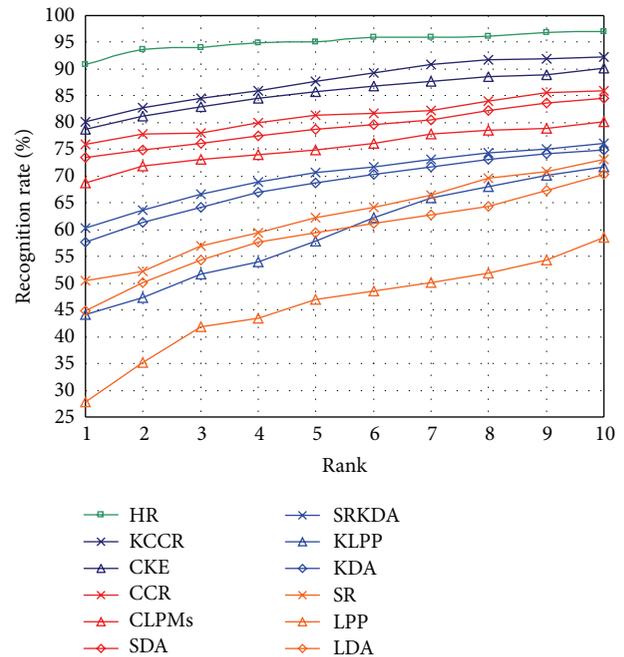


FIGURE 4: Cumulative recognition results for FERET database with the probe size 8×8 .

gap between them is larger than that in 12×12 . It implies that KCCR and CCR have a great potential in solving very LR problem. Compared with CMs methods, the performances of LGE methods, that is, KDA (57.7%), KLPP (44.1%), SRKDA (60.3%), LDA (44.9%), and SRLDA (50.4%), are degraded as a result of LR, especially LPP which only achieves the

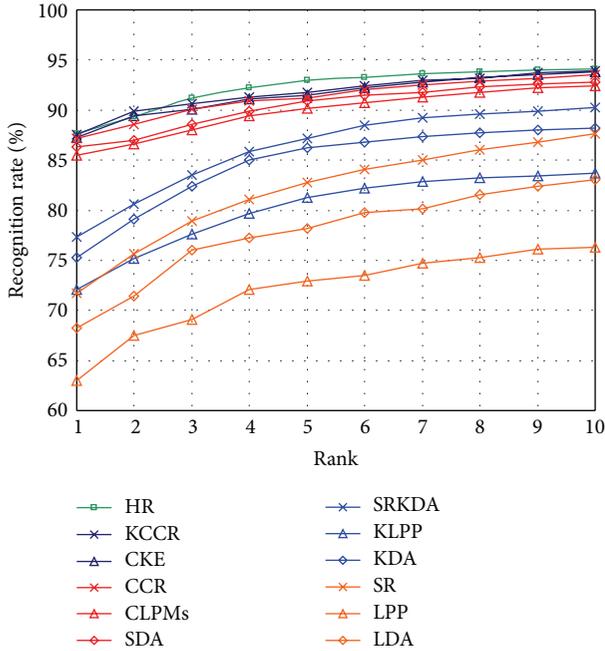


FIGURE 5: Cumulative recognition results for CMU PIE database with the probe size 12×12 .

