FACE IMAGE SUPER RESOLUTION BY LINEAR TRANSFORMATION

Hua Huang*, Ning Wu*, Xin Fan** and Chun Qi*

*School of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an 710049, China
**School of Software, Dalian University of Technology, Dalian 116620, China

ABSTRACT

A novel two-step super-resolution (SR) method for face images is proposed in this paper. The critical issue of global face reconstruction in the two-step SR framework is to construct the relationship between high resolution (HR) and low resolution (LR) features. We choose the Principal Component Analysis (PCA) coefficients of LR/HR face images as the features for global faces. These features are considered as inputs and outputs of an unknown linear system. The mapping between the inputs and outputs is estimated from training sets as the system response. The HR features corresponding to a test LR image can be obtained by applying the learnt mapping to the LR features, and hence we can reconstruct the global face. Ultimately, an HR face image is generated by using the patch-based neighbor reconstruction that imposes facial details into the global face. Experiments indicate that our method produces HR faces of higher quality and is easier to implement than traditional methods based on two-step framework.

Index Terms— Face Super Resolution, linear mapping

1. INTRODUCTION

Image super-resolution (SR) is the process of combing one or a set of low resolution (LR) images to obtain a high resolution (HR) image. In practice, the resolution of face images captured by surveillance devices is too low to be recognized. Since Hallucinating faces were proposed by Baker and Kanade [1], the face SR problem is highly concerned by researchers and learning-based approaches are prevailing in face SR reconstruction during this decade. As face images exhibit geometrical structures unique to generic images, it is very important to represent (or learn) and reconstruct the structural feature (global information) besides detailed texture information. Liu et al. [2, 3] were the first to propose a two-step face image SR method that divided the SR problem into reconstructing global information and local information, both of which can be learnt from training sets. The two-step SR method is a nice framework upon which many SR algorithms are built.

In order to capture the appearance variations on global facial structures, face images are typically transformed into feature spaces by multivariate statistical techniques, e.g., principal component analysis (PCA) [2, 3, 4] and multi-linear analysis [5]. Since PCA can largely reduce the data dimension as well as preserve global face information, it is widely used as global face features in the existing SR methods based on the two-step framework. In this paper, we concentrate on the first step, i.e., reconstructing global face structures, and for the second step we employ the neighbor reconstruction [6] to compensate facial details.

The central issue of image SR, especially for global face reconstruction, is to construct the relationship between HR and LR images (or features). In the original two-step face hallucination algorithm [2], the relationship between HR and LR PCA features is bridged base on an image formation model assumed known. The global face images are estimated by Maximum a posterior (MAP). Zhuang et al. [6] introduced manifold learning techniques into the two-step hallucination. The local preserving projection (LPP) [7] and radial basis functions (RBFs) [8] are employed to establish the relationship between the corresponding LR and HR PCA coefficients to reconstruct global information. Instead of imposing an explicit mapping, Wang et al. [4] fit the input face image as a linear combination of LR training face images in the eigen transformation domain. The HR image is generated by replacing the LR training images with corresponding HR ones and retaining the same combination coefficients. In these approaches, an explicit down-sampling model has to be assumed known, or the mappings between LR/HR faces are relatively complicated and require high computational expense.

It is worth noting that the underlying structures of the PCA coefficients of LR/HR frontal faces are highly correlated. There exists a simpler form of mapping to represent the relationship between LR/HR face features. Specifically, we apply Procrustes analysis [9] to the LR/HR pairs of a given training set in order to estimate the linear mapping. Procrustes analysis, a least square estimator, is able to well correlate two point sets. Wang et al. [9] use Procrustes analysis for robust manifold alignment by taking advantage of the fact that it preserves the intrinsic structure of each data set. This property is also valuable for face SR based on the assumption of local geometry (or neighborhood) preservation [7]. The SR
reconstruction of global face turns out to projecting the PCA coefficient of a probe LR face to the HR space using the learnt mapping. The final SR face images can be obtained by cascading residual compensation to the recovered global face. Experiments show that our simple yet effective method produces higher-quality SR faces in subjective visual and objective PSNR aspects and is easier to implement compared with previous learning-based SR algorithms based on the two-step framework.

2. PROBLEM FORMULATION

The SR problem for frontal face images can be depicted as the inference of an HR image from one LR image $I$, given a training set of HR images and their corresponding LR versions, denoted by two vectors, $I^H = \{I^H_i\}_{i=1}^m \in \mathbb{R}^{n \times m}$ and $I^L = \{I^L_i\}_{i=1}^m \in \mathbb{R}^{n \times m}$, where $m$ is the training number and $n$ is the dimension of the image.

