# SUPERVISED-LEARNING BASED FACE HALLUCINATION FOR ENHANCING FACE RECOGNITION

*Weng-Tai Su*[1], *Chih-Chung Hsu*[2], *Chia-Wen Lin*[1,2], *Weiyao Lin*[3]

[1] Dept. of Electrical Engineering,
National Tsing-Hua University, Taiwan
E-mail: cwlin@ee.nthu.edu.tw

[2] Institute of Communication Engineering,
National Tsing-Hua University, Taiwan
Department of Electronic Engineering
Shanghai Jiao Tong University

## ABSTRACT

This paper presents a two-step supervised face hallucination framework based on class-specific dictionary learning. Since the performance of learning-based face hallucination relies on its training set, an inappropriate training set (e.g., an input face image is very different from the training set) can reduce the visual quality of reconstructed high-resolution (HR) face significantly. To address this problem, we propose to utilize supervised learning to learn a set of class-specific dictionaries so that one of the learned dictionaries can well fit the global and local characteristics of an input low-resolution (LR) face image. Besides, the representative coefficients of the input LR face image may be unreliable due to insufficient information contained in the LR input image. To resolve this issue, we propose a maximum a posteriori estimator to infer the global HR face. Experimental results demonstrate that our method cannot only effectively enhance the visual quality of a reconstructed HR face, but also significantly improves the accuracy of face recognition compared to existing hallucination methods.

***Keywords*** *Face hallucination; super-resolution; supervised learning; face recognition; Bayesian estimation;*

## 1. INTRODUCTION

In recent years, face hallucination has become an attractive technique in enlarging face photos because it has many applications such as security in surveillance video, face recognition, facial expression estimation, face age estimation, and image/video editing which usually require face images with enough fine details. The problem of face hallucination is, however, different from that for general image super-resolution because face images have a unified structure which people are very familiar with. Even only few reconstruction errors occurring in a face image will cause visually annoying artifacts.

Recently, several face hallucination methods have been proposed, including subspace-based methods [1][2], neighbor embedding [3], sparse representation based method [4][6], and other basis decomposition based methods [7]–[13]. In order to make full use of the structural information of facial images, Parker *et al*. [8] adopted locality preserving projection (LPP) to preserve local structures of reconstructed face images by using neighboring structures in the pixel domain. Moreover, Li *et al*. [13] proposed local-pixel structure to global domain (LPS-GIS) by using sparse representations to learn the local-pixel structures for face image super-resolution. Besides, nonnegative matrix factorization (NMF), like PCA, has been widely used for basis decomposition for face recognition [9] and face hallucination [4]. Since NMF-based face recognition schemes show the superior performance than PCA-based schemes [9][11], Yang *et al*. [4] adopted NMF as basis decomposition functions to upscale the input LR face image, followed by a sparse representation-based super-resolution to refine the result. In [10][11], overcomplete NMF (ONMF) basis decomposition is further used to improve the visual quality of the reconstructed face image, especially in local facial parts. In [4], the reconstruction of facial parts is based on incomplete NMF, implying the representing power may be restricted. However, once the input LR face image is dissimilar to the training set, the dictionaries leaned by the above methods cannot effectively represent the input LR face, making the visual quality of the reconstructed HR face images unacceptable.

To recognize the identities of LR face images taken by surveillance cameras is also a challenging problem, as the resolutions of the face images are usually too low to provide sufficient features. Although existing face hallucination techniques can reconstruct the missing HR details to some extent, they cannot guarantee faithful recovery of the missing details, especially when the input LR face is significantly different from the faces in the training samples. To overcome this problem, by extending our previous work [6], we propose a two-step supervised learning-based face hallucination scheme as shown in Fig. 1. The main contribution of this method is two-fold: (i) We propose a supervised learning scheme to select appropriate sets of training samples to learn class-specific dictionaries to better fit various input LR faces. (ii) We propose a maximum-a-posteriori (MAP) framework to estimate representative coefficients to avoid introducing artifacts in the reconstructed HR faces. Thanks to the new dictionary learning scheme, our method can guarantee that the selected class-specific dictionaries can better represent input LR face images, making it not only achieve good face hallucination quality but also provide effective additional information for enhancing face recognition performance.

