Position-Patch Based Face Hallucination Using Convex Optimization

Cheolkon Jung, *Member, IEEE*, Licheng Jiao, *Senior Member, IEEE*, Bing Liu, and Maoguo Gong, *Member, IEEE*

***Abstract*—We provide a position-patch based face hallucination method using convex optimization. Recently, a novel position-patch based face hallucination method has been proposed to save computational time and achieve high-quality hallucinated results. This method has employed least square estimation to obtain the optimal weights for face hallucination. However, the least square estimation approach can provide biased solutions when the number of the training position-patches is much larger than the dimension of the patch. To overcome this problem, this letter proposes a new position-patch based face hallucination method which is based on convex optimization. Experimental results demonstrate that our method is very effective in producing high-quality hallucinated face images.**

***Index Terms*—Convex optimization, face hallucination, least square estimation, position-patch, sparse representation.**

I. Motivation

FACE hallucination which is also referred as face super-resolution (SR) is a technology to estimate a high-resolution (HR) image from low-resolution (LR) image sequences or a single LR one [1]–[7]. A number of related face hallucination methods have been proposed in recent years. Among them, learning based methods have received much attention because they can achieve high magnification factor and produce good super-resolved results compared with other methods. Recently, a novel face hallucination method based on position-patch has been proposed [8], [9]. The position-patch based method hallucinated the HR image patch using the same position image patches of each training images. Accordingly, it was able to save computational time and produce high-quality super-resolved results compared to manifold learning based methods. In this method, the optimal weights of the training image position-patches were computed by a constrained least square estimation method. Then, the hallucinated patches were reconstructed based on the optimal weights. Our method has been inspired by the position-patch based face hallucination method [8]. However, in our method, we make use of constrained convex optimization based on sparse representation instead of least square estimation to obtain the optimal weights for face hallucination. That is to say, instead of the L_2-based optimization in [8], we use L_1-based optimization over image patches. L_1-norm is more suitable for this problem because each patch can be approximated with a smaller subset of patches than L_2-norm. On the contrary, L_2-norm provides nonzero weights for all patches. Thus, the proposed method can produce more stable face hallucination results when the number of the training position-patches is much larger than the dimension of the patch [10].

The rest of the paper is organized as follows. Section II briefly reviews the position-patch based face hallucination method. Section III describes the proposed face hallucination method based on convex optimization in detail. Experimental results and analysis are provided in Section IV, and we conclude this paper in Section V.

Manuscript received February 11, 2011; revised March 29, 2011; accepted April 01, 2011. Date of publication April 07, 2011; date of current version April 25, 2011. This work was supported by the National Natural Science Foundation of China (Grants 61050110144, 60803097, 60972148, 60971128, 60970066, 61072106, 61003198, 61001206, and 61077009), the National Research Foundation for the Doctoral Program of Higher Education of China (Grants 200807010003 and 20100203120005), the National Science and Technology Ministry of China (Grants 9140A07011810DZ0107 and 9140A07021010DZ0131), the Key Project of Ministry of Education of China (Grant 108115), and the Fundamental Research Funds for the Central Universities (Grants JY10000902001, K50510020001, and JY10000902045). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lisimachos Paul Kondi.

The authors are with the Key Lab of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, China (e-mail: zhengzk@xidian.edu.cn; jlc1023@163.com; 372715246@qq.com; gong@ieee.org).

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

Digital Object Identifier 10.1109/LSP.2011.2140370

Algorithm 1 Position-patch based face hallucination

Step 1: Denote the input LR image, LR training image and HR training image in overlapping patches as $\{X_L^P(i,j)_{p=1}^N\}$, $\{Y_L^{mP}(i,j)\}_{p=1}^N$, and $\{Y_H^{mP}\}_{p=1}^N$, respectively, for $m = 1,2,\ldots M$.
Step 2: For each patch $X_L^P(i,j)$:
(a) Compute the reconstruction weights $w(i,j)$
(b) Synthesize the HR patch $X_H^P(i,j)$
Step 3: Concatenate and integrate the hallucinated HR patches to form a facial image, which is the target HR facial image $\{X_H^P(i,j)\}_{p=1}^N$.

