

# Face Hallucination Using Bayesian Global Estimation and Local Basis Selection

Chih-Chung Hsu<sup>1</sup>, Chia-Wen Lin<sup>2</sup>, Chiou-Ting Hsu<sup>3</sup>, Hong-Yuan Mark Liao<sup>4</sup>, and Jen-Yu Yu<sup>5</sup>

<sup>1,2</sup>Department of Electrical Engineering, National Tsing-Hua University

<sup>3</sup>Department of Computer Science, National Tsing-Hua University  
101, Section 2, Kuang-Fu Road, Hsinchu 30013, Taiwan

{d9661805@oz, cwlin@ee, cthsu@cs}.nthu.edu.tw

<sup>4</sup>Institute of Information Science, Academia Sinica  
128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan

liao@iis.sinica.edu.tw

<sup>5</sup>Information & Comm. Research Lab., Industrial Technology Research Institute  
195, Sec. 4, Chung Hsing Rd., Chutung, Hsinchu 310, Taiwan  
KavinYu@itri.org.tw

**Abstract**—This paper proposes a two-step prototype-face-based scheme of hallucinating the high-resolution detail of a low-resolution input face image. The proposed scheme is mainly composed of two steps: the global estimation step and the local facial-parts refinement step. In the global estimation step, the initial high-resolution face image is hallucinated via a linear combination of the global prototype faces with a coefficient vector. Instead of estimating coefficient vector in the high-dimensional raw image domain, we propose a maximum a posteriori (MAP) estimator to estimate the optimum set of coefficients in the low-dimensional coefficient domain. In the local refinement step, the facial parts (i.e., eyes, nose and mouth) are further refined using a basis selection method based on overcomplete nonnegative matrix factorization (ONMF). Experimental results demonstrate that the proposed method can achieve significant subjective and objective improvement over state-of-the-art face hallucination methods, especially when an input face does not belong to a person in the training data set.

## I. INTRODUCTION

In recent years, face hallucination [1]-[3] has become an attractive technique in super-resolving 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. simple interpolation schemes are difficult to achieve. Example-based super-resolution (SR) schemes [1][5] have proven to be able to reconstruct significantly finer details of a low-resolution (LR) image compared to interpolation-based schemes, provided that a comprehensive set of training images are used to collect prior knowledge of the structures and patterns of images using machine learning techniques.

The problem of super-resolution for face images is, however, different from that for generic images because face images have a unified structure which people are very familiar with. Even only few reconstruction errors occurring in a face image can cause visually annoying artifacts. For example, geometry distortion in the mouth and eyes on a reconstructed face image may only reduce the image's objective quality slightly, whereas the subjective quality of the reconstructed face can be degraded significantly. Therefore, both the global face shape and textures and local geometric structures (e.g., mouth, nose, and eyes) need to be treated carefully in face hallucination [2].

Example-based face hallucination schemes can be classified into two major kinds of approaches: prototype-face-based approaches [3][6] and model-based approaches [2][8][10]. The prototype-faces-based methods proposed in [3] decomposes an input LR face into a set of prototype faces (e.g., eigenfaces) as prior models using principle components analysis (PCA). After reconstructing the HR face, the method in [3] uses a recursive error back-propagation method to further refine the reconstructed high-resolution (HR) face image. The method proposed in [6] uses a similar approach but applies a different basis decomposition scheme.

The model-based approaches generally utilize probabilistic models to find the relationship between HR images and their corresponding LR versions. The model-based method proposed in [10] uses a two-step approach that captures the global and local geometrical features of a face using a parametric Gaussian model and a non-parametric local model based on Markov random field (MRF), respectively. The method then uses the maximum a posterior (MAP) approach to approximate the output HR face image under different situations. Similarly, in the method proposed in [2], the global and the local models are partitioned into two parts: soft and hard constraints, which can address the problems of hallucination under different situations (e.g., the number of training samples is greater/less than the dimension of the training samples). Recently, the manifold learning techniques have been used in face hallucination methods and more details of the hallucinated face image can be preserved since manifold learning can preserve the neighboring structures in the pixel domain. For example, the locality preserving projection (LPP) is used in [8] to hallucinate a HR face image, where the MAP approach is applied to estimate the optimal HR face image in the LPP domain.