The rest of this paper is organized as follows. Sec. II presents the proposed supervised face hallucination scheme for global face reconstruction. In Sec. III, the refinement of local facial parts is briefly discussed. In Sec. IV, experimental results are demonstrated. Finally, Sec. V concludes this paper.

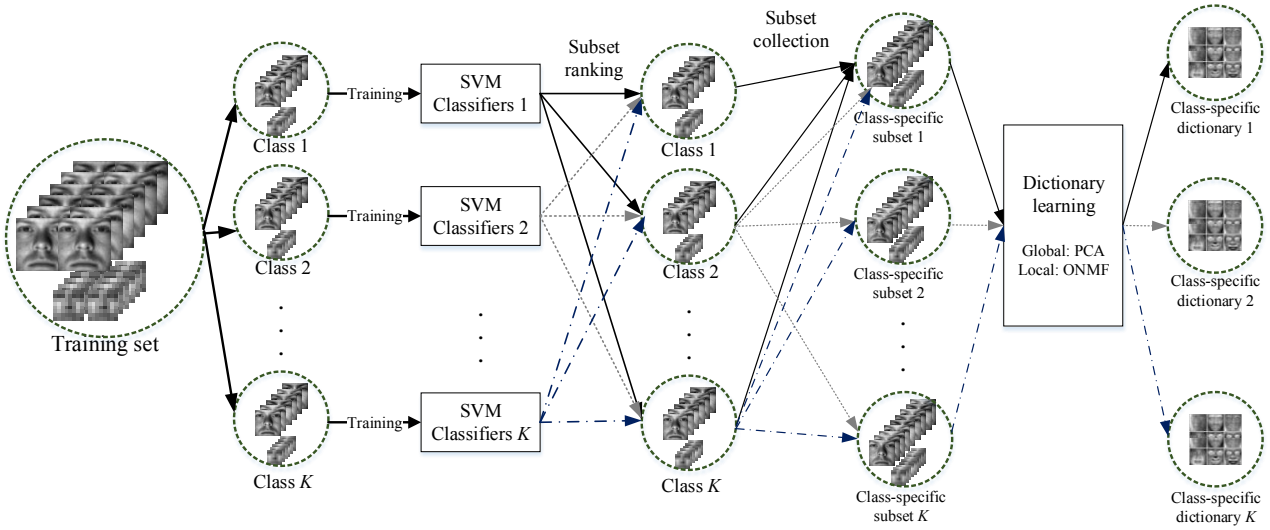


**Fig.** 1. Block diagram of the proposed class-specific dictionary learning.

## 2. CLASS-SPECIFIC LEARNING BASED FACE HALLUCINATION

### A. Preprocessing

Given a training face image set, we adopt the RASL face alignment scheme [5], which is based on sparse priors and low-rank decomposition, to align all face images in the training set. RASL is also used to align the input LR face. Since the local and global facial features have different characteristics, besides the global face, three local facial parts including eyes, nose, and mouth, are extracted from the training set to learn the corresponding dictionaries separately. In total, there are four training sets including global face, eyes, nose, and mouth images.

### B. Class-Specific Dictionary Learning

To learn dictionaries for effectively representing $K$ individual identities for both face hallucination and recognition, as shown in Fig. 1, our method uses $K$ SVM classifiers to divide the offline collected training set into $K$ subsets, and then learn $K$ class-specific dictionaries from the $K$ subsets. For each input LR face image, the SVM classifiers will be used to select the most representative dictionary for face hallucination. Suppose that we have a training set consisting of $N$ HR training face images and its LR counterpart pre-collected from $K$ people. Each HR-LR face image pair in the training set is concatenated as a vector as follows:

$$I_T = (\underbrace{i_1, i_2, \dots, i_L}_{\text{LR face}}, \underbrace{i_{L+1}, \dots, i_{H+L}}_{\text{HR face}}), \quad (1)$$

where $L$ and $H$ denote the pixel numbers of the LR and HR face images, respectively.