II. Position-Patch Based Face Hallucination

The position-patch based face hallucination method is briefly summarized in Algorithm 1, where the notations are explained in Table I [8]. Each patch $X^P(i,j)$ in the face image $\{X^P(i,j)\}_{p=1}^N$ can be represented by

$$X^P(i,j) = \sum_{m=1}^{M} w_m(i,j) \cdot Y^{mP}(i,j) + e \tag{1}$$

where e is the reconstruction error. Each patch matrix of $X^P(i,j)$ and $Y^P(i,j)$ is converted into a column vector as in [8]. Accordingly, we can determine a weight $w_m(i,j)$ by the minimization of e as follows:

TABLE I
MEANING OF NOTATIONS

Notation	Meaning
X_L	LR testing image
Y_L	LR training image
Y_H	HR training image
$\{X_L^P(i,j)\}_{p=1}^N$	patch of X_L
$\{Y_L^{mP}(i,j)\}_{p=1}^N$	patch of the m^{th} Y_L
$\{Y_H^{mP}(i,j)\}_{p=1}^N$	patch of the m^{th} Y_H
M	the number of training images
N	the number of patches
$w(i,j)$	reconstruction weight at position (i,j)

$$w(i,j)=\arg\min_{w_m(i,j)}\left\|X^P(i,j)-\sum_{m=1}^{M} w_m(i,j)\cdot Y^{mP}(i,j)\right\| \tag{2}$$

where $w(i,j)$ is a M-dimensional weight vector of each reconstruction weight $w_m(i,j)$ for $m=1,2,\ldots,M$. To solve (2), a constrained least square problem has been employed as follows [8]:

$$w(i,j)=\frac{\mathtt{Z}^{-1}C}{C^T\mathtt{Z}^{-1}C} \tag{3}$$

where C is a column vector of ones which has M elements, and $\mathtt{Z}$ is local covariance matrix obtained by $\mathtt{Z}=\mathtt{S}^T\mathtt{S}$. Here, the matrix $\mathtt{S}$ is defined by $\mathtt{S}=X^P(i,j)\cdot C^T-\mathtt{Y}$, where $\mathtt{Y}$ is a matrix with its columns being training patches $Y^{mP}(i,j)$. A more efficient way to obtain $w(i,j)$ of (3) is to normalize the weights obtained by solving the linear equation $\mathtt{Z}\cdot w(i,j)=C$. Consequently, (2) can be represented in the following matrix form:

$$\mathtt{w}=\arg\min_{\mathtt{w}}\|\mathtt{X}-\mathtt{Y}\cdot\mathtt{w}\|. \tag{4}$$

Equations (2) and (4) are least square problems. Given $\mathtt{Y}\in\mathrm{R}^{d\times M}$, it is stable to use least square estimation to solve (2) and (4) when the matrix $\mathtt{Y}$ satisfy $d>M$. The dimension of the patch should be larger than the number of the training position-patches. However, the patch with 3×3 size (i.e., $d=9$) has been selected even if the number of training images is more than 100 (i.e., $M>100$) [8]. That is, M is much larger than d, and thus (2) and (4) are actually under-determined equations. Accordingly, the solutions of (2) and (4) are not unique. In this case, the least square estimation can produce biased solutions [10].

III. FACE HALLUCINATION BASED ON CONVEX OPTIMIZATION

Fortunately, the compressed sensing theory provides some possible solutions of this problem. Equations (2) and (4) can be converted as follows:

$$\min_{\mathtt{w}}\|\mathtt{w}\|_1 \quad \texttt{subject to} \ \|\mathtt{X}-\mathtt{Y}\cdot\mathtt{w}\|_2^2\le\varepsilon. \tag{5}$$

Fig. 1. Face hallucination results. (a) Input 25×25 LR faces. (b) Bicubic interpolated images. (c) Freeman *et al.* [6] (d) Chang *et al.* [7] (e) Ma *et al.* [8] (f) Our method. (g) Original 100×100 HR faces.

Equation (5) is a convex constrained optimization problem and we have employed it for solving the NP-hard problem. By compressed sensing theory, we can regard the training position-patch set as an overcomplete dictionary whose base elements are the position-patches. In addition, the testing LR patches can be represented using the overcomplete dictionary. This representation is naturally sparse if the size of the training position-patches is reasonably large [10]. Accordingly, we can get a more stable reconstruction weight $w(i,j)$ for face hallucination. The hallucinated HR patch $X_H^P(i,j)$ is obtained by

$$X_H^P(i,j)=\sum_{m=1}^{M} Y_H^{mP}(i,j)\cdot w_m(i,j). \tag{6}$$

Consequently, the target HR image $\mathtt{X}_H$ is reconstructed by combining the hallucinated HR patches. The proposed face hallucination method is described in Algorithm 2.

