J Supercomput
DOI 10.1007/s11227-014-1257-z

Face hallucination and recognition in social network services

Feng Jiang · Seungmin Rho · Bo-Wei Chen · Xiaodan Du · Debin Zhao

Abstract Due to the rapid growth of social network services such as Facebook and Twitter, incorporation of face recognition in these large-scale web services is attracting much attention in both academia and industry. The major problem in such applications is that visual data provided are in general noisy, low resolution, prone to degradation due to lighting and other adverse effects. This paper proposes a novel face hallucination method with an aim of improving face recognition performance with the photography in social media data, which is modeled as a progressive process which explores the specific face characteristics and priors based on frequency bands analysis. In the first stage of our algorithm, initially estimated middle-resolution images are generated based on a patch-based learning method in discrete cosine transformation (DCT) domain as the first-scale restoration image. According to the relationship between the high-resolution face images and their lower resolution ones, the DC coefficients and AC coefficients are estimated separately. In the second stage, aiming at generating more refined high-resolution face images, a DCT up-sample algorithm emphasizing on low-frequency

F. Jiang (✉) · X. Du · D. Zhao
School of Computer Science, Harbin Institute of Technology, Harbin 150001, China
e-mail: fiang@hit.edu.cn

X. Du
e-mail: xddu@hit.edu.cn

D. Zhao
e-mail: dbzhao@hit.edu.cn

S. Rho
Department of Multimedia, Sungkyul University, Anyang, Korea
e-mail: smrho@sungkyul.ac.kr

B.-W. Chen
Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan
e-mail: dennisbwc@gmail.com

Published online: 27 July 2014

bands preserving is applied. Meanwhile, an interpolation-based method is presented in spatial domain to obtain high-frequency bands. Extensive experiments show that the proposed algorithm effectively helps improving the recognition performance of low-resolution faces.

Keywords Social media data · Face hallucination · Learning-based · Frequency bands

1 Introduction

In recent days, social network services (SNS) such as Twitter, Facebook and Blogger strongly affect our daily life, the industry and the academy. In order to understand and analyze this new life style, new research fields emerged, including computational social science [13], computational social network analysis [17] and computational journalism [7]. Among these social media data shared and analyzed among members of SNS, facial photographs occupy a significant share. Photography analysis containing faces is attracting attention from the academy and the industry. Face.com offers a platform for developers and publishers to automatically detect and recognize faces in photos using a robust, free, REST API. The rememberAR App for Android platform allows a user to look up the Facebook account from his Facebook friends with same event participants. Augmented ID provided a new interface combining face recognition, tracking, augmented reality and SNS. Moreover, Google has acquired NevenVision, Riya, and PittPatt and deployed face recognition into Picasa [25]. Recent advances in the computer vision and multimedia scientific communities have greatly improved the performance of face recognition in social networks media data. Several works on face recognition under social network service have been proposed recently [8,20]. In [1], the authors focused on the democratization of surveillance, faces as conduits between online and offline data, the emergence of personally predictable information, the rise of visual, facial searches and the future of privacy in the world of augmented reality. Face recognition under SNS is a very challenging and yet useful application. The major problem in such applications is that visual data provided are in general noisy, low resolution, prone to degradation due to lighting and other adverse effects. Accordingly, in order to enhance the face recognition performance, face hallucination technique is proposed, inspired with the reasonable assumption that the recognition performance can be improved if the image resolution is large enough. Face hallucination aims at inferring additional pixels to increase the resolution for very low-resolution face images. Aiming at improving face recognition performance with the photography in social media data, this paper proposes a novel face hallucination method and our contributions are summarized as follows:

- A novel learning-based method in discrete cosine transformation (DCT) domain is proposed to generate middle-resolution images (MRIs) as the first-scale preserving image. The redundancy of training set is reduced effectively and simplified clustering-based offline training scheme is adopted with obtained MRIs.
- The specific face characteristics and priors are explored differently based on frequency bands analysis. The evidence in the experiments shows that significant

performance can be achieved even with relative simple process and low computational costs.
– The proposed face hallucination method can significantly improve the recognition performance and the relationship between the face hallucination method and the recognition performance is further investigated.

The remainder of the paper is organized as follows. Related works are overviewed in Sect. 2. Section 3 elaborates the proposed learning-based face hallucination in DCT domain for low- and middle-frequency components restoration. The high-frequency components restoration with spatial domain interpolation is presented in Sect. 4. Extensive experimental results are reported in Sect. 5. Section 6 summarizes this paper.