To simplify the notation, we denote by $I_{\text{LR}}$ and $I_{\text{HR}}$ the LR and HR parts of $I_T$. Then, the training set $\mathbf{I}_T = \{I_T^1, I_T^2, \dots, I_T^N\}$ is further partitioned into $K$ subsets according to the label information, and $K$ linear SVM classifiers are learned from the $K$ subsets, respectively. The ranking score $r_{n,c}$ of the $n$-th HR face image $I_{T_c}^n$ obtained by the $c$-th linear SVM classifier is computed by its class probability value [14]. To enrich the $c$-th subset $\mathbf{I}_c = \{I_{T_c}^1, I_{T_c}^2, \dots, I_{T_c}^{n_c}\}$, where $n_c$ is the number of training samples in $\mathbf{I}_c$, all face images in training set $\mathbf{I}_T$ are re-ranked by the scores obtained by the $c$-th classifier. To maximize the similarity value $R_c$ in $\mathbf{I}_c$ for class-specific dictionary learning, we find

$$R_c^* = \max_c \textstyle\sum_n r_{n,c}, \forall n, \quad (2)$$

To solve (2), the ranking scores of the HR training face images corresponding to class $c$ is calculated and then sorted in the descend order as follows:

$$\mathbf{I}_c' = \{I_{T_c}^1, I_{T_c}^2, \dots, I_{T_c}^{m_c}\},$$
$$\text{s.t. } r_{1,c} \geq r_{2,c} \geq \cdots \geq r_{m_c,c}. \quad (3)$$

where $m_c$ is a predefined number and $m_c > n_c$.

In this way, we can maximize the similarity of training samples in $\mathbf{I}_c'$. Note that training subset $\mathbf{I}_c'$ is collected not only from $\mathbf{I}_c$ but also from the other subsets. Then, a basis decomposition method (e.g., PCA, NME, or ONMF) can be applied on the $K$ training subsets to obtain the class-specific dictionary set $\mathbf{P_T} = \{P^1, P^2, \dots, P^k\}$.

Since the global face and local facial parts of the input face image are separately hallucinated, at first a facial mask is used to extract the global face $I_{\text{LR}}^G$ and local facial parts $I_{\text{LR}}^L$ from the input LR face image. For $I_{\text{LR}}^G$, we apply the global SVM classifiers to determine the class of $I_{\text{LR}}^G$, and then select the corresponding class-specific dictionary $P$ from $\mathbf{P_T}$ accordingly. Similarly, the class-specific dictionary for $I_{\text{LR}}^L$ are selected using local SVM classifiers. For the sake of clarity in notation, we use $\mathbf{F_T}$ and $F$ to denote the set of class-specific dictionaries learned for local facial parts and the class-specific dictionary selected for hallucination, respectively.

### C. Global Face Hallucination

Once the class-specific dictionaries are selected by the SVM classifiers, we use exemplar-based face hallucination to reconstruct the HR global face. Since a global face image is relatively smooth, Principal Component Analysis (PCA) can usually do a good job in learning a representative global

dictionary. With the learned dictionary $P_{\mathrm{LR}}$, the global LR face can be represented by

$$I_{\mathrm{LR}}^{G} = P_{\mathrm{LR}} \cdot \boldsymbol{\alpha}_{\mathrm{LR}}^{G}, \quad (4)$$

where $P_{\mathrm{LR}}$ is the LR part of class-specific dictionary $P$, $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$ is the LR PCA coefficient vector with respect to $P_{\mathrm{LR}}$. The LR coefficients $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$ can be estimated via the following least-squares approximation:

$$\boldsymbol{\alpha}_{\mathrm{LR}}^{G} \cong (P_{LR}^{\mathrm{T}} \cdot P_{\mathrm{LR}})^{-1} \cdot P_{LR}^{\mathrm{T}} \cdot I_{\mathrm{LR}}^{G}. \quad (5)$$

The approximation of $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$ in (5), however, may be inaccurate due to insufficient information contained in $I_{\mathrm{LR}}^{G}$. To address the problem, we propose a MAP estimator to more accurately estimate $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$. In general, we can model the relationship between HR and LR face images using a downsampling matrix $D$ as

$$I_{\mathrm{LR}}^{G} = DI_{\mathrm{HR}}^{G} = DP_{\mathrm{HR}}\boldsymbol{\alpha}_{\mathrm{HR}}^{G}. \quad (6)$$