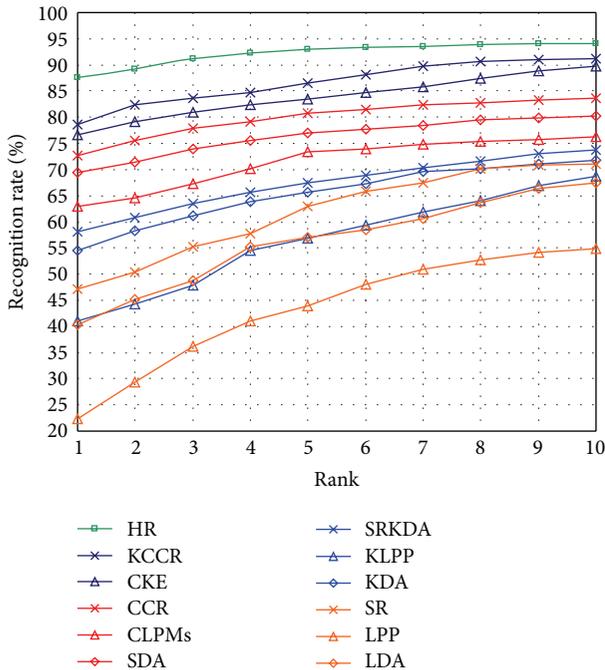


FIGURE 6: Cumulative recognition results for CMU PIE database with the probe size 8×8 .

accuracy of 27.8%. Therefore, our KCCR and CCR methods are competitive for LR FR task.

Furthermore, Figures 5 and 6 show the recognition performances in CMU PIE with the probe size 12×12 and 8×8 , respectively. The results of different methods are similar to those in FERET. That is, our KCCR & CCR method still

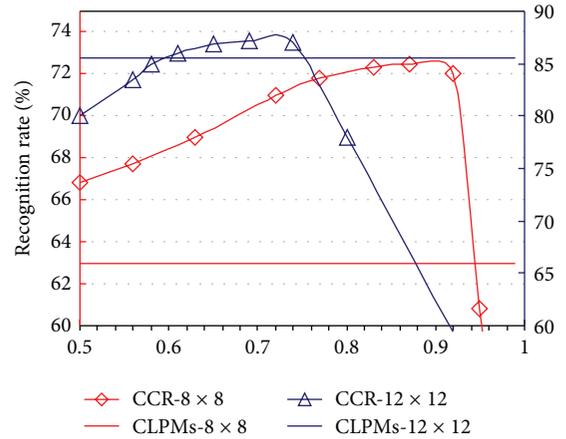


FIGURE 7: Model selections for KCCR and CCR. The curve shows the recognition rates of CCR with respect to $\tau/(1 + \tau)$ (τ denotes β/α) with two different resolutions 12×12 (blue line) and 8×8 (red line). The two solid lines show the recognition rates of CLPMs.

outperforms the baseline methods, though the performances of all the methods are slightly degraded by different illumination variations involved in CMU PIE. However, the baseline methods are affected much more as the resolution declines.

4.3. Model Selections. The $\alpha, \beta \geq 0$ are two essential parameters in our CCR method which control the smoothness of the estimator under LR and HR cases, respectively. When the sample vectors are linearly independent, CCR provides the same solution as the coupled cross version of LGE as α, β decrease to zero. Also the relationships between α and β determine the performance of the model to some extent. In this subsection, we try to examine the impact of the two parameters on the performance of CCR.

In experiments, we find that the smoothness parameters mainly depend on different resolutions. For example, CCR obtains the best performance as α, β are set to 0.001 and 0.01, respectively, at the resolution of 8×8 , while $\alpha = 0.005$ and $\beta = 0.01$ at 12×12 . Obviously, $\beta \geq \alpha$ is also a possible condition to guarantee that HR samples can provide more information for LR classification in our CCR model. Therefore, we take the performance of CCR as a function of β/α under different resolutions, which is shown in Figure 7. For convenience, the x -axis is plotted as $\tau/(1 + \tau)$ (τ denotes β/α) which is strictly in the interval $[0.5, 1]$. It is easy to see that CCR can achieve significantly better performance than CLPMs over a large range of τ especially at the resolution of 8×8 , which further shows the potential of our CCR method in solving very LR problem. Thus, the parameter selection is not a very crucial problem in our model. The selection of α, β in KCCR model shows similar property to CCR, although they are slightly different. Here, we empirically set α, β as 0.0001, 0.001 at the resolution of 8×8 , and 0.0005, 0.001 at 12×12 .