2 Related work

Face images analyses such as detection, recognition, 3D modeling and tracking have been studied extensively in the fields of computer vision, pattern recognition and machine learning. Various commercial products and research methods have been proposed for the face recognition-related applications. Research on face recognition has been popular for over a decade. In general, face recognition system consists of three steps [30]: face detection, feature extraction and face recognition. Face detection techniques have been widely applied in our daily lives, such as the face detection function in digital cameras. In terms of feature extraction and face recognition, many approaches have been proposed such as principal component analysis (PCA) [21], linear discriminant analysis (LDA) [4], locality preserving projections (LPP) [10] and face recognition via sparse representation (SRC) [23]. Face recognition algorithms are also widely implemented nowadays, for example, the face recognition function in Facebook and iphoto. In many cases, the face images captured by live cameras are often of low resolutions due to the environment or equipment limitations. Thus, how to recover human faces automatically has become an important problem for the further works such as face analysis and recognition. Some current methods work quite well under certain circumstances. However, low-resolution face image may degrade seriously due to variations in real-world applications such as video surveillance. Face hallucination algorithms dealing with less-degradation input may not work well for these cases. In surveillance data, when faces are captured in a far distance, their images are generally regarded too small to be recognized. Compared to general image super-resolution [2,28], face hallucination always employs specific face characteristics and priors. Face hallucination can handle more challenging tasks than generic image super-resolution due to the regular face structure. In the last decade, various face hallucination methods have been proposed. Baker and Kanade [3] was the first to introduce the face hallucination theory. Based on the image super-resolution theory, they proposed a learning-based algorithm, which learns the prior on the spatial distribution of the face image gradient and yields high-resolution face images. Numerous face hallucination approaches have been proposed ever since [15,22,26,31]. Wang and Tang proposed an efficient hallucinating algorithm [16]. They treated the low-resolution and high-resolution faces as two groups, and tried to find the linear transformation relation between those two groups. They derived this through PCA, where both the low- and high-resolution

images are projected into their eigen subspaces, and the linear transformation could be interpreted in these two subspaces. Zhuang et al. [31] adopted locality preserve projection (LPP) and radial basis function (RBF) to produce the global faces and simplify the non-parametric model to generate the local features. Yang et al. [26] adopted the non-negative matrix factorization (NMF) algorithm to generate global faces and found the local residues through sparse representation method. Zhang and Cham [29] proposed an approach in frequency domain. They transformed faces through the DCT, and estimated the high-resolution DC components and AC components separately. Through the inverse DCT, hallucinated faces can be produced. Other super-resolution approaches for generic images also can be used in face hallucination. Chang et al. [5] proposed a super-resolution approach based on the locality linear embedding. In addition, Liu et al. [15] developed a two-step statistical modeling approach, which integrates a PCA-based global model with a local Markov network, corresponding to the common and specific face characteristics, respectively. Park and Lee [19] presented a face hallucination method to reconstruct a HR face image by combining an example-based reconstruction method with an extended morphable face model. Li and Lin [14] improved the face hallucination based on manifold alignment by projecting the point-pairs from the original coupled manifolds into the embeddings of the common manifold. Inspired by the position-patch based face hallucination method, Ma et al. [16] proposed a position-patch face hallucination, which uses position patches to estimate the optimal weights for reconstructing a HR face image. Based on the assumption that two similar face images should have similar local pixel structures, Hu et al. [11] proposed a new face hallucination framework which learns the local pixel structures for the target HR face and then uses the input LR face and the learned pixel structures as priors to estimate the target HR face. Yang et al. [27] proposed an image super-resolution method via sparse representation. In their method, a medium resolution face image is obtained first, and then a sparse prior is applied on low dimensional image patches for recovering local details. Zhang and Cham [29] propose a learning-based face hallucination method in the DCT domain, which can produce a high-resolution face image from a single low-resolution one.

3 Face hallucination based on frequency bands analysis

The frameworks of the proposed face hallucination method are shown as Fig. 1. Instead of working in one domain, we tend to explore the specific face characteristics and priors in both transform and spatial domains. In addition, rather than in the same scale, middle-resolution image as the first-scale restored image, which is generated by middle-resolution DC coefficient and a middle-resolution AC coefficient inference model. Moreover, in the second-scale image restoration, we adopt different preserving models based on diverse frequency bands analysis. Utilizing different scales provide better performance.

The proposed method is a two-step framework where the face hallucination is formulated as a learning-based problem and a super-resolution afterwards and we infer and obtain DC and AC coefficients with different strategies. In the first step, aiming at generating the first-scale middle-resolution face images, the LR input I_L is