Besides, PCA, nonnegative matrix factorization (NMF) has also been used for basis decomposition for face recognition [9] and face hallucination [4]. The NMF method, like PCA, represents a facial image as a linear combination of basis images, whereas the NMF method can better learn localized parts-based representation for face images. For example, NMF can yield a decomposition of human faces into parts reminiscent of features such as lips, eyes, and nose. Compared to NMF, the reconstruction results of PCA are not that intuitive and hard to interpret as PCA allows subtractive combinations of the basis images [9]. Moreover, the NMF-based face recognition schemes have been shown to achieve better performance than PCA-based schemes do [9], [14]. However, the NMF basis decomposition function used in [4] is an incomplete basis since the

number of the bases is restricted by its definition. Usually, an overcomplete bases set provides better performance in term of the quality of the reconstructed images [13], especially in local facial features. For example, an overcomplete NMF (ONMF) method was proposed in [14] which can represent the local facial features well.

The rest of this paper is organized as follows. In Section 2, we give an overview of the proposed face hallucination scheme. Section 3 presents the proposed Bayesian global face estimation scheme based on prototype face decomposition and ONMF. In Section 4, we present the proposal overcomplete basis selection scheme which can further refine the visual quality of facial parts. Experimental results of the proposed scheme are demonstrated in Section 5. Finally, Section 6 concludes this paper.

## II. OVERVIEW OF THE PROPOSED SCHEME

Estimating the corresponding HR details from a LR face is in nature an ill-posed inverse problem. Therefore, priors such as spatio-temporal consistency, sparsity of signal representation, and structures of faces, are used to relax the ill-posedness of the HR face reconstruction. Based on the priors, MAP estimators have been proposed to estimate the HR face image from a learned low-dimensional subspace with [2][10][12]. However, it is difficult to estimate a HR face image accurately due to the complexity of face images and the limited information carried in the LR input face, even though we can use prior models to relax the ill-posedness. Therefore, the hallucinated HR face images usually present annoying artifacts caused by the inaccurate estimation.

Recently, it was shown in [4][5] that learning the statistical structures of faces (e.g., the prototype faces) as priors is an effective way to address the ill-posedness problem. The method proposed in [4] performs an iterative back projection scheme to refine the estimated HR details, whereas the method proposed in [5] uses sparsity priors to obtain better estimation. Our method is built on top of the prototype-faces-based framework proposed in [3][4]. This method exploits a whole LR face image, rather than dividing the face into small patches like in [5], to reconstruct the HR face. Let  $\mathbf{I}_H$  denote the HR image to be estimated, represented as

$$\mathbf{I}_H = (i_1, i_2, \dots, i_L, i_{L+1}, i_{L+2}, \dots, i_H) \quad (1)$$

where  $i_j$  represents the value of the  $j$ -th pixel,  $L$  and  $H$  denote the pixel numbers of LR and HR images, respectively. The reconstructed face image can be obtained by

$$\mathbf{I} \approx \mathbf{P} \cdot \boldsymbol{\alpha} = \mathbf{R} \quad (2)$$

where  $\mathbf{P}$  denotes the prototype faces,  $\boldsymbol{\alpha}$  stands for the reconstruction coefficients, and  $\mathbf{R}$  denotes the reconstructed face.

The coefficients of the input LR face image corresponding to a set of LR prototype faces can be obtained by using the following least-squares projection:

$$\boldsymbol{\alpha} = ((\mathbf{P}_L)^T \cdot \mathbf{P}_L)^{-1} \cdot (\mathbf{P}_L)^T \cdot \mathbf{I}_L, \quad (3)$$

where  $\mathbf{P}_L$  denotes the LR prototype faces and the  $\mathbf{I}_L$  denotes the input LR face image.

After computing the coefficients, the HR face image can be hallucinated via a linear combination of HR prototype faces weighted by the coefficients as illustrated in Fig. 1. The coefficients obtained from decomposing the LR face image, however, are typically not very accurate for synthesizing the HR counterpart. Therefore, in the reconstruction phase, an iterative method based on back projection was proposed in [3] to refine the estimated the coefficients corresponding to the HR prototype faces for face hallucination.

As shown in Fig. 2, the proposed face hallucination method consists of two phases: the training phase and the hallucination phase. Before face hallucination, pre-processing including face alignment using active appearance model (AAM) [7], scaling, and

normalization is performed on the LR image to ensure that the prior models can be reconstructed well when the inputs are obtained under different situations (e.g., different poses, environments and ages). In the training phase of our method (see Fig. 2(a)), each LR face is divided into a global face and a local-parts-only face containing main facial parts (i.e., eyes, nose and mouth) extracted using masks. Subsequently, two different prototype faces decomposition scheme, PCA and ONMF, are applied on the two faces, respectively.

