An effective framework of IBP for single facial image super resolution

Cong Wang[1], Weifeng Li[1], Longbiao Wang[2] and Qingming Liao[1]

[1]Department of Electronic Engineering/Graduate School at Shenzhen, Tsinghua University, China
[2]Nagaoka University of Technology, Japan

Abstract. Super-resolution was used to refine the LRI (low resolution image) which was acquired from the existing electric devices. Moreover, it is significant to do this work on the human face. To solve the special facial problem, a novel framework based on IBP (Iterative Back Projection) is proposed, in which the edge detection and edge preserving are taken into account. In the iterative part, we adopt wavelet for edge detection to preserve the details and eliminate the loss of visual information. Furthermore, PCA (Principle Component Analysis) technique is introduced to extract more local high frequency information to refine the LRI . Experimental results show that our framework could efficiently enhance the visual effect. Both PSNR and SSIM of the super-resolution of LRI could be increased. Moreover, some experiments on face recognition are conducted and our method achieves more inspiring results.

Keywords: super-resolution, IBP, wavelet edge detection, PCA.

1 INTRODUCTION

Super resolution is known as a technique in signal processing area, which makes use of the input LRI to generate HRI(High Resolution image). There are a range of applications in remote sensing, public security, video communication, etc. It was first proposed by Harris [1] and Goodman [2] to resolve the single image problem. It obtained a huge improvement in 1984, when Tsai and Huang utilized multi-frame images to reconstruct the HRI in the spectral space [3].

[1] Weifeng Li (✉)
Graduate School at Shenzhen, Tsinghua University, Lishui Road 1147, Xili town,Shenzhen, Guangdong, China
e-mail: li.weifeng@sz.tsinghua.edu.cn

In recent years, public safety has aroused more and more attention. As one of the most important means to security the public safety, video surveillance could provide a lot of information when the crime is committed. Moreover, it is significant to refine human face of low resolution in the region of interest (ROI), which has received much attention in the past few years.

According to the number of the input LRI, the SR (Super Resolution) methods for facial images can be divided into two categories: single facial SR and multi-frame facial SR.

For the single image super resolution, interpolation is known as a kind of popular method, such as Bilinear, NEDI (New Edge-directed Interpolation) [4], etc. Although computational complexity of Bilinear interpolation is low, blurring and zigzagging at the contour area will degrade the quality of the results. The NEDI considered the local similarity of the corresponding LRI and HRI, and performed well at edge preserving. Inspired by the wide and efficient application of sparse coding, many researchers applied the sparse coding to enhance the LRI. Zhou proposed a method based on sparse coding, which contains two-step procedure combined with dictionary learning [5]. It took the incoherent property of each subset of samples into consideration by learning a sub-dictionary. Different from Zhou's work, Lu's [6] method, not only retained the geometrical property of the dictionary, but also preserved the sparse coefficients of the data. Hu [7] adopted the local pixel structure and global constraints to solve the face hallucination. To optimize the result, it took the iterative procedure to learn the local and global constraints. The above methods are operated in spatial domain. In [8], Zhang induced the face hallucination in DCT domain to reduce the computational complexity. In his works, The MRF framework was simplified by estimating the AC coefficients with method of LLE. Replaced the all 63 AC coefficients by the first 15 reduced the computational complexity significantly with little loss of information. In addition, researchers considered the feature-domain to solve the face super resolution. Nguyen focused on non-linear Gabor feature domain and improved the face recognition significantly [9].

There are also many scholars are interested in the multi-frame super-resolution. In [10], a two-step procedure was proposed, where fitting a surface for each point followed by determining with the MAP criterion.

In this paper, we focus on single frame super resolution for human face, and a novel framework was proposed based on IBP. The IBP method is fast in both computing and convergence, but its ability in retrieving the high frequency is poor because much high frequency information is lost in the down-sampling procedure. To preserve the high frequency part, we take wavelet for edge detection in the iterative process. Also, PCA technique is introduced to extract more details to refine the superresolution results.

The rest of the paper is organized as follows: We briefly explain how the IBP works in Section 2; In Section 3, we describe how our framework works. The experimental results are illustrated in Section 4; Finally, we conclude this paper and discuss the work in future.