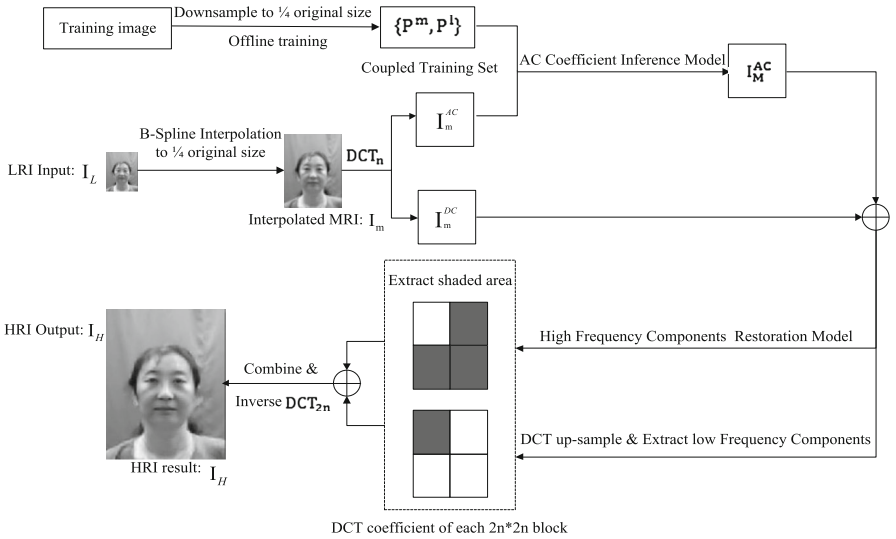

Fig. 1 Framework of proposed method

up-sampled to obtain the interpolated middle-resolution image (MRI): I_m with half size of HRI height and width, respectively, then we employ a $n \times n$ DCT transform DCT_n to get DCT coefficients: I_m^{DC} and I_m^{AC}. Given the coupled training set $\{P^m, P^l\}$ generated by an offline cluster-based training strategy, we can apply the proposed learning-based AC coefficient inference model to get more accurate AC coefficients I_M^{AC}. Meanwhile, I_m^{DC} is taken as the final MRI DC coefficient I_M^{DC} directly. In the second step, two different strategies are applied simultaneously according to different frequency bands. A DCT up-sample method is employed to get final low-frequency information. However, it still suffers from block artifacts. To tackle incompleteness of local features, an interpolation-based high-frequency components restoration model is put forward. After that, an inverse DCT_{2n} is applied to get final HRI output: I_H.

3.1 Learning-based face hallucination in DCT domain

A LRI or MRI y can be formulated by Eq. 1, where x is the original image, D and H denote down-sampling and blurring processing, respectively, and gn is Gaussian noise.

$$y = DHx + \mathrm{gn}. \tag{1}$$

P^m is a set of MRIs which are of 1/4 original HRI size. Once a LRI is given, we first interpolate it into 1/4 original size as the P^l in coupled training set $\{P^m, P^l\}$. The $n \times n$ DCT_n transform is applied on both P^m and P^l afterwards. DC coefficient stands for the average pixel intensity of the target image and can be estimated quite accurate by classic interpolation-based methods such as bicubic and B-spline [29], while AC

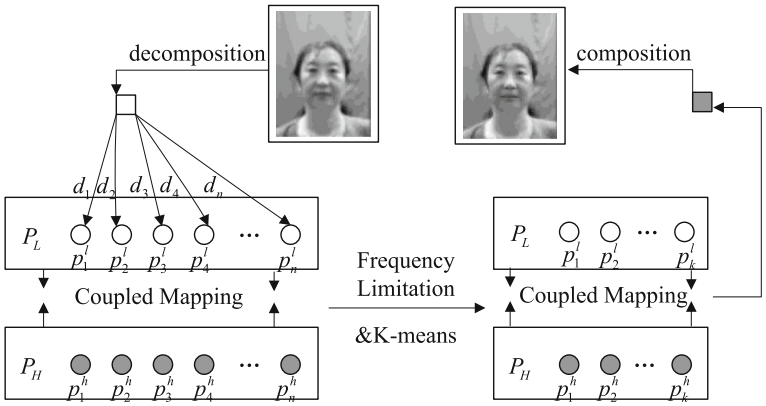

Fig. 2 Scheme of recognition phase

coefficients, which contain the information of local features cannot be estimated well by interpolation. For this reason, training processing embodied in AC coefficients only.

3.2 AC coefficient inference model

As shown in Fig. 2, we adopt a k-pass criterion in learning procedure to reconstruct target block from multiple samples instead of only one. In our experiments, the number of nearest neighbors: k is set to 3.

For $s_t h$ block, $j_t h$ image in training set $\{P^m, P^l\}$, we define d_{sj} as a metric by Eq. 2 to get suitable matching blocks, where X^l is the interpolated MRI from low-resolution input.

$$d_{sj} = \|P^l_{sj} - X^l_s\|_2. \tag{2}$$

An image I can be represented by the sum of $B(x, y)$ weighted by $W(x, y)$ formulated as:

$$I = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} W(x, y) B(x, y), \tag{3}$$

where $B(x, y)$ denotes the basis image, while $W(x, y)$ is the DCT coefficient and specifies the contribution of each basis image to I. An error measurement is defined by Eq. 4, via solving a minimization problem, we can get weight W of k nearest neighbors, respectively.

$$E(X^l_s, W_s) = \|X^l_s - \sum_{t=1}^{k} W_s(t) P^l_s(t)\|_2, \tag{4}$$

We assume that LRI blocks on the low-resolution manifold close to a locally-linear structure. And for coupled mapping LRI and MRI blocks, they share similar structure

on some manifolds [5,6]. Based on this assumption, we formulated corresponding MRI block $\widehat{X_s^m}$ as Eq. 5:

$$\widehat{X_s^m} = \sum_{t=1}^{k} W_s(t) P_s^m(t). \tag{5}$$

Combining with DC coefficient which is generated by simply cubic-spline interpolation, we obtain all DCT coefficients of the middle-resolution image. Details of the AC coefficient inference model are summarized in Algorithm 1.