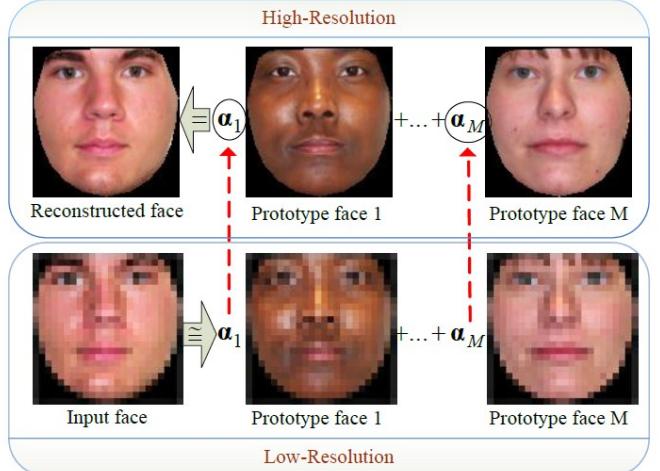


Fig. 1 Illustration of prototype-faces-based face hallucination.

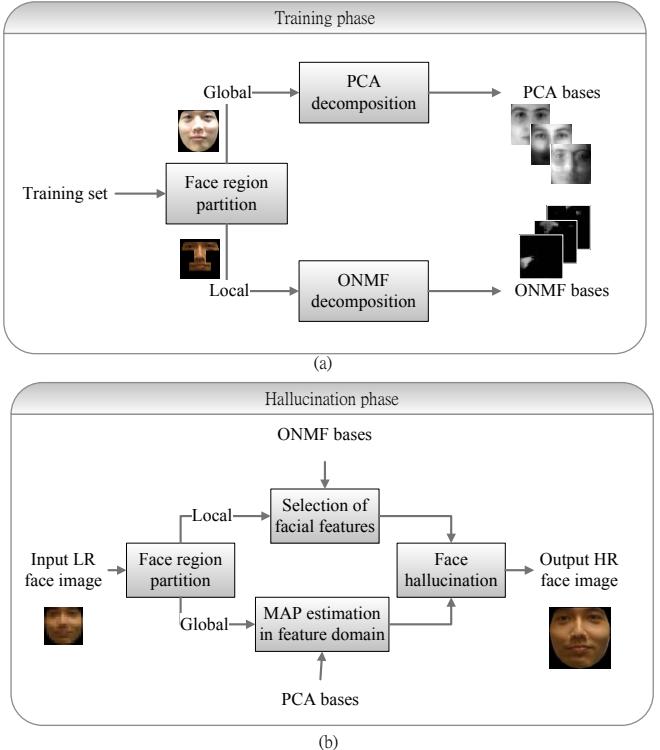


Fig. 2 The proposed two-step face hallucination scheme consists of two phases: (a) the training phase, and (b) the hallucination phase.

The proposed face hallucination method consists of two major steps: the MAP-based global face estimation and the local features refinement. As illustrated in Fig. 2(b), the input LR face is again divided into a whole face and a facial-parts-only portion. The whole face is used to estimate the global face while the local-parts-only portion is used to guide the selection of local-parts-based prototype

faces (i.e., ONMF) so as to reconstruct fine facial parts. In the global face estimation step, the input face is initially decomposed using the LR PCA prototype faces obtained in the training phase using least-squares projection. The LR PCA coefficient vector is then projected to a low-dimensional Orthogonal Locality Preserving Projection (OLPP) feature domain. The HR PCA coefficient vector is estimated from the low-dimensional OLPP feature vector using a MAP estimator. As a result, the global face is obtained using the estimated HR PCA coefficient vector.

In the local feature refinement step, a set of ONMF bases is used to represent the individual facial parts. The facial parts of the input LR face including eyes, nose, and mouth, are extracted using a facial mask. The entropy values of the individual local parts are calculated and are then used to guide the selection of ONMF bases that best represent the individual facial parts. The coefficients corresponding to the selected bases are computed accordingly.

After the two steps, the local facial parts are hallucinated using the selected set of ONMF bases and the corresponding coefficients. The remaining facial region is subsequently hallucinated using the result of the global face estimation step. Finally, the boundaries of the two hallucinated results are blended.