In addition, how to select appropriate kernels for LR face data is a more important problem in the KCCR, which is still an unsolved problem without any selection criteria in theory for all kernel-based methods. In this paper, Gaussian

TABLE 3: Time complexity of different methods.

Database	FERET(s)	CMU PIE(s)
LDA	0.142	0.122
LPP	0.161	0.134
SRLDA	0.034	0.028
CLPMs	0.586	0.513
SDA	0.604	0.525
CCR	0.098	0.084
KDA	0.452	0.334
KLPP	0.552	0.382
SRKDA	0.086	0.058
CKE	0.648	0.422
KCCR	0.171	0.139

kernel is selected for building kernel representation for fair comparisons in previous experiments. Empirical analysis shows that kernel-based methods are usually sensitive to the selection of the hyperparameters, that is, the width of Gaussian kernel σ . Thus, finding the hyperparameters with a good generalization performance is crucial for a successful application. In our experiments, we adopt the same strategy used in [17] to select the optimal kernel parameter. The parameter σ is set as $\sigma = 2^{(n-10)}\sigma_0, n = 0, 1, \dots, 20$, where σ_0 is the standard deviation of the training samples. The best result is selected from the 21 configurations by cross-validation. The details can be referred as in [17]. However, it is unclear how to choose the optimal σ , which is still a future work for kernel-based models effectively used in LR FR applications.

4.4. Time Complexity Analyses. The average time of each method for recognizing per face images on CPU with 2.33 GHz is shown in Table 3. The properties of time complexity with 12×12 and 8×8 are similar. Here, we just give some analyses with 12×12 . Although the time requirement of the original CMs methods (with $O((M + m)^3)$ time complexity) is generally higher than the original LGE methods (with $O(MN^2 + N^3)$ time complexity), compared with other CMs methods such as CLPMs and SDA, the speed of our CCR method (with $O((M + m)N)$ time complexity) is relatively close to SRLDA and even faster than LDA and LPP. With the introduction of kernel trick, the time complexity ($O(MN^2)$) of kernel LGE method increases slightly compared with the original ones. The same situation occurs at CKE and KCCR with the time complexity of $O((M + m)N^2)$. However, our KCCR combining with the regression framework for LR FR task shows more efficiency than CKE or other baseline methods except for SRLDA and SRKDA.

5. Conclusions

In this paper, a novel structure-based method called kernel coupled cross-regression (KCCR) is proposed for LR FR. Our method combines coupled mappings model and graph embedding analysis to describe the relationships between LR and HR in low-embedding space. Kernel trick is further adopted for describing these relationships in an infinite

dimensional and nonlinear space. Moreover, cross-regression is developed to fully utilize the HR embedding to increase the information of the LR space. As a result, the proposed KCCR method is more discriminative and faster. It outperforms LGE methods with LR probe images, the kernel versions of LGE methods, and other existing structure-based methods such as CLPMs, SDA, and CKE, especially in time complexity.

Acknowledgments

This work is supported by the National Key Technology R&D Program of China (2012BAH01F03), the National Natural Science Foundation of China (60973061), the National Basic Research (973) Program of China (2011CB302203), Ph.D. Programs Foundation of Ministry of Education of China (20100009110004), and China Postdoctoral Science Foundation (2013M530020).