Algorithm 1 AC coefficient inference model

Require: the interpolated MRI from low-resolution, X^l
Ensure: middle-resolution AC coefficient.
1: **for** each interpolated MRI block X_s^l **do**
2: Find its k nearest low-resolution neighbors based on Eq. 2 in P^l;
3: Take the corresponding k middle-resolution blocks in P^m;
4: Calculate W_s by minimizing Eq. 4.
5: Generate $\widehat{X_s^m}$ based on Eq. 5;
6: **end for**
7: **return** middle-resolution AC coefficient $\widehat{X^m}$.

4 Face hallucination in accordance with different frequency bands

Since natural images are often flat and smooth, high-frequency components exist often around edges. Hence, most energy lies at low-frequency coefficients. However, some subtle characteristics which represent high-frequency information have crucial impact on visible quality. Therefore, it is reasonable to deal with low- and high-frequency separately.

4.1 Low-frequency band generation

With the assumption that most of the energy lies in the low frequency range, an approximated solution is employed as followed Eq. 6, where $n \times n$ and $C_L(x)$ denote the block size and its DCT coefficients of the MRI input, respectively. $C(x)$ is the DCT coefficient of its corresponding HRI output with block size $2n \times 2n$ [12,18].

$$C(x) = \begin{cases} \cos\left(\frac{\pi x}{4n}\right) C_L(x), & x \in \{0, 1, \ldots, n-1\} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

As shown in Fig. 3, we present the DCT up-sampling method used for low-frequency band generation, while the high-frequency coefficients are set to be null in the proposed method. Experiments indicate a higher objective quality compared with some spatial domain interpolation methods. However, visual quality of this DCT up-sampling method is still not satisfactory because of blocking artifacts arisen from zero appending in high-frequency bands [24].

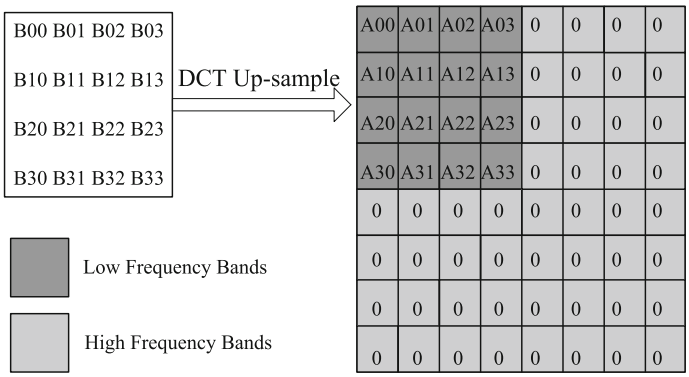

Fig. 3 Low-frequency band generation via DCT up-sampling

4.2 High-frequency band generation

In the proposed learning-based method, the generation MRI has two distinctive characteristics. DCT is suitable because of its nice energy concentration capability, and we can take different strategies on DC and AC coefficient restoration. Learning-based methods attempt to generate a model which represents the relationship between MRIs and LRIs from training set to predict MRI. Combined with patch-based methods, it often achieves quality improvement. However, transform-based and patch-based algorithms are often annoyed by blocking artifacts. As explained above, high frequency absence would also bring details missing.

To tackle this problem, we can benefit from our high-frequency components restoration model which is an interpolation-based method to recover some of high-frequency information in spatial domain. First, bicubic interpolation is used on MRI to reach original size in spatial domain. Interpolation from less distorted input can fulfill some high-frequency information. In addition, a blurring process, which aims at eliminating block artifacts apply afterwards. In our experiment, the blur kernel is Gaussian function with standard deviation 0.5. Then back to DCT domain, we only extract high-frequency components as the final high-frequency bands with low-frequency bands remained.

It is of significant necessity to combine spatial domain and DCT domain. In our second-scale restoration frame, it is able to preserve low-frequency components in DCT domain, while the high-frequency elements can benefit from the interpolation-based method in the spatial domain. Therefore, hybrid scheme is naturally expected to achieve better performance in both subjective and objective measures.

5 Experimental results

Public face databases CAS-PEAL, YaleB and FERET are selected for the experiments. For CAS-PEAL, a subset that consists of 1,040 frontal view face images (one per individual) is used. For YaleB, a subset of frontal view images from 38 persons with

64 different illuminations is used. For FERET, a subset of frontal view images of 1,196 persons is used. All images are manually aligned by the position of the eyes. Images in different databases are normalized to different resolutions.

For CAS-PEAL, images are normalized to the resolutions of 360×480 (HR) and 45×60 (LR). For YaleB, images are normalized to the resolutions of 96,128 (HR) and 12×16(LR). For FERET, images are normalized to the resolutions of 128×192 (HR) and 16×24 (LR). All the images are well aligned. Since there is no general method for aligning images with different poses, only frontal view images are used in our experiments.