2 ITERATIVE BACK PROJECTION

According to the practical generation of the LRI, it can be taken as the combination of several ways of effect, such as motion blurring and down-sampling. Denoting g as the filter of the blur effect, we can formulate the generation of the LRI as

$$I_L = (I_H * g) \downarrow_s \tag{1}$$

where I_L and I_H are respectively LRI and HRI, and $\downarrow_s$ represents the down-sampling by factor s.

According to the aim of the super-resolution, we attempt to recover the HRI by minimizing the reconstruction error. Mathematically, the reconstruction error is formulated as

$$error = I_L - (I_{HL} * g) \downarrow_s \tag{2}$$

where I_{HL} is the estimated HRI and I_L is the input LRI, respectively. When the result of Eq.2 is small enough, we take the result as acceptable HRI.

IBP is the method to minimize the quantization error, which focuses on minimizing the estimation error in Eq.2. It was proposed by Irani and Peleg in 1991[11]. In this method, people estimates the high resolution images by back projecting the reconstruction error iteratively. It performs well on the SR problems and has many applications. However, we have to realize that due to the down-sampling, the result obtained by IBP is always too smooth to satisfy requirement. Obtaining better performance at the edge area is helpful for super-resolution.

We denote $I^{(n)}$ as the result after the $n-th$ iterative procedure. $I^{(n)}$ is down-sampled and degradation blurred to simulate the LRI. The estimated error can be obtained as

$$I_e = (I_L - I_L^{(n)}) \uparrow_2 \tag{3}$$

Then we can integrate the estimated error in (3) and the high frequency component obtained from the interpolation of the input image $HF(I^{(0)})$ into the next iterative process to obtain the result after the $(n+1)-th$ iterative procedure. Eq.4 summarizes the above process:

$$I^{(n+1)} = I^{(n)} + I_e + HF(I^{(0)}) \tag{4}$$

3 THE OVERALL FRAMEWORK

Although the IBP is fast in convergence and easy to implement, it has some disadvantages, such as the ringing artefacts in the edge area. Moreover, the smoothing effect would occur due to the down-sampling in the iterative procedure. To improve the reconstructed result, we proposed a novel framework based on IBP.

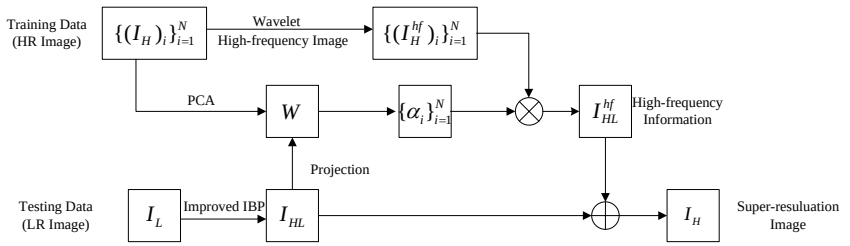

Fig. 1 The overall framework

Canny is known as an effective technique in edge-detection. However, canny-based edge detection methods are often affected by the noise. Therefore, we introduce wavelet for edge detection, which not only can preserve the edge information better in multi-scale, but also is insensitive to the noise. Meanwhile, to ensure the effectiveness of the procedure, we adopt the Bicubic interpolation as the initial estimation of the HRI.

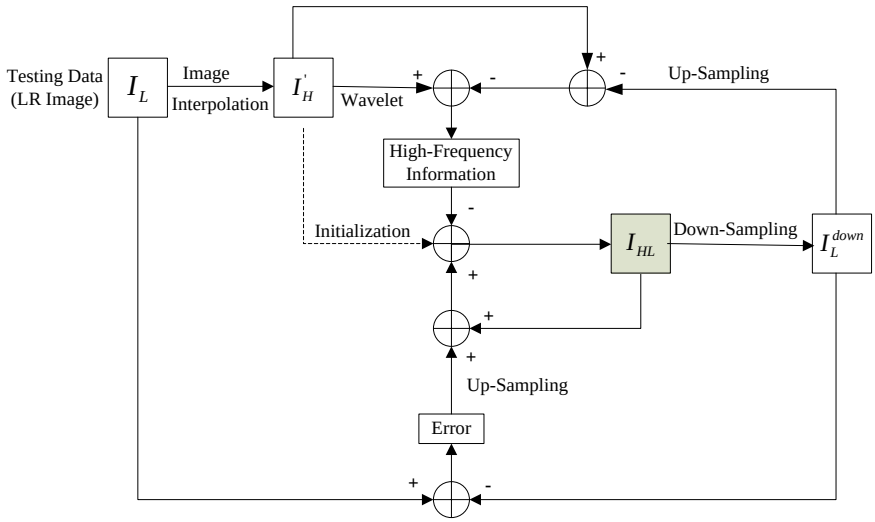

Fig. 2 The improved IBP

According to our proposed framework, the first step of iterative procedure, which is the initialization, can be written as follows:

