# LOCALITY-CONSTRAINED MATRIX REGRESSION FOR POSITION-PATCH BASED FACE HALLUCINATION

*Guangwei Gao[1], Xiaoyuan Jing[2], Quan Zhou[3], Songsong Wu[2], Dong Yue[1]*

1. Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China
2. School of Automation, Nanjing University of Posts and Telecommunications, Nanjing, China
3. Key Lab of Ministry of Education for Broad Band Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Nanjing, China

## ABSTRACT

Position-patch based face hallucination approaches have been proposed to replace the probabilistic graph-based or manifold learning-based models recently. In this paper, we propose a novel position-based face hallucination method based on locality-constrained matrix regression (LcMR). LcMR uses nuclear norm to characterize the reconstruction error straightforward, thus preserving the essential structural information of the input. On the other hand, LcMR imposes a locality constraint onto the combination coefficients to reach sparsity and locality simultaneously. The locality constraint can derive an analytical solution to the optimization problem. Moreover, LcMR can be solved using alternating direction method of multipliers. Experimental results demonstrate the superiority of the proposed method over some state-of-the-art approaches.

***Index Terms*—** Face hallucination, matrix regression, position-patch, locality

## 1. INTRODUCTION

The details of facial features are crucial for identifying an individual from surveillance video. However, due to the limitations of surveillance system, such as server storage, and long distance to the interest object, it is sometimes difficult to acquire high-definition face images. Face hallucination, or face super-resolution, is a technology to obtain high-resolution (HR) face images from low-resolution (LR) inputs [1]. Numerous approaches have been proposed 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 [2].

The patch-based methods exhibit the ability to render the local details well. Following the well-known locally linear embedding (LLE) [3], Chang et al. [4] developed a super-resolution method through neighbor embedding (NE). Assuming the LR image patch space and the HR one share the same geometry, the local geometry of the LR patch space is directly mapped to the HR patch space to obtain the HR image patches through linear combination. Different from those patch-based approaches using a fixed number of neighbors for reconstruction, Yang et al. [5] introduced sparse representation technique [6] that can adaptively select the most relevant neighbors to minimize the reconstruction error. They assumed that there exist coupled dictionaries of LR and HR images, which have the same sparse representation for each pair of LR and HR patches.

For a class of highly structured objects, such as human faces, the prior position information can be incorporated into face super-resolution reconstruction. Ma et al. [7] utilized this prior and proposed a position-patch based face hallucination method. The HR image patches are hallucinated using the training patches at the same position by solving a constrained least square problem. However, when the number of the training samples is much larger than the dimension of a patch, the solution of the least square estimation may be not unique. To alleviate this problem, Jung et al. [8] employed sparsity-constrained optimization to replace least square estimation so that the reconstruction results can be improved by adaptively choosing several principal training patches. Li et al. [9] reconstructed the HR image via several similarity constraints, which formulate face hallucination as a local linear filtering process. Jiang et al. [10] incorporated a locality constraint into the least square inversion problem to maintain locality and sparsity simultaneously. Their scheme is capable of capturing the non-linear manifold structure of image patch samples while exploiting the sparse property of the redundant data representation. Recently, with the development of sparse representation techniques, lots of position-patch based methods were proposed to solve the face hallucination problem [11-17].

In this paper, we propose a locality-constrained matrix regression (LcMR) model for position-patch based face hallucination. The method applies the matrix regression to directly compute the combination coefficients (without the matrix-to-vector conversion). In addition, LcMR incorporates a locality constraint into the objective function to reach sparsity and locality simultaneously. Compared with previous methods, where $l_2$ norm (or $l_1$ norm) is used to characterize the reconstruction error, LcMR makes full use of the structure information of image matrices and use the minimal rank of representation residual matrix as a criterion to determine the combination coefficients. On the other hand, the locality constraint can capture fundamental similarities between neighbor patches and derives an analytical solution to the constrained problem, significantly reducing the computational complexity. Experimental results demonstrated that our method can gain better hallucination performance than other state-of-the-art methods.