For each database, images are divided into two non-overlapped sets according to their class label. In the CAS-PEAL database, images from 520 persons (one per person) are randomly selected as the training data, and for the YaleB database, images of 19 persons (64 per person) are randomly selected as the training set. In the FERET database, images from about 600 persons (one per person) are randomly selected as the training data. The rest of images are used as the testing (probe) set.

The training data pairs will be clustered using linearity clustering, so that the relationship between the data pairs in each cluster can be linearly approximated. One relationship is learned for each cluster. In the testing phase, the testing image will be classified into one of the clusters, the reconstructed image is obtained by applying proposed method on the input testing image. For each image, it will be divided into small patches, and we conduct the super-resolution patch by patch.

5.1 Face hallucination evaluation on CAS-PEAL-R1 databases

In this experiment, face images are under normal conditions, i.e., frontal pose, even illumination, and neutral expression. The Normal subset CAS-PEAL-R1 [9] contains 1,040 images for one image per person. Among these samples, about 1,020 images are selected for training and the remaining for testing. In our method, training images are well aligned before experiments. Then normalize them to 360×480 images as HRIs. By smoothing and down-sampling, their corresponding 180×240 MRIs and 45×60 LRIs are obtained. In addition, aiming at imitating the process of real video surveillance, test HRIs are segmented into 88 blocks. For each block, DC coefficient is quantized and extracted with AC coefficient abandoned. Thus, test images of 45×60 LRIs* are obtained. In the experiment, we evaluate the proposed methods in two different conditions: the uncompressed and compressed. In the context of uncompressed, the original down-sampled face images are input as test images directly. While in the context of compressed, the down-sampled face images are further compressed with JPEG.

Table 1 represents PSNRs of all 20 test images with diverse methods. It indicated that our final uncompressed result has convincing improvement in objective measures. Our compressed result is still favorable compared to the following methods with uncompressed inputs. Subjective result is shown in Fig. 4. We observe that cubic B-spline interpolation suffers from severe blurring problem. Zhangs method has a more satisfactory visual quality but still loss details. Meanwhile, slight fuzziness in smooth face region can be observed. Our method even with compressed inputs possesses best visual quality among them especially in subtle characteristics.

Table 1 Comparison of several methods in PSNR (db)

Method	Cubic B-spline	Zhang and Cham [29]	The proposed (uncompressed)	The proposed (compressed)
test1	24.36	24.57	26.95	26.80
test2	22.97	22.98	24.84	24.72
test3	23.90	24.45	27.50	27.29
test4	25.31	25.30	28.10	27.99
test5	22.90	22.98	25.90	25.97
test6	23.68	24.11	27.24	26.97
test7	24.59	25.11	28.59	28.35
test8	25.31	25.54	29.17	28.99
test9	24.93	25.42	28.74	28.59
test10	25.83	26.41	29.00	28.84
Average	24.42	24.79	27.50	27.32

5.2 Face hallucination evaluation on YaleB and FERET databases

We evaluate the proposed method on YaleB database and FERET database, and the comparison of average PSNR is provided in Table 2. Figure 5 shows some of the reconstructed images using the proposed and existing methods on the FERET database. From top to bottom are the original HR image, the results using the bicubic interpolation, the results of Zhangs method [29] and our proposed method, respectively. It is shown that the bicubic interpolation gives a relatively blurred image and high-frequency details cannot be recovered. Zhangs method could recover some high-frequency details. However, this method generates some artifacts that degrade human visual quality. The visual quality of the reconstructed images of our proposed method is good. It is shown that the proposed method gives a good visual quality image, which also looks like the original one (Table 2).

Normally, the performance of learning-based method often depends on how well the input LRI matches the samples in the training set. Obviously, the more training samples are collected, the more robust the learning-based algorithm is. However, a huge training set requires taxing computation and heavy memory load. In our method, all collected training images are first aligned by affine transform based on three marked points: the centers of the two eyes and the center of the mouth. Then each image is cropped to normalized size as the HRI. We can see that the blocks cropped from the face images do not have much variation since human facial features are similar.

The learning-based method is applied on the first-scale restoration procedure which only attempts to generate the middle-resolution image. Hence, the block number of the training set is reduced. In addition, our training set contains only AC coefficients to represent local features. Hence, the raw training set should have much redundancy and it is possible to learn those most representative blocks and build a compact block dictionary by clustering.

In the second-scale restoration, the image resizing in the compressed domain is employed to generate low-frequency band. Using spatial relationship of the block DCT