Given HR and LR training sets we can use the principal component analysis (PCA) to obtain the mean face $\mu^H$ and $\mu^L$ as well as the base vectors of the face subspace $P^H$ and $P^L$ composed of the orthogonal eigenvectors of the HR/LR face covariance matrices. We have the PCA coefficients $Y^H = \{y^H_i\}_{i=1}^m = \{y^H_1, y^H_2, \ldots, y^H_m\}$ and $Y^L = \{y^L_i\}_{i=1}^m = \{y^L_1, y^L_2, \ldots, y^L_m\}$ of HR and LR training images by projecting faces onto the face subspace as:

$$y^H_i = P^H(I^H_i - \mu^H)$$ 
$$y^L_i = P^L(I^L_i - \mu^L)$$

Specifically, we exploit the Procrustes analysis to build the PCA coefficient mapping by which the global reconstruction can be obtained, and apply the neighbor reconstruction for residual compensation.

3. ALGORITHMS

Our algorithm follows the two-step SR framework, which includes global face reconstruction and detail compensation. Specifically, we exploit the Procrustes analysis to build the PCA coefficient mapping by which the global reconstruction can be obtained, and apply the neighbor reconstruction for residual compensation.

3.1. GLOBAL FACE

Our global reconstruction method originates from the “black-box” theory. We are able to estimate a system model by mathematical techniques from observed inputs/outputs of the system if it is difficult or unnecessary to establish the physical model. This model may not have rigorous physical explanation for the problem, but can directly bridge the inputs and outputs, which is called “black-box theory”. In this paper, we take the down-sampling process as a “black box”, and the PCA coefficients of LR and HR face images are considered as corresponding inputs and outputs of an unknown system. Thus, the problem of global face reconstruction is posed as a problem of estimating the response of this unknown system.

For the generic SR problem, it is not practical to establish mapping from LR images (features) to HR ones using a relatively simple mathematical form. But there is a different story for images of frontal faces that are highly structured. The intensities at the same position exhibit similar texture for relatively simple mathematical form. But there is a different story for images of frontal faces that are highly structured.

It is reasonable to assume that the system is linear based on the above mentioned analysis. In our method, although the non-linear face SR reconstruction is approximated by a linear
system to obtain global face in step-one, the non-linear part is supplemented by residue compensation in step-two. Given the corresponding PCA coefficients of HR and LR training sets, we formulate the reconstruction of global face image as computing a transformation matrix $R$ that minimize the reconstruction error defined as:

$$\sum_{i=1}^{m} \|x_i^H - Rx_i^L\|_2 = \|((X^H)^T - (X^L)^T R)^T\|_F$$  

where $R$ presents the response of the system and $R^T$ is the transpose of $R$. We set the dimensions of PCA coefficients of LR and HR face images equal without loss of generality. Based on Procrustes analysis [9], we have:

$$\|((X^H)^T - (X^L)^T R)^T\|_F = \|((X^H)^T - k(X^L)^T Q)\|_F$$  

where the matrix $Q$ is orthonormal and $k$ is a scaling factor [9]. Thus, the minimization problem is equivalent to:

$$\{k_{opt}, Q_{opt}\} = \arg \min_{k,Q} \|((X^H)^T - k(X^L)^T Q)\|^2$$

$$= \arg \min_{k,Q} (k^2 \text{trace}(X^L(X^L)^T) - 2k \text{trace}(Q^T X^L(X^H)^T))$$

Denoting $y = k^2 \text{trace}(X^L(X^L)^T) - 2k \text{trace}(Q^T X^L(X^H)^T)$ and assuming $Q$ known, we can get $k$ by letting $\frac{\partial y}{\partial k} = 0$:

$$k_{opt} = \text{trace}(Q^T X^L(X^H)^T)/\text{trace}(X^L(X^L)^T)$$

Then the solution for $Q$ is shown by equation as follows:

$$Q_{opt} = \arg \max_{Q} (\text{trace}(Q^T X^L(X^H)^T))^2$$  

We denote the singular value decomposition (SVD) of $X^L(X^H)^T$: $X^L(X^H)^T = U \Sigma V^T$.