Combining (4) and (6), we have

$$P_{\mathrm{LR}}\boldsymbol{\alpha}_{\mathrm{LR}}^{G} = DP_{\mathrm{HR}}\boldsymbol{\alpha}_{\mathrm{HR}}^{G}. \quad (7)$$

To estimate the relationship between coefficient $\boldsymbol{\alpha}_{\mathrm{HR}}^{G}$ and $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$, we can use the following least-squares approximation:

$$\boldsymbol{\alpha}_{\mathrm{LR}}^{G} \cong (P_{\mathrm{LR}}^{\mathrm{T}}P_{\mathrm{LR}})^{-1}P_{\mathrm{LR}}^{\mathrm{T}}DP_{\mathrm{HR}}\boldsymbol{\alpha}_{\mathrm{HR}}^{G} = H\boldsymbol{\alpha}_{\mathrm{HR}}^{G}, \quad (8)$$

where $H = (P_{\mathrm{LR}}^{\mathrm{T}}P_{\mathrm{LR}})^{-1}P_{\mathrm{LR}}^{\mathrm{T}}DP_{\mathrm{HR}}$.

To estimate $\boldsymbol{\alpha}_{\mathrm{HR}}^{G}$ from $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$, we first collect the PCA coefficient vectors of the $N$ HR and LR training pairs: $A_{\mathrm{HR}} = \{\boldsymbol{\alpha}_{\mathrm{HR}_1}^{G}, \boldsymbol{\alpha}_{\mathrm{HR}_2}^{G}, \dots, \boldsymbol{\alpha}_{\mathrm{HR}_N}^{G}\}$ and $A_{\mathrm{HR}} = \{\boldsymbol{\alpha}_{\mathrm{LR}_1}^{G}, \boldsymbol{\alpha}_{\mathrm{LR}_2}^{G}, \dots, \boldsymbol{\alpha}_{\mathrm{LR}_N}^{G}\}$, respectively. By estimating a posteriori model $p(\boldsymbol{\alpha}_{\mathrm{HR}}|\boldsymbol{\alpha}_{\mathrm{LR}})$ from $A_{\mathrm{HR}}$ and $A_{\mathrm{HR}}$, we can then estimate $\boldsymbol{\alpha}_{\mathrm{HR}}^{G}$ from $\boldsymbol{\alpha}_{\mathrm{LR}}^{G}$. Based on Bayesian's rule, the posteriori probability can be factorized into a likelihood tem and a prior term: $p(\boldsymbol{\alpha}_{\mathrm{HR}}|\boldsymbol{\alpha}_{\mathrm{LR}}) \propto p(\boldsymbol{\alpha}_{\mathrm{LR}}|\boldsymbol{\alpha}_{\mathrm{HR}})p(\boldsymbol{\alpha}_{\mathrm{HR}})$. As a result, to maximize the posteriori probability is equivalent to maximize the product of its likelihood and prior terms. Thus, the MAP estimate of $\boldsymbol{\alpha}_{\mathrm{HR}}$ is

$$\boldsymbol{\alpha}_{\mathrm{HR}}^{*} = \arg\max_{\boldsymbol{\alpha}_{\mathrm{HR}}} p(\boldsymbol{\alpha}_{\mathrm{LR}}|\boldsymbol{\alpha}_{\mathrm{HR}})p(\boldsymbol{\alpha}_{\mathrm{HR}}). \quad (9)$$

Because solving (9) directly may be inaccurate and unstable as the training data set is still of high dimension, we propose to utilize orthogonal locality preserving projection (OLPP) [18] to project the coefficient vector to a compact feature domain. As a result, the compact representation of the coefficient vector becomes

$$\mathbf{y}_{\mathrm{HR}} = B_{\mathrm{OLPP}}^{\mathrm{T}}\boldsymbol{\alpha}_{\mathrm{HR}}^{G}, \quad (10)$$

where $B_{\mathrm{OLPP}}$ is the OLPP projection matrix, which is orthogonal (i.e. $B_{\mathrm{OLPP}}^{\mathrm{T}} = B_{\mathrm{OLPP}}^{-1}$). Then, the MAP problem can be rewritten as:

$$\mathbf{y}_{\mathrm{HR}}^{*} = \arg\max_{\mathbf{y}_{\mathrm{HR}}} p(\boldsymbol{\alpha}_{\mathrm{LR}}|\mathbf{y}_{\mathrm{HR}})p(\mathbf{y}_{\mathrm{HR}}). \quad (11)$$

where $\mathbf{y}_{\mathrm{HR}}^{*}$ denotes the compact representation of the coefficient vector.