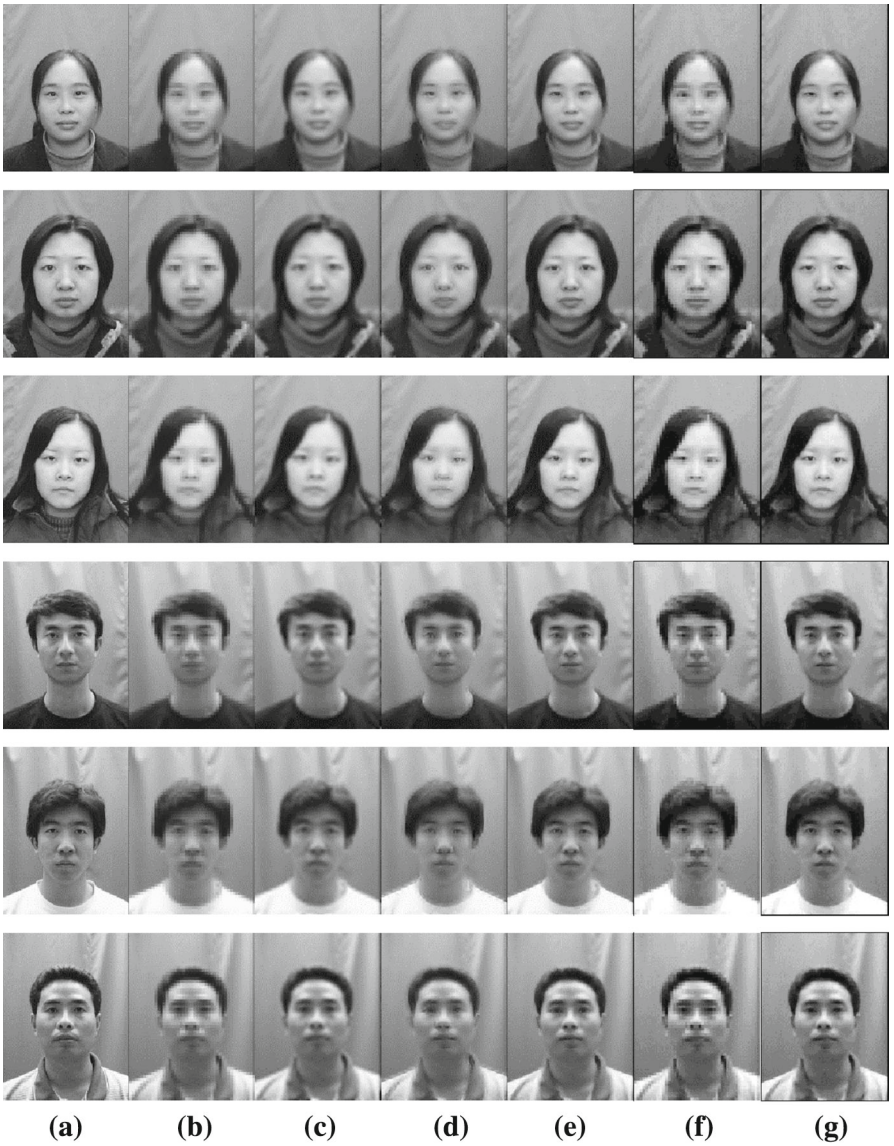

Fig. 4 Face hallucination results of frontal face images with different compression ratio. **a** Original uncompressed HRIs. **b** Low-resolution face inputs. **c** Cubic B-spline interpolation. **d** Zhang and Cham [29]. **e** Our final uncompressed results. **f** Compressed low-resolution face inputs. **g** Our final compressed results

coefficients and sub-band approximations, a simple computational DCT up-sample framework is adopted for image doubling operations. An extensional application is the conversion of images from one format to another in the compressed domain. Whats more, patches will be optimized with a high-frequency band generation model as a post-processing framework. This process significantly improves the results, whereas

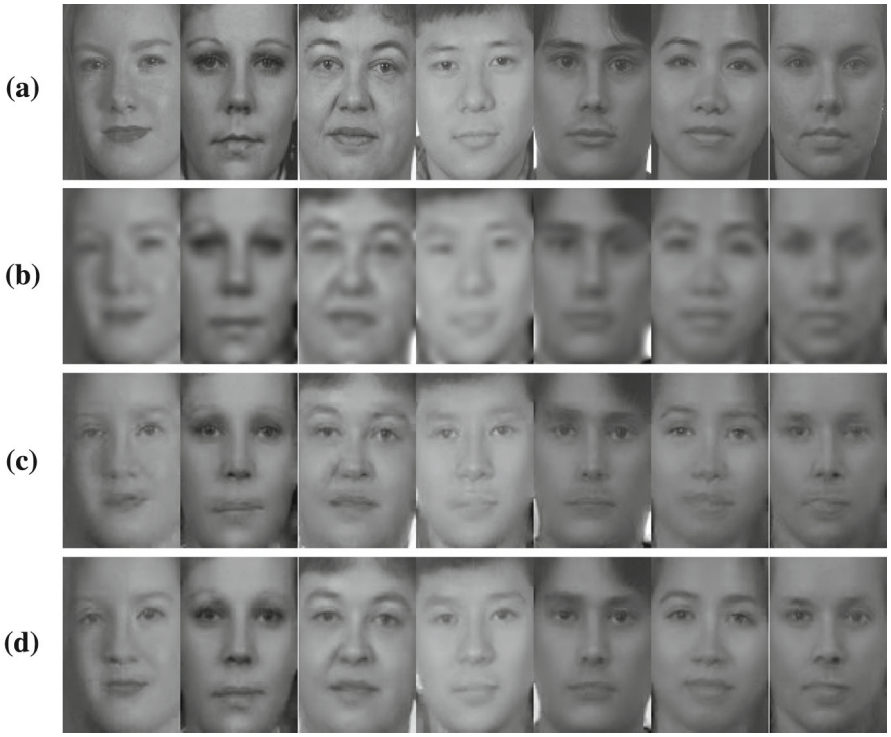

Fig. 5 Face hallucination results of frontal face images with different compression ratio FERET database. **a** Original uncompressed HRIs. **b** Cubic B-spline interpolation. **c** Zhang and Cham [29]. **d** Our final compressed results