IV. EXPERIMENTAL RESULTS

To verify the superiority of our method, experiments were performed on the CMU PIE database[11]. The database contains 41 368 images obtained from 68 subjects. We took the frontal face images with 21 different illumination conditions. Thus, the total number of images were 1428 in our experiments. Among them, 630 images of 30 subjects were used in the training stage, and the rest were used in the synthesis stage. We compared our method with conventional methods based on the same training set. They are bicubic interpolation, Freeman *et al.* [6], Chang *et al.* [7], and Ma *et al.* [8]'s methods. In [7], the HR patch size of Y_H^m was 12×12 while corresponding LR patch size of Y_L^m was

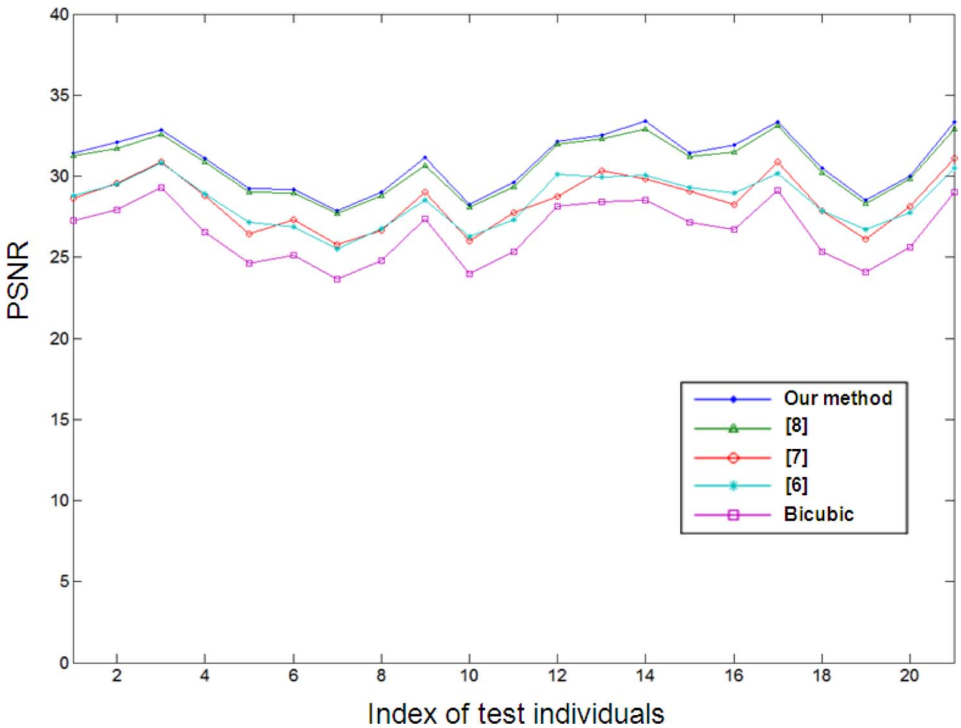

Fig. 2. PSNR values of face hallucination results in the 40th subject. The magnification factor is 4 and the unit of the PSNR is dB.

Algorithm 2 Position-patch based face hallucination using convex optimization

Input: Training set $\{Y_L^m\}_{m=1}^M$ and $\{Y_H^m\}_{m=1}^M$, LR image $\mathbf{X}_L$.
Output: Estimated HR image $\mathbf{X}_H$.

Training Stage:
Obtain the position-patches $\{Y_L^{mP}(i,j)\}_{p=1}^N$ and $\{Y_H^{mP}(i,j)\}_{p=1}^N$ $(m = 1,2,..,M)$ from the training set.
Synthesis Stage:
Step 1: For each LR-patch $X_L^P(i,j)$ of $\mathbf{X}_L$:
a) Compute the reconstruction weight $w(i,j)$ by solving the following convex optimization problem:

$$\min \|\mathbf{w}\|_1 \quad \texttt{subject to} \ \|\mathbf{X}_L^P - \mathbf{Y}_L^P \cdot \mathbf{w}\|_2^2 \le \varepsilon$$

where $\mathbf{Y}_L^P$ is a column matrix of the training patches $Y_L^{mP}(i,j)$ $(m = 1,2,...,M)$.
b) Construct the HR-patch representation by

$$X_H^P(i,j) = \sum_{m=1}^M Y_H^{mP}(i,j) \cdot w_m(i,j)$$

Step 2: Reconstruct the target HR image $\mathbf{X}_H$ by concatenating and integrating the hallucinated HR-patches.