Assuming the likelihood and prior terms are Gaussian distributed, the probability distribution of the prior term is modeled as

$$p(\mathbf{y}_{\mathrm{HR}}) = \frac{1}{Z_p}\exp\{-\mathbf{y}_{\mathrm{HR}}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{y}_{\mathrm{HR}}\}, \quad (12)$$

where $Z_p$ is a normalization constant and $\boldsymbol{\Sigma}$ is a covariance matrix. Likely, the likelihood term can be modeled as

$$p(\boldsymbol{\alpha}_{\mathrm{LR}}^{G}|\mathbf{y}_{\mathrm{HR}}) = \frac{1}{Z_l}\exp\{-\frac{\left\|\mathbf{H}B_{\mathrm{OLPP}}\mathbf{y}_{\mathrm{HR}} - \boldsymbol{\alpha}_{\mathrm{LR}}^{G}\right\|}{\lambda}\} \quad (13)$$

Taking the log operation on (13) and plugging (12) and (13) into (11), we have

$$\mathbf{y}_{\mathrm{HR}}^{*} = \arg\min_{\mathbf{y}_{\mathrm{HR}}}(\lambda\ \mathbf{y}_{\mathrm{HR}}^{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}_{\mathrm{HR}} + \left\|HB_{\mathrm{OLPP}}\mathbf{y}_{\mathrm{HR}} - \boldsymbol{\alpha}_{\mathrm{LR}}^{G}\right\|) \quad (14)$$

By taking partial derivative on the right of (14) with respect to $\mathbf{y}_{\mathrm{HR}}$ and setting it to zero to solve the minimization problem, we have

$$\mathbf{y}_{\mathrm{HR}}^{*} = (B_{\mathrm{OLPP}}^{\mathrm{T}}H^{T}HB_{\mathrm{OLPP}} + \lambda\,\boldsymbol{\Sigma}^{-1})^{-1}B_{\mathrm{OLPP}}^{\mathrm{T}}H\,\boldsymbol{\alpha}_{\mathrm{LR}}^{G}. \quad (15)$$

Now, we can obtain the refined coefficient $\boldsymbol{\alpha}_{LR}^{G*}$ using

$$\boldsymbol{\alpha}_{\mathrm{HR}}^{G*} = B_{\mathrm{OLPP}}\,\mathbf{y}_{\mathrm{HR}}^{*}, \quad (16)$$

Finally, the HR global face $I_{\mathrm{HR}}^{G}$ can be obtained as follows:

$$\hat{I}_{\mathrm{HR}}^{G} = P_{\mathrm{HR}}\,\boldsymbol{\alpha}_{\mathrm{HR}}^{G*}. \quad (17)$$

In this manner, the global face can be reconstructed without annoying artifacts based on two facts: 1) the most representative class-specific dictionary for the input LR face is selected by the proposed supervised dictionary learning and selection, and 2) given a class-specific dictionary, the proposed MSP estimator can more accurately estimate the HR coefficient vector based on some useful priors about $p(\boldsymbol{\alpha}_{\mathrm{LR}}|\boldsymbol{\alpha}_{\mathrm{HR}})$ and $p(\boldsymbol{\alpha}_{\mathrm{HR}})$.