Table 2 Comparison of average PSNR (db) in YaleB database and FERET database

Method	Bicubic	Zhang and Cham [29]	The proposed (uncompressed)	The proposed (compressed)
YaleB database	22.43	23.06	25.85	24.67
FERET database	25.54	25.98	28.57	26.32

we can still see some artifacts generated in the incorrect patch matching process. In addition, it should be noted that there is no complex process in our second-scale restoration framework. We simply use the gauss function to induce block artifacts. Our results have a large room to improve by coupling with some post-processing techniques.

5.3 Relationship between face hallucination and recognition

Since one of the most important aims of face hallucination is to assist in recognition, it is an important issue to determine what kind of face images cannot be recognized by human perceptions or machines. For human perceptions, faces are able to be recog-

Table 3 Recognition rates in terms of different recognition algorithms

Resolution	45 × 60	90 × 120	180 × 240	360 × 480
PCA	46.9	61.2	66.3	67.1
SRC	74.1	79.6	86.3	89.1
LPP	70.3	77.4	84.6	85.6
LDA	71.2	77.1	84.1	85.1

nized in reasonably high-resolution images since there are more details available in those images. When the image resolution decreases to a certain threshold, faces in the CAS-PEAL-R1 database are difficult to be recognized by eyes. The faces become hardly recognizable when the resolution is below 45 × 60.

If the resolutions are set as the variable, it is found that when the resolutions vary from 45 × 60 to 360 × 480, the recognition rates have different trends in terms of different recognition approaches. Table 3 shows the results of our experiments in which we recognize faces using different recognition approaches in the CAS-PEAL-R1 database. We implemented LDA, LPP, PCA and SRC for images with resolutions of 45×60, 90×120, 180×240 and 360×480. Recognition by the PCA algorithm has an obvious trend, where the recognition rate increases when face resolution increases. LDA, LPP and SRC produce satisfactory recognition rates. It can be seen that in the low resolution, SRC has a very good performance. Specifically in the resolution of 45 × 60, SRC produces a similar recognition rate as in the highest resolution of 360 × 480.

From this experiment, it has been shown that though it is not always the case that higher resolutions lead to the higher recognition rates, the recognition performance is in general becoming better with increasing high resolution. Resolution is not the only factor that influences the face recognition performances.

In order to clearly evaluate the performance of face hallucination in recognition, another experiment is conducted with CAS-PEAL-R1 dataset. We compared the assistant recognition performance of the proposed method with that of the other three typical face hallucination approaches: learning-based DCT [29], sparse representation super-resolution (ScSR) [26] and cubic interpolation (cubic). The image resolutions are enhanced from 45 × 60 to180 × 240. As a result a set of hallucinated high-resolution faces is derived. In the recognition experiment, the testing faces are randomly selected from the hallucinated faces and the training data are randomly selected from the original high-resolution images. LDA [4], LPP [10], PCA [21] and SRC [23] face recognition algorithms are adopted. Similarly, the recognition experiment is repeated twenty times and the recognition rates are averaged from them (Table 4).

The recognition rates of the low-resolution faces, hallucinated faces and high-resolution faces are compared in terms of different face recognition approaches, when the images resolution is enhanced from 45 × 60 to 180 × 240. Table 4 shows the experimental result for such cases. For most hallucinated 180 × 240 images, face recognition rates increase when compared with the original 45 × 60 faces using all four typical recognition approaches. This means that if the original face is in the

Table 4 Recognition performance of hallucinated faces in CAS-PEAL-R1 database from 45 × 60 to 180 × 240

Hallucination approaches	PCA	SRC	LPP	LDA
Low	46.9	74.1	70.3	71.2
Cubic	47.1	74.5	70.2	71.6
ScSR	51.5	78.9	75.1	75.7
learning-based DCT	53.2	79.1	77.2	77.1
Proposed	55.9	80.6	79.1	77.1
High	66.3	86.3	84.6	84.1

resolution of 45 × 60, the face recognition rates can be improved significantly, by enhancing the image resolutions through hallucinating faces.

6 Conclusions

Aiming at improving face recognition performance with the photography in social media data, this paper proposes a novel twofold face hallucinate framework, which is introduced to restore image progressively. The first-scale recovering image with middle-resolution is generated by a learning-based method in DCT domain. The second-scale recovering image is obtained differently based on frequency bands analysis. Extensive experiments show that the proposed algorithm effectively helps enhance face recognition with a selected classifier. The proposed face hallucinate method can significantly improve the face recognition accuracy which make a vital foundation for the incorporation of face recognition in these large-scale web services.