$$I_H^{(1)} = I_H^{'} + I_e - I_H^{(0)} \tag{5}$$

where $I_H^{'}$ is the initial estimated HRI interpolated by the input LRI; I_e is the estimated error; $I_H^{(0)}$ represents the estimated initial high frequency component:

$$I_H^{(0)} = (\mathrm{I}_H^{'} * HF) \tag{6}$$

where HF is the wavelet edge detection. In this paper, we take the wavelet edge detection. Similarly, the $n-th$ iteration $I_H^{'(n)}$ is taken as follows:

$$I_H^{'(n)} = I_H^{'(n-1)} + I_e^{(n-1)} - I_H^{(n-1)} \tag{7}$$

The estimation of the high frequency can be formulated as

$$I_H^{(n)} = I_H^{(0)} - (I_H^{'} - I_L^{down} \uparrow_2) \tag{8}$$

During the iterative procedure, to minimize the reconstruction error and accelerate convergence, we extract the acceptable pixels to avoid redundant computing and unnecessary error when the reconstruction error is small enough. This is good for saving both space and time complexity. When the number of iterations achieves the maximum or the reconstruction error is small enough, the iterative procedure is terminated. In this way, we acquire the global reconstruction result of the input facial LRI, I_{HL}.

At last, we utilize PCA technique to extract more local information to refine it. The procedure can be described as follows:

1) 200 facial images picked from the FERET are taken as the training set, which is denoted as $\{(\mathrm{I}_H)_i\}_{i=1}^N$, where $N = 200$ is the number of the training test. We extract the local facial information $\{(\mathrm{I}_H^{hf})_i\}_{i=1}^N$ with wavelet from the training

2) The result I_{HL} is projected on the eigenvector matrix W, which is obtained from the training set by PCA, and the weight α is calculated. And the edge information I_{HL}^{hf} is obtained by the corresponding wavelet edge images with the same weight α.

3) At last, we synthesize both the global I_{HL} and local information I_{HL}^{hf} to acquire the final HRI I_H.

4 EXPERIMENT RESULT

4.1 Database used and settings

To begin the work, we choose the publicly available face image database FERET.

In our experiment, the size of the test images is 36x36, which is obtained by down-sampling the HRI, together with the 3x3 Gauss blurring of standard deviation $\sigma = 1.0$. The LRI is viewed as the test images and the HRI is viewed as the ground-truth images. We attempt to obtain the corresponding reconstruction images with the single input LRI. We take the 300 targets of the FERET to conduct the experiment. For the recognition part, the size of the images in the gallery set is 72x72.

To validate the effectiveness of our methods in the preserving of the global and local feature in texture, we also induce the LBP (Local Binary Pattern) to calculate the recognition rate. Besides, we compare with some other interpolation methods, such as the NEDI, Bicubic and [12]. We take the regular criteria PSNR to evaluate our method. Furthermore, as facial images are different from the other natural images, we also induced recognition rate as another criteria. Furthermore, since PSNR is poor in verifying structural preserving property due to the direct calculation of the intensity difference, we utilize SSIM as the third criteria, which is more consistent with human visual perception.

4.2 Experimental Results

We adopt the wavelet and PCA to preserve more contour and edges, and we obtain suitable result according to the evaluation criteria. In **Fig. 2**, we show some comparisons of the result to verify the effectiveness of the proposed method. As we can see from it, our proposed method can achieve better visual effects, especially, the result is sharper in the edge area when compared with some other methods.

To illustrate the effectiveness of the proposed method, we compare our results with Bicubic, NEDI, IBP and CIBP[12] method. Our proposed method achieve higher PSNR and SSIM.

We extract the texture feature of the facial images by LBP histogram and take the intersection distance as the distance measure to implant face recognition system. We take the recognition rate of Bicubic as the base-line. As shown in Table 3, our method achieve higher recognition rate than other methods.