The rest of the paper is organized as follows. In Section 2, we present our face hallucination method based on LcMR. Section 3 evaluates the performance of the proposed methods on commonly used face hallucination database. Section 4 concludes this paper.

## 2. FACE HALLUCINATION VIA LOCALITY-CONSTRAINED MATRIX REGRESSION

### 2.1. Problem formulation

In this sub-section, all the face image patches are represented in the matrix form. In other words, the patches of the input and training face image at position $(i,j)$ can be denoted as $Y^p(i,j) \in \Re^{d\times d}$ and $A^{mp}(i,j) \in \Re^{d\times d}$ ($p$=1,…, $N$, $m$=1,…, $M$), respectively, where $N$ is the number of the patches and $M$ is the number of the training samples. Then, $Y^p(i,j)$ can be represented linearly using $A^{1p}$,…, $A^{Mp}$:

$$Y^p(i,j) = \sum_{m=1}^{M} A^{mp}(i,j)x_m(i,j) + E, \tag{1}$$

where $x_1(i,j)$,…,$x_M(i,j)$ is the combination coefficients and $E$ is the representation residual. Following the intuition of Locality-constrained Linear Coding (LLC) [18], we incorporate a locality constraint into our objective function:

$$\min_x \left\{ \begin{array}{l} \lambda \| d(i,j) \circ x(i,j) \|_2^2 + \\ \left\| Y^p(i,j) - \sum_{m=1}^{M} A^{mp}(i,j)x_m(i,j) \right\|_* \end{array} \right\}, \tag{2}$$

where $\circ$ denotes point-wise vector product and $d(i,j)$ is a $M$-dimensional vector that penalizes the distance between $Y^p(i,j)$ and each training patch at the same position. It is simply defined by the Euclidean distance:

$$d_m(i,j) = \left\| Y^p(i,j) - A^{mp}(i,j) \right\|_2^2, 1 \le m \le M. \tag{3}$$

### 2.2. Optimization

For the convenience of expression, we use $A^m \in \mathrm{R}^{d\times d}$ ($m$=1,…, $M$) and $Y \in R^{d\times d}$ to denote the training face image patches $A^{mp}(i,j) \in \mathrm{R}^{d\times d}$ and input face image patch $Y^p(i,j) \in \mathrm{R}^{d\times d}$, respectively. The objective function (2) can be represented in the following matrix form:

$$\min_x \lambda \|Dx\|_2^2 + \|F(x) - Y\|_*, \tag{4}$$

where $F(x) = x_1A^1 + x_2A^2 + \ldots + x_MA^M$, $D$ is a $M\times M$ diagonal matrix with entries

$$D_{mm} = d_m(i,j), 1 \le m \le M. \tag{5}$$

The objective function (4) can be converted into the following equivalent problem:

$$\min_{x,E} \lambda \|Dx\|_2^2 + \|E\|_*, \quad s.t. \;\; F(x) - Y = E. \tag{6}$$

The alternating direction method of multipliers (ADMM) [19, 20], also known as the augmented Lagrange multipliers (ALM) method, can been used to solve the nuclear norm optimization problem defined in (6). Here, we briefly describe the process to solve Eq. (6) using ADMM.