## 3. RECONSTRUCTION OF LOCAL FACIAL PARTS

In local facial part hallucination, we further partition a local face image into three facial parts: eyes, nose, and mouth. For each facial part, SVM classifiers are used to select class-specific dictionary $F$ from the dictionary set for local facial parts. However, since the local facial parts contain more complex details, PCA may not have enough representing power for those facial parts. Instead, as reported in [6][11], overcomplete dictionary has more representing power for local facial parts. Therefore, in this work, we choose ONMF [10][11] as our basis decomposition method for analyzing and synthesizing local facial parts. Then, the local facial parts can be respectively reconstructed in a similar manner described in Sec. 2.C. Consequently, the final reconstructed HR face image $\hat{I}_{\mathrm{HR}}^{*}$ can be obtained by

$$\hat{I}_{\mathrm{HR}}^{*} = \hat{I}_{\mathrm{HR}}^{L} + \hat{I}_{\mathrm{HR}}^{G}. \quad (7)$$

where $\hat{I}_{\mathrm{HR}}^{L}$ is the reconstructed local facial part image obtained by combining the three local-facial-part images.

In post-processing, the discontinuity in boundary between local facial parts and smoothing region is addressed using alpha-matting [17], and then a smoothing filter with Gaussian kernel is further used to mitigate the possible minor artifacts in boundary.

## 4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed method, we compare our method with five recent face hallucination methods for performance evaluation, including eigentransformation-based (EF) method [1], the PCA-based method [2], the Two-Step method [7], relationship learning method (RLSR) [12], and sparse local-pixel structure (LPS-GIS) [13]. Since there is still no widely-accepted objective quality metric for face hallucination now, besides subjective evaluation, we further utilize the face recognition rate with the reconstructed HR face image as a quality metric to evaluate whether the reconstructed HR details are useful in face recognition. Our experiments are performed on a public face dataset Yale B [15], which contains 16128 frontal view images from 38 people under 9 poses and 64 illumination conditions. Since the pose variation issue is beyond the scope of this paper, we select 910 face images with the frontal view for the 38 identities, which are aligned by the RASL method [5]. The size


**Fig. 2.** Subjective performance comparison: (a) 8x8 input LR images and the reconstructed $32 \times 32$ HR face images using (b) EF [1], (c) the PCA-Based method [2], (d) the Two-Step method [7], (e) RLSR [12], (f) LPS-GIS [13], (g) the proposed method, and (h) ground-truth HR images.

of the input LR face image is $8 \times 8$ and that of the hallucinated HR face image is $32 \times 32$. We randomly select 607 images for training and the remaining 303 images for testing. Parameter $m_c$ in the proposed method is set to be 20. In the test dataset, for each test identity only a few training samples are used for the same identity (say, $n_c = 15$ on average).

### A. Subjectively Visual Quality Evaluation

Fig. 2 shows hallucination results obtained by the proposed method and the five compared methods. Fig. 2 shows that the visual quality of the reconstructed face images obtained by the our method [Fig. 2(g)] is better than those in Fig. 2(c)–(e). Besides, the proposed method achieves higher similarity between the reconstructed HR faces and the ground-truths compared with the others. The performance improvement lies in the fact that the class-specific dictionaries learned in our method embed the label information so that usually an effective dictionary suitable for representing the input face image can be selected, implying that most useful information for the input LR face can be extracted for face hallucination and recognition. Even when the input LR face does not belong to any person whose faces are used in the training set, the common intrinsic features shared in the dictionary of a specific class are still useful in enhancing the visual quality of a hallucinated HR face.

### B. Objective Quality Evaluation by Recognition Rate

To evaluate the objective face recognition performances of the HR faces reconstructed by the proposed method and the compared methods, three state-of-the-art face recognition engines SRC [16], MFL [19], and CR [20] are adopted. In this experiment, 38 HR face images with 38 subjects (1 for each subject) are randomly selected from 303 ground-truths of the test images as the training set. The rest of the reconstructed HR face images are used as the test images. For fair comparison, we resize the size of ground-truth face images to $32 \times 32$, the same as that of the reconstructed HR faces. As shown in Table I, the average recognition rates with our method are significantly higher than that with the others and are very close to the rates obtained using the ground-truth HR face, meaning that the HR faces reconstructed by our method offer useful additional HR details for identifying a person, which is very important for surveillance applications. In contrast, the recognition rate with Liu's method [7] is rather low due to the dissimilarity and artifacts in the reconstructed HR faces.