Acknowledgments This work was supported in part by the National Science Council of the Republic of China under Grant No. 103-2917-I-564- 058, and the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2061978), the National Natural Science Foundation of China under Grant No. 61272386 and 61100096. We are grateful for Zhang kindly providing source code of [29].

References

1. Acquisti A, Gross R, Stutzman F (2011) Faces of facebook: privacy in the age of augmented reality. BlackHat USA
2. Anastassiou D, Jensen K (1995) Sub-pixel edge localization and the interpolation of still images. IEEE Trans Image Process 4(3):285C295
3. Baker S, Kanade T (2000) Hallucinating faces. In: Proceedings fourth IEEE international conference on automatic face and gesture recognition, IEEE, pp 83–88
4. Belhumeur PN, Hespanha JP, Kriegman D (1997) Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720
5. Chang H, Yeung DY, Xiong Y (2004) Super-resolution through neighbor embedding. In: Proceedings of the 2004 IEEE Computer Society conference on computer vision and pattern recognition. CVPR 2004, vol 1, IEEE, pp I-I
6. Chang TL, Liu TL, Chuang JH (2006) Direct energy minimization for super-resolution on nonlinear manifolds. In: Computer vision-ECCV 2006, Springer, pp 281–294

7. Cohen S, Hamilton JT, Turner F (2011) Computational journalism. Commun ACM 54(10):66–71
8. Dantone M, Bossard L, Quack T, Van Gool L (2011) Augmented faces. In: 2011 IEEE International conference on computer vision workshops (ICCV workshops), IEEE, pp 24–31
9. Gao W, Cao B, Shan S, Chen X, Zhou D, Zhang X, Zhao D (2008) The CAS-PEAL large-scale Chinese face database and baseline evaluations. IEEE Trans Syst Man Cybern Part A Syst Hum 38(1):149
10. He X, Niyogi P (2003) Locality preserving projections. NIPS 16:234–241
11. Hu Y, Lam KM, Qiu G, Shen T (2011) From local pixel structure to global image super-resolution: a new face hallucination framework. IEEE Trans Image Process 20(2):433–445
12. Jung SH, Mitra SK, Mukherjee D (1996) Subband dct: definition, analysis, and applications. IEEE Trans Circuits Syst Video Technol 6(3):273–286
13. Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M et al (2009) Life in the network: the coming age of computational social science. Science (New York, NY) 323(5915):721
14. Li Y, Lin X (2004) Face hallucination with pose variation. In: Sixth IEEE international conference on automatic face and gesture recognition, IEEE, pp 723–728
15. Liu C, Shum HY, Freeman WT (2007) Face hallucination: theory and practice. Int J Comput Vis 75(1):115–134
16. Ma X, Zhang J, Qi C (2010) Hallucinating face by position-patch. Pattern Recognit 43(6):2224–2236
17. Mavridis N, Kazmi W, Toulis P (2010) Friends with faces: how social networks can enhance face recognition and vice versa. In: Computational social network analysis, Springer, pp 453–482
18. Mukherjee J, Mitra SK (2002) Image resizing in the compressed domain using subband DCT. IEEE Trans Circuits Syst Video Technol 12(7):620–627
19. Park JS, Lee SW (2008) An example-based face hallucination method for single-frame, low-resolution facial images. IEEE Trans Image Process 17(10):1806–1816
20. Stone Z, Zickler T, Darrell T (2010) Toward large-scale face recognition using social network context. Proc IEEE 98(8):1408–1415
21. Turk MA, Pentland AP (1991) Face recognition using eigenfaces. In: IEEE Computer Society conference on computer vision and pattern recognition, IEEE, pp 586–591
22. Wang X, Tang X (2005) Hallucinating face by eigentransformation. IEEE Trans Syst Man Cybern Part C: Appl Rev 35(3):425–434
23. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2009) Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell 31(2):210–227
24. Wu Z, Yu H, Chen CW (2010) A new hybrid DCT-Wiener-based interpolation scheme for video intra frame up-sampling. IEEE Signal Process Lett 17(10):827–830
25. Xu X, Liu W, Li L (2013) Face hallucination: how much it can improve face recognition. In: 2013 3rd Australian control conference (AUCC), IEEE, pp 93–98
26. Yang J, Tang H, Ma Y, Huang T (2008) Face hallucination via sparse coding. In: 15th IEEE international conference on image processing. ICIP 2008, IEEE, pp 1264–1267
27. Yang J, Wright J, Huang TS, Ma Y (2010) Image super-resolution via sparse representation. IEEE Trans Image Process 19(11):2861–2873
28. Zhang D, Wu X (2006) An edge-guided image interpolation algorithm via directional filtering and data fusion. IEEE Trans Image Process 15(8):2226–2238
29. Zhang W, Cham WK (2011) Hallucinating face in the DCT domain. IEEE Trans Image Process 20(10):2769–2779
30. Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. ACM Comput Surv (CSUR) 35(4):399–458
31. Zhuang Y, Zhang J, Wu F (2007) Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation. Pattern Recognit 40(11):3178–3194