The augmented Lagrange function of Eq. (6) is

$$\begin{aligned} L_\mu(x,E) =& \lambda \|Dx\|_2^2 + \|E\|_* + tr\left(Z^T\left(F(x) - Y - E\right)\right) \\ &+ \frac{\mu}{2}\|F(x) - Y - E\|_F^2, \end{aligned} \tag{7}$$

where $\mu$>0 is a penalty parameter, $Z$ is the Lagrange multipliers, $tr(\cdot)$is the trace operator. Following some simple algebraic steps, Eq. (7) can be rewritten as

$$L_\mu(x,E) = \lambda \|Dx\|_2^2 + \|E\|_* + \frac{\mu}{2}\left\| F(x) - Y - E + \frac{1}{\mu}Z \right\|_F^2 - \frac{1}{2\mu}\|Z\|_2^2. \tag{8}$$

**Updating $x$**

Given $E$, the optimization problem can be reformulated as

$$x^{k+1} = \arg\min_x \left\{ \|Dx\|_2^2 + \frac{\mu}{2\lambda}\left\| \left( Y + E - \frac{1}{\mu}Z \right) - F(x) \right\|_F^2 \right\}. \tag{9}$$

Following [4, 7], the solution of a regularized least square in equation (9) can be derived analytically as

$$x^{k+1} = \left( G + \tau D^2 \right) \backslash ones(M,1), \tag{10}$$

Where $ones(M,1)$ is a $M\times 1$ column vector of ones, the operator "\" denotes the left matrix division operation, $\tau$ =$2\lambda/\mu$, and $G$ is the covariance matrix $G = C^TC$ with

$$C = \left( Y + E - \frac{1}{\mu}Z \right) ones(M,1)^T - H, \tag{11}$$

where $H$ = [$Vec(A^1)$, …, $Vec(A^M)$], $Vec(\cdot)$ is the vectorization operator. The final optimal weight is obtained by rescaling to satisfy $\sum_{m=1}^{M} x_m = 1$.

**Updating $E$**

Given $x$, the optimization problem can be reformulated as

$$E^{k+1} = \arg\min_E \frac{1}{\mu}\|E\|_* + \frac{1}{2}\left\| E - \left( F(x) - Y + \frac{1}{\mu}Z \right) \right\|_F^2. \tag{12}$$

The optimal solution can be obtained by the singular value thresholding operator. Specifically, given a matrix $Q$

**Algorithm 1. Locality-constrained matrix regression via ADMM**

**Input:** Training patch matrices $A^1,\ldots,A^M \in R^{p\times q}$ and test patch matrix $Y\in R^{p\times q}$, parameters $\lambda$, the termination condition parameter $\varepsilon$.

1: Update $x^{k+1}$ while fixing the others by Eq. (10);
2: Update $E^{k+1}$ while fixing the others by Eq. (15);
3: Update the multiplies

$$Z^{k+1}=Z^k+\mu\left(F(x^{k+1})-Y-E^{k+1}\right);$$

4: If the termination condition is satisfied, go to **Output**; otherwise repeat 1-3.

**Output:** Optimal coding vector $x^{k+1}$

**Algorithm 2. Face hallucination via LcMR**

**Input:** LR training images $A_L^{1},\ldots,A_L^{M}$, corresponding HR training images $A_H^{1},\ldots,A_H^{M}$, input LR images $Y_L$.

1: **For** each input patch matrix in $Y_L$:
  a) Calculate $d_m(i,j)$ using Eq. (3);
  b) Compute $x^*(i,j)$ using **Algorithm 1**;
  c) Construct the desired HR patch by

$$Y_H^p(i,j)=\sum_{m=1}^{M} A_H^{mp}(i,j)x_m^*(i,j).$$

2: **End for**
3: Integrate all the reconstructed HR patch matrices to form target HR image $Y_H$.

**Output:** The Hallucinated HR face image $Y_H$.

$\in\Re^{p\times q}$ of rank $r$, its singular value decomposition (SVD) is

$$Q=U_{p\times r}\Sigma V_{q\times r}^T,\ \Sigma=diag(\sigma_1,\ldots,\sigma_r), \tag{13}$$

where $\sigma_1,\ldots,\ \sigma_r$ are singular values, $U$ and $V$ are corresponding matrices with orthogonal columns. For a given $\tau$>0, the singular value shrinkage operator [21] $T_\tau(\cdot)$ is defined as follows

$$T_\tau(Q)=U_{p\times r}\mathrm{diag}\left(\left\{\max\left(0,\sigma_j-\tau\right)\right\}_{1\le j\le r}\right)V_{q\times r}^T. \tag{14}$$

The optimal solution of (12) is

$$E^{k+1}=T_{\frac{1}{\mu}}\left(F(x)-Y+\frac{1}{\mu}Z\right). \tag{15}$$

The detailed algorithm via ADMM to solve problem (4) is summarized in **Algorithm 1**.