### III. BAYESIAN GLOBAL FACE ESTIMATION

Fig. 3 shows the flowchart of the proposed global face estimation step. At the training stage, the prototype faces are trained as prior models by applying PCA on the training set. The initial coefficients for HR face image reconstruction can be calculated from LR PCA prototype images by using least-squares approximation. The initial coefficient vector is then projected to a low-dimensional feature subspace using OLPP. As a result, the final HR PCA coefficients for hallucination are then estimated using the MAP estimator. Finally, the HR global face image is hallucinated using the estimated coefficients. Different from the methods in [2][8][10], the proposed MAP estimator is applied in the OLPP coefficient domain rather than in the pixel domain.

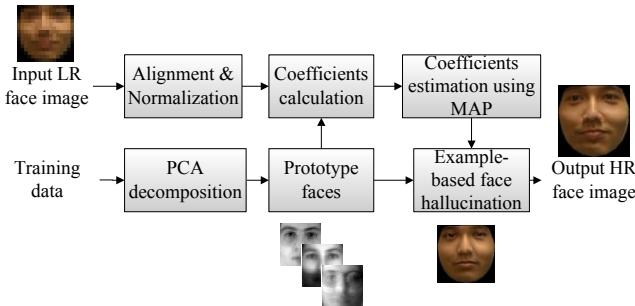


Fig. 3 Flowchart of global face estimation.

After the initial coefficient vector  $\mathbf{a}_L^g$  for reconstructing the global face is calculated by (3) with PCA prototype faces, the coefficients can be further refined using a MAP estimator. Recall the hallucination formulation in [2], the HR face image can be downsampled using a downscaling matrix by

$$\mathbf{I}_L^g = \mathbf{A} \mathbf{I}_H^g, \quad (4)$$

where  $\mathbf{A}$  denotes the transformation matrix for downscaling an image from HR to LR. As a result, the relationship between the coefficients  $\mathbf{a}_L^g$  and  $\mathbf{a}_H^g$  can be formulated as

$$\mathbf{P}_L \mathbf{a}_L^g = \mathbf{A} \mathbf{P}_H \mathbf{a}_H^g, \quad (5)$$

where  $\mathbf{P}_L$  and  $\mathbf{P}_H$  denote the matrices of LR and HR prototype faces, respectively.

The least-squares estimate of  $\mathbf{a}_L^g$  becomes

$$\mathbf{a}_L^g \cong (\mathbf{P}_L^T \mathbf{P}_L)^{-1} \mathbf{P}_L^T \mathbf{A} \mathbf{P}_H \mathbf{a}_H^g = \mathbf{A}_{LS} \mathbf{a}_H^g, \quad (6)$$

where  $\mathbf{A}_{LS} = (\mathbf{P}_L^T \mathbf{P}_L)^{-1} \mathbf{P}_L^T \mathbf{A} \mathbf{P}_H$ .

As explained above, in general, dimensionality reduction algorithms are used to project the face images to a low-dimensional subspace (e.g., a feature space). The purpose of dimensionality reduction is to extract the intrinsic features from the original high dimensional data. As a result, we can easily address optimization problems in the feature domain. In this work, we use OLPP for reducing the dimensionality of PCA coefficient vectors. The OLPP seeks the embedding which can preserve the manifold structure by measuring the “neighboring distance information” using the  $k$ -NN algorithm.

Given a set of PCA coefficient vectors of  $N$  HR images in the training set  $\mathcal{A}_H = \{\mathbf{a}_{H,1}^g, \mathbf{a}_{H,2}^g, \dots, \mathbf{a}_{H,N}^g\}$  and its LR counterpart  $\mathcal{A}_L = \{\mathbf{a}_{L,1}^g, \mathbf{a}_{L,2}^g, \dots, \mathbf{a}_{L,N}^g\}$  where  $N$  is the number of training face images. As a result, the estimated HR coefficient vector  $\mathbf{y}_H$  is expressed by

$$\mathbf{y}_H = \mathbf{B}_{OLPP}^T \mathbf{A}_H^g, \quad (7)$$

where  $\mathbf{B}_{OLPP}$  is an orthogonal transform, that is,  $\mathbf{B}_{OLPP}^T \mathbf{B}_{OLPP} = \mathbf{B}_{OLPP}^{-1}$ .