**Table 1.** Comparison of face recognition rates using the HR faces reconstructed by the proposed method and the others using three SRC-based face recognition engines: SRC [16], MFL [19], CR [20]

| Method\ Recognition rate | SRC | MFL | CR |
|---|---|---|---|
| EF [1] | 60% | 61% | 62% |
| PCA-based [2] | 50% | 49% | 59% |
| Two-Step [7] | 3% | 4% | 5% |
| RLSR [12] | 62% | 63% | 66% |
| LPS-GIS [13] | 63% | 66% | 69% |
| Proposed | **73%** | **74%** | **78%** |
| Ground-truth HR | 77% | 78% | 80% |

## 5. CONCLUSION

In this paper, we proposed a supervised learning-based face hallucination scheme which learns class-specific dictionaries based on label information to well fit the characteristics of input LR faces. We have also proposed a two-step scheme to utilize PCA and ONMF to learn the most representative dictionaries for global face and local facial parts. Our experimental results demonstrate that the proposed method outperforms the state-of-the-art face hallucination methods in terms of the visual quality of reconstructed HR face image and face recognition performance

## REFERENCES


[1] X. Wang, and X. Tang, "Hallucinating face by eigentransformation," *IEEE Trans. Syst. Man, Cybernet.*, vol. 35, no. 3, pp. 425-434, 2005.

[2] J. S. Park, and S. W. Lee, "An example-based face hallucination method for single-frame, low-resolution facial images," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1806-16, Oct, 2008.

[3] H. H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, pp. 275-282, 2004.

[4] J. C. Yang, S. W. Ma, and T. Huang, "Face hallucination via sparse coding," in *Proc. IEEE Int. Conf. Image Process.*, pp. 1264-1267, 2008.

[5] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Yi Ma, "Robust batch alignment of images by sparse and low-rank decomposition" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2233-2246, Nov., 2012.

[6] C.-C. Hsu, C.-W. Lin, C.-T. Hsu, and H. Y. Mark Liao, "Face hallucination using Bayesian global estimation and local basis selection," in *Proc. IEEE Workshop Multimedia Signal Process.*, Saint-Malo, France, 2010.

[7] C. Liu, H. Y. Shum, and W. T. Freeman, "Face hallucination: theory and practice," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 115–134, 2007.

[8] S. W. Park and M. Savvides, "Breaking the limitation of manifold analysis for super-resolution of facial images," in *Proc. of IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 1, pp. I-573–I-576, 2007.

[9] T. P. Zhang, B. Fang, Y. Y. Tang, and G. H. He, "Topology preserving non-negative matrix factorization for face recognition," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 574–584, 2008.

[10] A. Hyvärinen, J. Hurri, and P. O. Hoyer, Ch 13: Overcomplete and nonnegative models: Springer, 2009.

[11] J. Eggert, and E. Korner, "Sparse coding and NMF," in *Proc. of IEEE Int. Joint Conf. Neural Net.*, vol. 4, pp. 2529-2533, 2004.

[12] W. W. Zou, and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Trans. Image Process.*, pp.1,6, 27-29, Sept. 2012.

[13] Y. Li, C. Q. Cai, G. Quiu, and K. M. Lam, "Face hallucination based on sparse local-pixel structure," *Pattern Recognit.*, vol. 47, pp. 1261-1270, 2014.

[14] C.C. Chang and C.J. Lin, LIBSVM: A Library for Support Vector Machines, 2001 [online] Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm

[15] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.23, no.6, pp.643-660, June 2001.

[16] M. Yang, L. Zhang, J. Yang, D. Zhang, "Robust sparse coding for face recognition," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, pp. 625-632, 2010.

[17] E. S. L. Gastal and M. M. Oliveira, "Shared sampling for real-time alpha matting," in *Proc. Eurographics*, 2010.

[18] D. Cai, X. He, J. Han, and H.-J. Zhang, "Orthogonal Laplacianfaces for face recognition," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3608–3614, Nov. 2006.

[19] M. Yang, L. Zhang, J. Yang and D. Zhang, "Metaface learning for sparse representation based face recognition," in *Proc. IEEE Int. Conf. Image Process.*, pp. 1601–1604, 2010.

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