3×3. In addition, the number of the neighbor patches for reconstruction was 5. The size of the image patches in our method and [8] was 3×3. The size of LR images for training and synthesis was 25×25, and they were generated by the following equation.

$$\mathtt{I}_L = \mathtt{D} \cdot \mathtt{H} \cdot \mathtt{I}_H \tag{7}$$

where $\mathtt{I}_L$ and $\mathtt{I}_H$ are LR and HR images, respectively; and $\mathtt{H}$ and $\mathtt{D}$ are a smoothing filter and down-sampling operator, respectively. Interpolation factor was 4, and thus the size of hallucinated results was 100×100. The LR patch overlapped with its adjacent patch by 3 pixels (i.e., one column) and its corresponding HR patch overlapped by 48 pixels (i.e., four columns). To solve the convex optimization problem of (5), we used a primal-dual algorithm for linear programming based on [12]. In the equation, the error tolerance ε was set to 1.0.

TABLE II
AVERAGE PSNR AND SSIM VALUES OF DIFFERENT METHODS (THE MAGNIFICATION FACTOR IS 4 AND THE UNIT OF THE PSNR IS IN dB)

Method	Bicubic	Freeman et al. [6]	Chang et al.[7]	Ma et al.[8]	Our method
PSNR	24.5388	26.0954	26.3785	28.1613	**28.2437**
SSIM	0.7278	0.7544	0.7444	0.8146	**0.8178**

Some representative hallucinated results are shown in Fig. 1. The hallucinated results of [6] and [7] are somewhat blurred and with some artifacts. However, our method and [8] produce more natural looking facial images than them. Further examination of the results reveals that our method is more effective in preserving the edge and image details in the nose and mouth areas than [8] (see the 6th-8th rows of the figure). Moreover, Fig. 2 shows the PSNR values of face hallucination results in the 40th subject. It can be observed that our method outperforms the other methods in terms of the PSNR. In addition, average PSNR and SSIM values of the face hallucination results are provided in Table II. The SSIM is a complementary measure of the PSNR, which gives an indication of image quality based on known characteristics of the human visual system [13]. As shown in the table, our method achieves the best hallucination performances in terms of the PSNR and SSIM. Here, the bold numbers represent the best PSNR and SSIM values.

V. CONCLUSION

In this letter, we have proposed a new position-patch based method for face hallucination using convex optimization. The proposed method is based on the compressed sensing theory and computes the optimal weights for face hallucination by solving the convex optimization problem. Experimental results show that our method consistently achieves good performance on face hallucination tasks, outperforming the latest state-of-the art methods in terms of PSNR and SSIM. We believe that our method can be effectively employed for enhancing image quality in various image processing applications.

REFERENCES

[1] X. Ma, H. Huang, S. Wang, and C. Qi, "A simple approach to multiview face hallucination," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 579–582, 2010.

[2] B. Li, H. Chang, S. Shan, and X. Chen, "Aligning coupled maifolds for face hallucination," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 957–960, 2009.

[3] J. Yang, H. Tang, Y. Ma, and T. Huang, "Face hallucination via sparse coding," in *Proc. IEEE Conf. Image Processing*, 2008, pp. 1264–1267.

[4] 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, 2008.

[5] W. Zhang and W. K. Cham, "Learning-based face hallucination in DCT domain," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[6] W. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Applicat.*, vol. 22, no. 2, pp. 56–65, 2002.

[7] H. Chang, D. Y. Yeung, and Y. M. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2004, pp. 1275–1282.

[8] X. Ma, J. Zhang, and C. Qi, "Hallucinating face by position-patch," *Pattern Recognit.*, vol. 43, pp. 2224–2236, 2010.

[9] X. Ma, J. Zhang, and C. Qi, "Position-based face hallucination method," in *Proc. IEEE Conf. Multimedia and Expo*, 2009, pp. 290–293.

[10] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, 2009.

[11] T. Sim, S. Baker, and M. Bsat, "The CMU pose illumination and expression (PIE) database," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2002, pp. 46–51.

[12] E. Candes and J. Rombergt, ℓ^1-Magic: Recovery of Sparse Signals via Convex Programming 2005 [Online]. Available: http://www.acm.caltech.edu/l1magic/

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.