The a posteriori model  $p(\mathbf{a}_L | \mathbf{a}_H)$  can be factorized into a likelihood term and a prior term by applying Bayes' rule as follows.

$$p(\mathbf{a}_H | \mathbf{a}_L) \propto p(\mathbf{a}_L | \mathbf{a}_H) p(\mathbf{a}_H) \quad (8)$$

The ill-posed problem of maximizing the a posteriori probability  $p(\mathbf{a}_L | \mathbf{a}_H) p(\mathbf{a}_H)$  in (8) can be approximated by maximizing  $p(\mathbf{a}_L | \mathbf{y}_H) p(\mathbf{y}_H)$ . Therefore, we can reformulate the MAP problem as

$$\mathbf{y}_H^* = \arg \max_{\mathbf{y}_H} p(\mathbf{a}_L | \mathbf{y}_H) p(\mathbf{y}_H) \quad (9)$$

where  $\mathbf{y}_H^*$  denotes the optimal projected OLPP coefficients.

Assume the prior term can be simply modeled as a Gaussian distributed function [10] as

$$p(\mathbf{y}_H) = \frac{1}{Z} \exp \left\{ -\mathbf{y}_H^T \Sigma^{-1} \mathbf{y}_H \right\} \quad (10)$$

where  $\Sigma$  denotes a diagonal matrix and  $Z$  is a normalization constant. Similarly, the likelihood term can also be formulated as

$$p(\mathbf{a}_L^g | \mathbf{y}_H) = \frac{1}{Z} \exp \left\{ -\frac{\|\mathbf{A}_L \mathbf{B}_{OLPP} \mathbf{y}_H - \mathbf{a}_L^g\|^2}{\lambda} \right\} \quad (11)$$

The optimization problem is to minimize the following cost function:

$$\mathbf{y}_H^* = \arg \min_{\mathbf{y}_H} \left( \lambda \mathbf{y}_H^T \Sigma^{-1} \mathbf{y}_H + \|\mathbf{A}_{LS} \mathbf{B}_{OLPP} \mathbf{y}_H - \mathbf{a}_L^g\|^2 \right) \quad (12)$$

To obtain the optimal solution for (12), we take the partial derivative of (12) with respect to  $\mathbf{y}_H^*$ , and set the partial derivative to zero. This leads to

$$\mathbf{y}_H^* = \left( \mathbf{B}_{OLPP}^T \mathbf{A}_{LS}^T \mathbf{A}_{LS} \mathbf{B}_{OLPP} + \lambda \Sigma^{-1} \right)^{-1} \mathbf{B}_{OLPP}^T \mathbf{A}_{LS}^T \mathbf{a}_L^g \quad (13)$$

The optimal coefficients  $\mathbf{a}_H^{g*}$  can then be calculated by

$$\mathbf{a}_H^{g*} = \mathbf{B} \hat{\mathbf{y}}_H^* \quad (14)$$

Finally, the global HR face image  $\mathbf{I}_{..}$  can be reconstructed via a linear combination of the optimal coefficients  $\mathbf{a}_H^{g*}$  and the HR PCA prototype faces  $\mathbf{P}_H$ .

#### IV. LOCAL REFINEMENT USING LOCAL-PARTS-GUIDED BASIS SELECTION

When an input LR face is significantly different from those in the training database, the prototype-faces-based approach usually cannot do a good job because it is difficult to reconstruct the HR face well using the prototype faces that lack of the local features of the input face. This problem can be resolved using an overcomplete basis set which can well capture the local structures of facial parts (e.g., eyes, nose, and mouth) since the individual facial parts of an input face usually look similar to the combination of the parts from different faces, even if the face is largely different from the training faces. Since ONMF [14] is a parts-based basis decomposition scheme, it can better represent individual facial parts compared to global decomposition schemes such as PCA. An overcomplete basis set denotes the number of the bases is greater than the dimensionality of the bases. In general, the overcomplete bases contain more information from the training set [13]. The ONMF algorithm is formulated as the following optimization problem:

$$\begin{aligned} \{\mathbf{W}^*, \mathbf{H}^*\} = \arg \min_{\mathbf{W}, \mathbf{H}} & \|\mathbf{D} - \mathbf{WH}\|_2^2 + \lambda_s \cdot S(\mathbf{H}) \\ \text{s.t. } & \mathbf{W} \geq 0, \mathbf{H} \geq 0, \end{aligned} \quad (15)$$

where  $\mathbf{D} \in \mathbb{R}$  denotes data matrix,  $\mathbf{W} \in \mathbb{R}$  and  $\mathbf{H} \in \mathbb{R}$  denote bases and coefficients matrix, respectively,  $S$  denotes a sparsity constraint for the coefficients, and  $\lambda_s$  is a weight for the sparsity constraint. The solution for (15) can be obtained through the following updating rules [14]:

$$\begin{aligned} r_{ij} = & \odot \frac{\mathbf{D}^T \mathbf{W}^T}{\|\mathbf{W}^T\|_2^2} \mathbf{H}_i \\ & \sum \odot \mathbf{W}_j, \quad (16) \\ & \mathbf{W}_j \end{aligned}$$

where  $\mathbf{R} = \mathbf{D} - \mathbf{WH}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ ,  $1 \leq l \leq k$ . Note that the sparseness function can constrain the norm of  $\mathbf{H}$ .