### 2.3. Face hallucination via Locality-constrained matrix regression

For face hallucination, the training set is composed of LR and HR face image pairs. HR face images are denoted as $A_H^m$ while their LR counterparts are denoted as $A_L^m$ ( $m$=1,…, $M$). The face hallucination task aims to reconstruct the HR face image $Y_H$ from the observed LR face image $Y_L$.

For each LR input image patch matrix in $Y_L$, it is approximated by a linear combination of the LR training patch matrices using LcMR. By replacing the LR training image patch matrices with the corresponding HR ones, a new HR patch matrix can be synthesized. By concatenating all the HR patch matrices to their corresponding positions and averaging pixel values in the overlapping regions, we can obtain the target HR face image. The entire face hallucination algorithm is summarized in **Algorithm 2**.

**Fig. 1:** Some sample images from CAS-PEAL database.

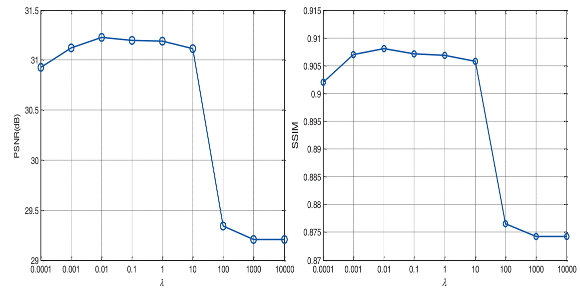


**Fig. 2:** PSNR (left) and SSIM (right) values of results from our method using different $\lambda$.

## 3. EXPERIMENTAL RESULTS AND DISCUSSIONS

### 3.1. Dataset

In this sub-section, we perform experiments of frontal face image hallucination on the CAS-PEAL database [22]. We select a subset that contains 1040 frontal view face images with normal lighting and normal expression. Each individual has only one single image. All the face images are manually aligned using the locations of three points: centers of left and right eyeballs and center of the mouth (some examples are shown in Fig. 1). We crop the region of the faces and normalize the HR images to the size of 128×112. The LR images are formed by smoothing (an averaging filter of size 4×4) and down-sampling the corresponding HR images, thus the LR images have size 32×28.

### 3.2. Parameter analysis

In this sub-section, we evaluate the performance of our method with different $\lambda$. Fig. 2 plots the average PSNR and SSIM values according to different values of $\lambda$. We can see that the regularization parameter plays an important role on the performance of the proposed method: as $\lambda$ increases, more benefits on performance can be gained. However, it should be noted that the value of $\lambda$ could not be set too high. With a proper regularization parameter $\lambda$, our proposed method will gain good results. We recommend to set $\lambda$ between 0.001 and 0.1.

### 3.3. Comparisons with the state-of-the-art methods

We have compared our method with other local patch based face hallucination methods, such as Chang's neighbor embedding method [4], Ma's least square representation method [7], Jung's sparse representation method [8] and Jiang's locality-constrained representation method [10]. For face hallucination, we randomly choose 250 images for training, and 40 images for testing. To pursue the best performance, we tune the parameters for all comparative methods to achieve their best possible results. Specifically, the number of neighbors in Chang's method is set to 50. As for these local patch based methods, we recommend using the size of 12×12 pixels for HR image patch and the overlap between neighbor patches is 12×4 pixels, while corresponding LR image patch size is set to 3×3 pixels with an overlap of 3 pixels.