Table 1 The SSIM results compared with several other methods

SSIM	*Bicubic*	*NEDI*	*IBP*	*CIBP*	*Proposed*
ID_0201	0.818	0.830	0.873	0.913	**0.927**
ID_0202	0.832	0.847	0.864	0.901	**0.910**
ID_0203	0.804	0.836	0.851	0.892	**0.918**
ID_0204	0.789	0.802	0.836	0.884	**0.901**
ID_0205	0.836	0.839	0.872	0.909	**0.914**

Table 2 The PSNR results compared with several other methods

PSNR	*Bicubic*	*NEDI*	*IBP*	*CIBP*	*Proposed*
ID_0201	23.37	24.33	24.75	26.34	**26.05**
ID_0202	23.87	23.96	25.28	26.87	**27.64**
ID_0203	23.13	23.39	24.32	25.90	**26.79**
ID_0204	23.60	23.86	24.83	26.41	**27.30**
ID_0205	22.95	23.44	24.33	25.91	**26.85**

Table 3 Face recognition rate in LBP, where 300 targets were selected for testing

	Bicubic	*NEDI*	*IBP*	*CIBP*	*Proposed*
Recgrate(%)	74.3	76.3	82.3	82.7	**85.3**

Fig. 3 The super-resolution results compared with several methods. It shows the LRI, Original, results of Bicubic, NEDI, IBP, CIBP and the proposed method from left to right. The size of each figure is 72x72.

5 CONCLUSIONS

In this paper, we propose an effective method to enhance the LRI.

The wavelet is introduced to preserve and enhance the edge area of the images and the acceptable pixels are extracted to avoid redundant computation and speed

up the convergence. Meanwhile, to achieve more detail information, we introduce PCA technique into our framework. The experimental results on facial images show that our method can efficiently enhance the visual effect and improve face recognition performance.

6 References

[1] J L Harris. “Diffraction and Resolving Power”. Journal of the Optical Society of America, 1964,54(7):931-936.

[2] J W Goodman. “Introduction to Fourier Optics”. New York: McGraw-Hill, 1968.

[3] R Y Tsai, T S Huang. “Multiple frame image restoration and registration, Advances in Computer Vision and Image Processing. Greenwich”, CT: JAI Press Inc. 1984: 317-339.

[4] Xin Li; Orchard, M.T., "New edge-directed interpolation," Image Processing, IEEE Transactions on , vol.10, no.10, pp.1521,1527, Oct 2001.

[5] Fei Zhou; Wenming Yang; Qingmin Liao, "Single image super-resolution using incoherent sub-dictionaries learning," Consumer Electronics, IEEE Transactions on , vol.58, no.3, pp.891,897, August 2012.

[6] Xiaoqiang Lu; Haoliang Yuan; Pingkun Yan; Yuan Yuan; Xuelong Li, "Geometry constrained sparse coding for single image super-resolution," Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on , vol., no., pp.1648,1655, 16-21 June 2012.

[7] Yu Hu; Kin-Man Lam; Guoping Qiu; Tingzhi Shen, "From Local Pixel Structure to Global Image Super-Resolution: A New Face Hallucination Framework," Image Processing, IEEE Transactions on , vol.20, no.2, pp.433,445, Feb. 2011.

[8] Wei Zhang; Wai-Kuen Cham, "Learning-based face hallucination in DCT domain," Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on , vol., no., pp.1,8, 23-28 June 2008.

[9] Nguyen, K.; Sridharan, S.; Denman, S.; Fookes, C., "Feature-domain super-resolution framework for Gabor-based face and iris recognition," Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on , vol., no., pp.2642,2649, 16-21 June 2012.

[10] Mueller, N.; Lu, Y.; Do, M.N. "Image interpolation using multiscale geometric representations," Proc. SPIE Conf. on Electronic Imaging, San Jose, USA, 2007.

[11] M. Irani and S. Peleg, “Improving resolution by image registration”, CVGIP: Graphical Models and Image Processing, vol. 53, pp. 231-239, May 1991

[12] Bareja, M.N.; Modi, C.K., "An Effective Iterative Back Projection Based Single Image Super Resolution Approach," Communication Systems and Network Technologies (CSNT), 2012 International Conference on , vol., no., pp.95,99, 11-13 May 2012.