The local refinement step aims to reconstruct a HR face image with fine and natural facial parts using the ONMF prototype faces. To this end, the LR facial-parts-only image can be formulated as

$$\mathbf{I}_L^l \equiv \mathbf{W}_L^l \mathbf{a}_L^l = \mathbf{R}_L^l, \quad (17)$$

where  $\mathbf{I}_L^l$  denotes the masked facial-parts-only image and  $\mathbf{W}_L^l \in \mathbb{R}$  denotes the local basis matrix of ONMF as given in (16).

The corresponding coefficients  $\mathbf{a}_L^l$  for the facial-parts-only image can then be obtained by

$$\mathbf{a}_L^l = \left( (\mathbf{W}_L^l)^T \cdot \mathbf{W}_L^l \right)^{-1} \cdot \mathbf{W}_L^l \cdot \mathbf{I}_L^l. \quad (18)$$

The HR facial-parts-only image  $\hat{\mathbf{I}}_L^l$  is then obtained by

$$\hat{\mathbf{I}}_L^l = \mathbf{H} \mathbf{a}_L^l. \quad (19)$$

The proposed basis selection algorithm is shown in Fig. 4. First, a facial mask is used to extract the facial parts, including eyes, nose, and mouth, of an input LR face. The entropy values of individual facial parts and computed and sorted. Second, some bases not contain all of the facial features. Therefore, the unnecessary bases should be removed for each facial feature. If one of the facial features in the prototype faces are not presented, then their entropy are corresponding low. Therefore, the entropy-based bases selection method is proposed.

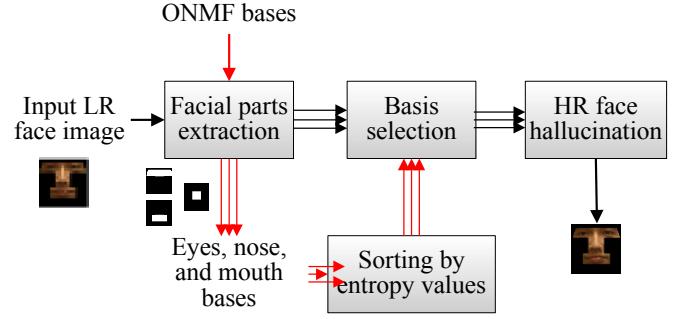


Fig. 4 Flowchart of the proposed local-parts-guided basis selection.

Let  $\mathbf{I}_{L,m}^l$  denote the  $m$ -th facial feature (e.g.,  $\mathbf{I}_L^1, \mathbf{I}_L^2$ , and  $\mathbf{I}_L^3$  are eyes, nose and mouth, respectively.) and  $\mathbf{W}^{l_m}$  the corresponding basis matrix, respectively. The entropy value of the ONMF can be calculated for each facial part as

$$H(\mathbf{W}_k^{l_m}) = - \sum_{i \in \text{Mask}_m} p(\mathbf{W}_k^{l_m}(i)) \log p(\mathbf{W}_k^{l_m}(i)) \quad (20)$$

where  $\mathbf{W}_k^{l_m}$  represents the  $k$ -th ONMF prototype face for the  $m$ -th facial part,  $i$  stands for the pixel index and  $\text{Mask}_m$  indicates the region of the  $m$ -th facial part.

Subsequently, the bases are sorted by their entropy values. Therefore, the prototype faces with the  $K$  largest entropy values corresponding to the  $m$ -th facial part are obtained as follows.

$$\begin{aligned} \overline{\mathbf{W}}^{l_m} &= \{\mathbf{W}_0^{l_m}, \mathbf{W}_1^{l_m}, \dots, \mathbf{W}_K^{l_m}\} \\ \text{s.t. } & H(\mathbf{W}_1^{l_m}) \geq H(\mathbf{W}_2^{l_m}) \geq \dots \geq H(\mathbf{W}_K^{l_m}) \end{aligned} \quad (21)$$