Fig. 3 shows some representative hallucinated results generated by different methods. Chang's method tends to generate noises on the face images especially on locations around eyes. Ma's method induces smoothness effects on the eyes and mouth. Jung's method enhances the edges and textures on the nostrils and mouth to some extent. However, the ringing effects also appear around the margin of faces. In some cases, the artifacts are even more serious than Ma's method. In contrast, Ma's method and the proposed LcMR generate competitive results with more facial details in the eye, mouth and face contour than the others and meanwhile are much more similar to the original HR faces.

To further assess the objective quality of different methods, we give the quantitative comparisons in terms of PSNR and SSIM as shown in Fig. 4. The average PSNR and SSIM values of the hallucinated results are provided in Table 1. It can be seen that the proposed LcMR gains the best performance for all the test face images. We can learn that: (i) by introducing the position information of face, position-patch based methods can perform better than NE method, which takes no consideration of the position priors; (ii) our method can achieve the best performances in both reconstruction error and visual quality.

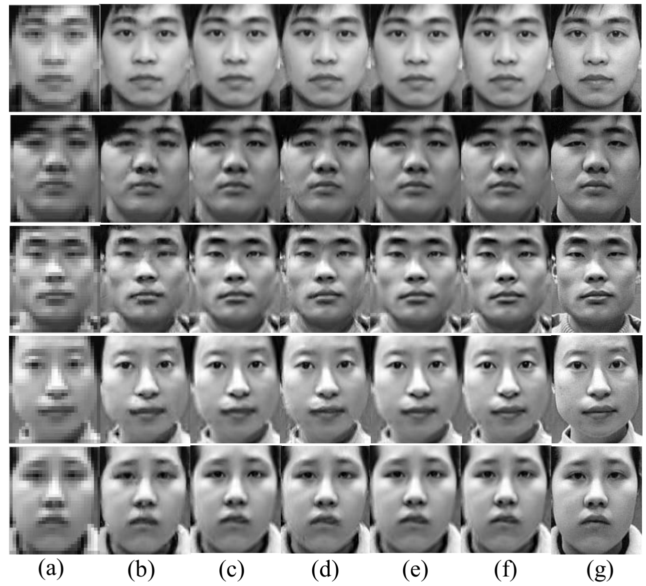


**Fig. 3:** Comparisons of super-resolved results based on different methods on the CAS-PEAL database. From left to right are: (a) the Input LR image; (b) Chang et al.'s method; (c) Ma et al.' method; (d) Jung et al.' method; (e) Jiang et al.' method; (f) the proposed method and (g) the original HR image.

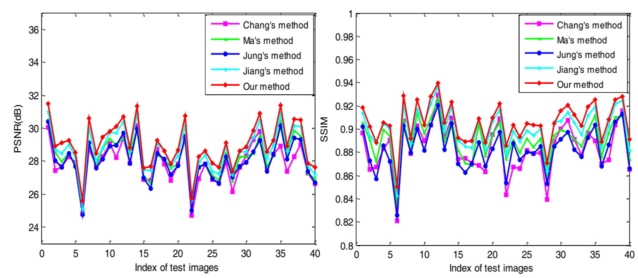


**Fig. 4:** PSNR (left) and SSIM (right) values on the CAS-PEAL database

Table 1: The Average PSNR and SSIM for each method

| Quality measure | PSNR(dB) | SSIM |
|---|---|---|
| Chang et al. | 27.3842 | 0.8333 |
| Ma et al. | 28.2078 | 0.8910 |
| Jun et al. | 28.0306 | 0.8827 |
| Jiang et al. | 28.5735 | 0.9005 |
| Our method | **28.9978** | **0.9054** |

## 4. CONCLUSION

In this paper, we have proposed a novel patch representation method, called Locality-constrained Matrix Regression (LcMR), for face hallucination. It uses nuclear norm rather than $l_2$ norm to characterize the reconstruction error, thus preserving the essential structural information of the input image patch matrix. On the other hand, it imposes a locality constraint onto the combination coefficients to reach sparsity and locality simultaneously, aiming at obtaining the optimal representation of the input image patch. Experimental results have demonstrated the superiority of the proposed method over some state-of-the-art methods.

## ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China under Grant nos. 61502245, 61503195, 61402238 and 61571236, the NUPTSF under Grant nos. NY214204 and NY214165, the Natural Science Foundation of Jiangsu Province under Grant no. BK20150849, the China Postdoctoral Science Foundation under Grant no. 2015M571786, the Natural Science Fund for Colleges and Universities in Jiangsu Province under Grant nos. 15KJB520026.

## 6. REFERENCES


[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine,* vol. 20, no. 3, pp. 21-36, May, 2003.

[2] 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, Sep, 2002.

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

[4] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in Proceedings of IEEE Computer Society Conference onComputer Vision and Pattern Recognition 2004, pp. 1275-1282.

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

[6] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 31, no. 2, pp. 210-227, Feb, 2009.

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

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

[9] H. Li, L. Xu, and G. Liu, "Face hallucination via similarity constraints," *IEEE Signal Processing Letters,* vol. 20, no. 1, pp. 19-22, Jan, 2013.

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

[11] J. Shi, X. Liu, and C. Qi, "Global consistency, local sparsity and pixel correlation: A unified framework for face hallucination," *Pattern Recognition,* vol. 47, no. 11, pp. 3520-3534, Nov, 2014.

[12] K. Huang, R. Hu, Z. Han, T. Lu, J. Jiang, and F. Wang, "Face hallucination via K-selection mean constrained sparse representation," in Proceedings of International Conference on Pattern Recognition, 2012, pp. 882-885.

[13] X. Ma, H. Q. Luong, W. Philips, H. Song, and H. Cui, "Sparse representation and position prior based face hallucination upon classified over-complete dictionaries," *Signal processing,* vol. 92, no. 9, pp. 2066-2074, Sep, 2012.

[14] Y. Li, C. Cai, G. Qiu, and K.-M. Lam, "Face hallucination based on sparse local-pixel structure," *Pattern Recognition,* vol. 47, no. 3, pp. 1261-1270, Mar, 2014.

[15] Y. Liang, J.-H. Lai, P. C. Yuen, W. W. Zou, and Z. Cai, "Face hallucination with imprecise-alignment using iterative sparse representation," *Pattern Recognition,* vol. 47, no. 10, pp. 3327-3342, Oct, 2014.

[16] Z. Wang, R. Hu, Z. Shao, and Z. Hou, "Parameter estimation in sparse representation based face hallucination," *Digital Signal Processing,* vol. 31, pp. 28-34, Aug, 2014.

[17] Z. Wang, R. Hu, S. Wang, and J. Jiang, "Face Hallucination Via Weighted Adaptive Sparse Regularization," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 24, no. 5, pp. 802-813, May, 2014.

[18] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 3360-3367.

[19] J. Yang, J. Qian, L. Luo, F. Zhang, and Y. Gao, "Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes," *arXiv preprint arXiv:1405.1207*, 2014.

[20] F. Zhang, J. Yang, Y. Tai, and J. Tang, "Double Nuclear Norm-Based Matrix Decomposition for Occluded Image Recovery and Background Modeling," *IEEE Transactions on Image Processing,* vol. 24, no. 6, pp. 1956-1966, Feb, 2015.

[21] J. F. Cai, E. J. Candes, and Z. W. Shen, "A Singular Value Thresholding Algorithm for Matrix Completion," *Siam Journal on Optimization,* vol. 20, no. 4, pp. 1956-1982, 2010.

[22] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The CAS-PEAL large-scale Chinese face database and baseline evaluations," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* vol. 38, no. 1, pp. 149-161, Jan, 2008.