Case 1: $\text{trace}(Q^T X^L(X^H)^T) \geq 0$

Equation (6) is equivalent to:

$$Q_{opt} = \arg \max_{Q} (\text{trace}(Q^T X^L(X^H)^T))$$

Because: $\text{trace}(Q^T X^L(X^H)^T) = \text{trace}(Q^T U \Sigma V^T) = \text{trace}(V^T Q^T U \Sigma)$, denoting $Z = V^T Q^T U$. It is known that $V$, $Q$ and $U$ are all orthonormal matrixes, so $Z$ is also orthometric. Based on [9] we can get: $\text{trace}(Z \Sigma) = Z_{1,1} \Sigma_{1,1} + Z_{2,2} \Sigma_{2,2} + \cdots \leq \Sigma_{1,1} + \Sigma_{2,2} \cdots = \text{trace}(I \Sigma)$. Based on $V^T V U^T U = I$, then we have:

$$Q = U V^T$$

Case 2: $\text{trace}(Q^T X^L(X^H)^T) < 0$

Following the similar derivation, we obtain:

$$Q = -U V^T$$

Because Eqs.(9) and (10) get the same results based on (7), we choose $Q_{opt} = U V^T$ in this paper. Summarizing Eqs.(4), (6) and (9) we have the solution to the mapping matrix $R$.

For an input test LR face image the corresponding PCA coefficient can be obtained:

$$y^l = P^L(I - \mu^L)$$

With the mapping matrix $R$, the PCA coefficient of the corresponding SR global face image is:

$$y^h = R(y^l - v^L) + v^H$$

The HR global face of the given LR image can be obtained by transforming $y^h$ into pixel-space.

### 3.2. RESIDUE COMPENSATION

A residual compensation can be appended in order to provide further facial details. Subtracting the reconstructed SR patches of the training LR sets obtained by Eqs.(11) and (12) from their corresponding original HR training images, we have the training residual images. Following the procedure in [6], we divide the test LR residual image into patches and

<table>
<thead>
<tr>
<th>Method</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
<th>Test 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our method</td>
<td>29.40</td>
<td>28.42</td>
<td>29.94</td>
<td>28.49</td>
</tr>
<tr>
<td>Zhuang’s method</td>
<td>28.29</td>
<td>27.359</td>
<td>29.19</td>
<td>27.22</td>
</tr>
<tr>
<td>Liu’s method</td>
<td>27.47</td>
<td>26.34</td>
<td>28.01</td>
<td>26.23</td>
</tr>
<tr>
<td>Wang’s method</td>
<td>27.67</td>
<td>26.58</td>
<td>28.15</td>
<td>26.16</td>
</tr>
</tbody>
</table>
reconstruct every corresponding SR residual patch by the linear combination of its neighbors in the training HR residual patches. The combination weights are taken as those that minimize the error of reconstructing the corresponding LR residual patch from its neighbors.

4. EXPERIMENTS AND ANALYSIS

We demonstrate the performance of the proposed SR method by comparing it with three representative face SR algorithms in [2, 4, 6]. The experiments are performed on CAS-PEAL face database [10] that has 1040 frontal face images with the dimension $128 \times 128$, and the LR images are obtained by 4-time down-sampling with the size of $32 \times 32$. We randomly select 40 images for testing and the rest 1000 for training. In Liu’s method [2], the variance accumulation contribution rate of PCA is 98% and $\lambda = 0.1$. In Zhuang’s method [6], the number of neighbors is 160 and the number of eigen values $h = 500$. In our method the variance accumulation contribution rate of PCA is 98%. The histograms of all the global and SR results were aligned to those of their corresponding LR observations for the sake of fair comparison.

Fig.1 gives the global faces obtained by ours, Liu’s and Zhuang’s methods. The global face images of the 40 test images in different methods are plotted by boxplot in Fig.2. It is shown that our algorithm can well reconstruct global information. Compared with the others our method is the simplest, but outputs the best quality.

Fig.3 shows the SR results of our method and the other three methods. In our method the size of the patch is $8 \times 8$ in residue compensation, which is the same as in Zhuang’s. Table 1 shows the PSNR values of the SR reconstruction results for the test images in Fig.3, from which we can see that the proposed method gains an improvement of 1-2 dB over other approaches and our results have finely chiseled features and least ringing effects, which are quite close to the original HR images. To further illustrate effectiveness of our method, we applied these methods to all 40 test images and plot the results by boxplot in Fig.4. Both Fig.4 and Table 1 show that the PSNR of our results is higher than others. These results indicate that our method produces higher-quality SR faces in subjective visual and objective PSNR aspects and is easier to implement than the other methods.

5. CONCLUSION

We use Procrustes analysis to establish the linear mapping matrix from the PCA coefficients of LR images to those of the corresponding HR ones based on the black-box theory. The SR reconstruction of global faces is achieved by applying the mapping to a given test LR image. Our method is quite simple and efficient, and the experiments show that the method is also able to produce higher quality reconstruction in both subjective visual and objective PSNR aspects, compared with state-of-the-art two-step SR methods for faces.

6. REFERENCES