Therefore, the  $m$ -th HR facial part image  $\hat{\mathbf{I}}_L^l$  can be reconstructed by using the corresponding basis matrix  $\overline{\mathbf{W}}^{l_m}$  without unnecessary information. The final HR face image can be represented as

$$\hat{\mathbf{I}}_L^l = \hat{\mathbf{I}}_L^1 + \hat{\mathbf{I}}_L^2 + \dots + \hat{\mathbf{I}}_L^K. \quad (22)$$

Note, to perverse the sharpness edges and smoothness global faces, we only take non-facial-parts region of the global face (defined by the mask as mentioned previously) as  $\hat{\mathbf{I}}_L^l$ . Then, the boundaries between the global face image and the facial-parts-only image are smoothed using a Gaussian smoothing filter.

#### V. EXPERIMENTAL RESULTS

Our training set contains 482 images with the resolution of 64x64 which are selected from the PAL (Productive Aging Lab) face database. These face images were aligned using AAM which is a semi-automatic process before being used for training with PCA and ONMF. We first downscaled the images to the resolution of 16x16 and then hallucinated them to the original resolution. The numbers of ONMF and PCA prototype faces are 800 and 100, respectively.

Fig. 5 compares the subjective visual qualities of the reconstructed face images obtained from three schemes: the proposed global face compensation scheme, the PCA prototype-faces-based scheme proposed in [3] with and without iterative back projection, and the NMF prototype-faces-based scheme with sparsity prior proposed in [4]. The mean-squared error (MSE) values between the reconstructed HR faces and the ground-truth are also indicated. As shown in Fig. 5(c) and 5(d), although the HR faces hallucinated using [3]'s method look smooth and clear, some facial parts (e.g., the eyes, nose, and mouth) actually are not similar to the ground-truth. The NMF-

prototype-faces-based scheme [4] results in sharper facial details due to the sparsity prior used, while also generating unnatural and annoying artifacts as can be observed from Fig. 5(e). Besides, the hallucinated facial parts are also not similar to the ground-truth. The proposed scheme achieves significant subjective and objective improvement over the other methods in terms of the similarity with the ground-truth and the reconstruction error (MSE).

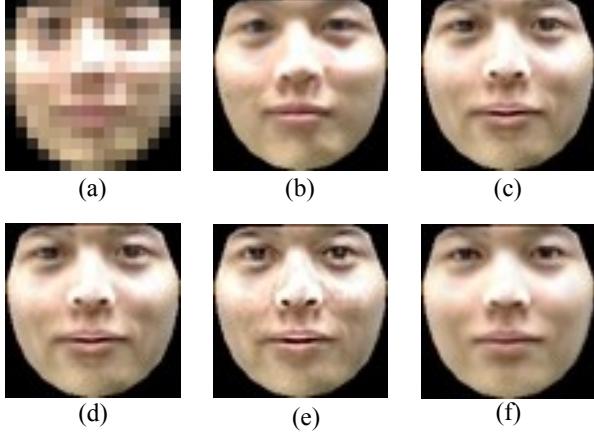


Fig. 5 Subjective quality comparison: (a) Input LR face image, (b) HR ground-truth, and the faces reconstructed using (c) [3]'s method without iterative back projection (MSE: 184.7), (d), [3]'s method with iterative back projection (MSE: 172.1) (e) [4]'s method (MSE: 242.3), (f) the proposed global estimation method (MSE: 161.3).

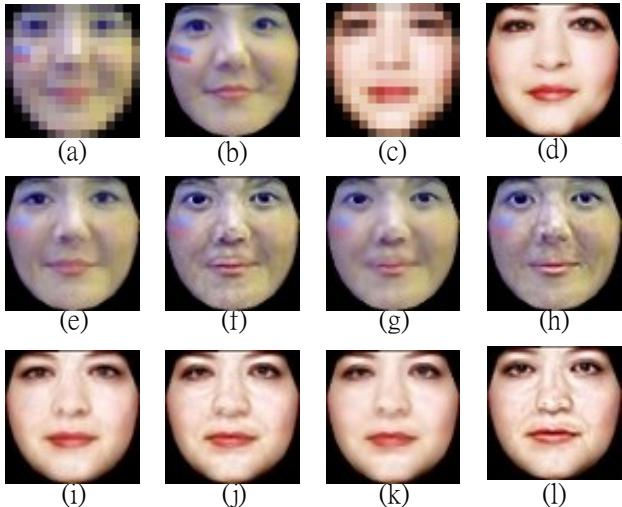


Fig. 6 Subjective and objective quality comparison: (a), (c) the two input LR face images, (b), (d) their corresponding ground truths, and the hallucinated HR face images using (e), (i) the proposed method (MSE: 99.1/115.5), (f), (j) [3]'s method (MSE: 222.0/164.6), (g), (k) [2]'s method (MSE: 143.3/150.1), and (h), (l) [4]'s method (MSE: 212.0/232.5).