References

- [1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.
- [2] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Color face recognition for degraded face images," *IEEE Transactions on Systems, Man and Cybernetics B*, vol. 39, no. 5, pp. 1217–1230, 2009.
- [3] Z. Lei, T. Ahonen, M. Pietikainen, and S. Z. Li, "Local frequency descriptor for low-resolution face recognition," in *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition and Workshops (FG '11)*, pp. 161–166, Santa Barbara, Calif, USA, March 2011.
- [4] R. Abiantun, M. Sawides, and B. V. K. Vijaya Kumar, "How low can you go? Low resolution face recognition study using Kernel correlation feature analysis on the FRGCv2 dataset," in *Proceedings of the Biometrics Symposium*, Baltimore, Md, USA, September 2006.
- [5] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, 2002.
- [6] W. W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 327–340, 2012.
- [7] P. H. Hennings-Yeomans, S. Baker, and B. V. K. V. Kumar, "Simultaneous super-resolution and feature extraction for recognition of low-resolution faces," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.
- [8] K. Jia and S. G. Gong, "Multi-modal tensor face for simultaneous super-resolution and recognition," in *Proceedings of the International Conference on Computer Vision (ICCV '05)*, vol. 2, pp. 1683–1690, Beijing, China, October 2005.
- [9] S. Biswas, K. W. Bowyer, and P. J. Flynn, "Multidimensional scaling for matching low-resolution face images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 2019–2030, 2012.
- [10] B. Li, H. Chang, S. G. Shan, and X. L. Chen, "Low-resolution face recognition via coupled locality preserving mappings," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 20–23, 2010.
- [11] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces," *IEEE Transactions on Pattern*

- Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328–340, 2005.
- [12] Z. X. Deng, D. Q. Dai, and X. X. Li, “Low-resolution face recognition via color information and regularized coupled mappings,” in *Proceedings of the Chinese Conference on Pattern Recognition (CCPR '10)*, pp. 1–5, Chongqing, China, October 2010.
- [13] C. T. Zhou, Z. W. Zhang, D. Yi, Z. Lei, and S. Z. Li, “Low-resolution face recognition via simultaneous discriminant analysis,” in *Proceedings of the International Joint Conference on Biometrics (IJCB'11)*, pp. 1–6, Washington, DC, USA, October 2011.
- [14] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces versus fisherfaces: recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.
- [15] C.-X. Ren, D.-Q. Dai, and H. Yan, “Coupled kernel embedding for low-resolution face image recognition,” *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3770–3783, 2012.
- [16] Z. Lei and S. Z. Li, “Coupled spectral regression for matching heterogeneous faces,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, vol. 2, pp. 1123–1128, Miami, Fla, USA, June 2009.
- [17] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, and S. Lin, “Graph embedding and extensions: a general framework for dimensionality reduction,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 40–51, 2007.
- [18] D. Cai, X. F. He, and J. W. Han, “Spectral regression for efficient regularized subspace learning,” in *Proceedings of the IEEE 11th International Conference on Computer Vision (ICCV '07)*, pp. 1–8, Rio de Janeiro, Brazil, October 2007.
- [19] W. W. W. Zou and P. C. Yuen, “Learning the relationship between high and low resolution images in kernel space for face super resolution,” in *Proceedings of the 20th International Conference on Pattern Recognition (ICPR '10)*, pp. 1152–1155, Istanbul, Turkey, August 2010.
- [20] S. J. An, W. Q. Liu, and S. Venkatesh, “Face recognition using kernel ridge regression,” in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–7, Minneapolis, Minn, USA, June 2007.
- [21] G. Baudat and F. Anouar, “Generalized discriminant analysis using a kernel approach,” *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.
- [22] J. Cheng, Q. Liu, H. Lu, and Y.-W. Chen, “Supervised kernel locality preserving projections for face recognition,” *Neurocomputing*, vol. 67, no. 1–4, pp. 443–449, 2005.
- [23] D. Cai, X. F. He, and J. W. Han, “Efficient kernel discriminant analysis via spectral regression,” in *Proceedings of the IEEE 7th International Conference on Data Mining (ICDM '07)*, pp. 427–432, Washington, DC, USA, 2007.
- [24] P. Jonathon Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, “The FERET evaluation methodology for face-recognition algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.
- [25] T. Sim, S. Baker, and M. Bsat, “The CMU Pose, Illumination, and Expression Database,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1615–1618, 2003.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