Fig. 6 compares the subjective qualities for two test face images hallucinated with the proposed two-step method, and the two other methods proposed in [3] and [8], respectively. We can observe that the mouth in Fig. 6(e), compared with the mouths in Fig. 6(f)~(h), is much more similar to the ground-truth in Fig. 6(b). Besides, the eyebrows, eyes, nose in the HR face hallucinated with the proposed method look more similar to the ground-truth than the others. Similarly, the hallucinated HR face in Fig. 6(e) shows better subjective visual quality than those in Fig. 6(j)~(l). In addition to the

improved subjective quality, the proposed method also achieves better objective qualities for the two test images in terms of reconstruction error. The proposed method results in MSE values of 99.1 and 115.5 for the two test images whereas the others result in 222.0 and 164.6, 143.3 and 150.1, and 212.0 and 232.5, respectively. Evidently, the proposed method achieves significantly better subjective quality of the reconstructed face images.

## VI. CONCLUSION

In this paper, we proposed a prototype-faces-based face hallucination method. The proposed method consists of MAP-based estimation for global face hallucination and entropy-based basis selection for local facial part refinement. In the training phase, PCA and ONMF are used to decompose the prototype faces for the global face and local (facial-parts-only) face, respectively. In the hallucination phase, the coefficients for global HR face image are estimated using a MAP estimator in the OLPP feature domain. Furthermore, the entropy-based bases selection is applied to choose the representative bases for individual facial parts to reconstruct the HR facial parts. Finally, the output HR face image can be obtained by fusing the global face and the local facial parts together. Experimental results demonstrate that the proposed method achieve significant subjective and objective improvement over other existing schemes.

## REFERENCES

- [1] B. Baker and T. Kanade, "Limits on superresolution and how to break them," *IEEE Trans. Pattern Anal. Match. Intell.*, vol. 24, no. 9, pp. 1167–1183, Sept. 2002.
- [2] 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, Oct. 2007.
- [3] 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. 1806–1816, Oct. 2008.
- [4] J.C. Yang, H. Tang, Y. Ma, and T. Huang, "Face hallucination via sparse coding," in *Proc. IEEE Int. Conf. Image Process.*, pp.1264-1267, Oct. 2008
- [5] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics & Applications*, vol. 22, no. 2, pp. 56–65, Mar. 2002.
- [6] W. Zhang and W. K. Cham, "Learning-based face hallucination in DCT domain," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, June 2008.
- [7] T. F. Cootes and C. J. Taylor, *Statistical Models of Appearance for Computer Vision*, Technical Report, University Manchester, U.K., Mar. 2004.
- [8] S. W. Park and M. Savvides, "Breaking the limitation of manifold analysis for super-resolution of facial images," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Process.*, vol.1, pp. I-573–I-576, Apr. 2007.
- [9] T. P. Zhang, B. Fang, Y. Y. Tang, G. H. He, and J. Wen, "Topology preserving non-negative matrix factorization for face recognition," *IEEE Trans. Image Process.*, vol.17, no.4, pp. 574–584, Apr. 2008.
- [10] C. Liu, H.Y. Shum, and C.S. Zhang, "A two-step approach to hallucinating faces: global parametric model and local nonparametric model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol.1, pp. 192–198, 2001.
- [11] L. K. Saul, K. Q. Weinberger, F. Sha, J. Ham, and D. D. Lee, *Spectral methods for dimensionality reduction*, in *Semisupervised learning*, O. Chapelle, B. Schölkopf, and A. Zien, Eds. MIT Press: Cambridge, MA, 2006.
- [12] Y. Zhuang, J. Zhang, and F. Wu, "Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation," *Pattern Recognit.*, vol. 40, issue 11, November 2007.
- [13] A. Hyvärinen, J. Hurri, and P.O. Hoyer, "Ch 13: Overcomplete and non-negative models," Springer, 2009.
- [14] J. Eggert and E. Korner, "Sparse coding and NMF," in *Proc. IEEE Int. Joint Conf. Neural Networks*, vol. 4, pp. 2529–2533, July 2004.