ICIC Express Letters
Part B: Applications
Volume **7**, Number **2**, February **2016**

 ISSN 2185-2766
pp. 337–344

EDGE DIRECTED SINGLE IMAGE SUPER RESOLUTION THROUGH THE LEARNING BASED GRADIENT REGRESSION ESTIMATION

Guogang Wang, Zongliang Gan and Xiuchang Zhu

Jiangsu Provincial Key Lab of Image Processing and Image Communication
Nanjing University of Posts and Telecommunications
No. 66, Xinmofan Road, Nanjing 210003, P. R. China
kingguogang@163.com; {ganzl; zhuxc}@njupt.edu.cn

Received August 2015; accepted November 2015

Abstract. *Single image super resolution (SR) aims to estimate high resolution (HR) image from the low resolution (LR) one, and estimating accuracy of HR image gradient is very important for edge directed image SR methods. In this paper, we propose a novel edge directed image SR method by learning simple priors by mapping functions. Recognizing that the training samples of the given sub-set for regression should have similar local geometric structure based on clustering, we employ high frequency of LR image patches with removing the mean value to perform such clustering. Then we learn effective mapping functions of gradient of bicubic image patches and that of HR image patches for each cluster. Experimental results suggest that the proposed method can achieve better gradient estimation of HR image and competitive SR quality compared with other SR methods.*

Keywords: Super resolution, Edge directed, Gradient estimation, Mapping function

1. **Introduction.** Image super resolution (SR) is a fundamental and significant issue in image processing community and computer vision applications. Generally, single image SR aims to recover a high resolution (HR) image from the low resolution (LR) one [1]. The SR problem is inherently ill-posed given that many different HR images can produce the same LR image when blurred and down-sampled.

Currently, approaches solving the SR problem, can be classified into four categories, i.e., interpolation based, learning based, reconstruction based and edge directed. Interpolation based approaches estimate the high-resolution image by interpolating the unknown pixels based on the surrounding known LR pixels, such as bicubic interpolation (Bicubic) and multiscale semilocal interpolation (MSI) [2]. The underlying assumption of learning based approaches [3-5] is that there exists an inherent relationship between LR and HR image patch pairs. Then the relationship is learned and applied to a new LR image to recover its HR version. In addition, reconstruction based approaches [6-13] highlight a reconstruction constraint and back-projection (IBP) [6] is a classical reconstruction based method. IBP introduces ringing or jaggy artifacts around edges because no regularization is imposed. Due to these artifacts of the IBP method, the super resolution method based on gradient profile prior (GPP) has made many improvements with different prior and regularization term imposed [7]. In the prior of GPP, the image gradients are represented by gradient profiles, which are 1-D profiles of gradient magnitudes perpendicular to image structures. Based on the gradient profile prior, a gradient field transformation is used to constrain the gradient fields of the high resolution image. The last category is about edge directed approaches. Edge directed approaches refer to the methods based on the edge models, where effective image edge priors [7-9] are enforced as a gradient domain constraint to estimate the target HR image [7,9]. Thanks to the new algorithm, many scholars pay much

attention to improve it. Adaptive gradient magnitude self-interpolation (GMSI) method [11] and the super-resolution convolutional neural network (SRCNN) method [12] appear in succession. By using an adaptive gradient magnitude self-interpolation, the GMSI algorithm estimates the high resolution gradient, which is regarded as an edge-preserving constraint to reconstruct the high-resolution image. Different from GMSI, SRCNN is a deep learning method for single image super resolution. The SRCNN approach learns an end-to-end mapping, which is represented as a deep convolutional neural network that takes the low-resolution image as the input and outputs the high-resolution one.

Drawing a conclusion from previous work, instead of making use of gradient relationship between LR and HR image patch pairs, most edge directed methods estimate HR gradients of images according to edge pixels position or gradient magnitude for whole image. Motivated by it, we propose a novel edge directed image SR method by learning simple priors of mapping functions [13]. In particular, the gradient of HR image is estimated through its bicubic gradient and the learned coefficients. The main step of our method is about gradient estimation and reconstruction. The step about gradient regression estimation can further be divided into sample training and gradient estimation specifically. Recognizing that the training samples of the given sub-set for regression should have similar local geometric structure based on clustering, we employ high frequency of LR image patches with removing the mean value to perform such clustering. Then for each cluster, we learn the coefficients by mapping function of gradient of bicubic patches and that of corresponding HR patches. In reconstruction, the estimated gradient is regarded as a gradient constraint to guarantee that the resulted HR image preserves sharpness and refrains from artifacts such as jaggy and blurring. Experimental results suggest that the proposed method can achieve better gradient estimation of HR image and competitive SR quality compared with other SR methods.

The remainder of the paper is organized as follows. Section 2 shortly introduces the ordinary methods of the edge directed single image super resolution. Section 3 presents the proposed new edge directed SR method. Simulation results are shown in Section 4. Finally, conclusions are provided in Section 5.

2. Edge Directed Single Image Super Resolution.

In the conventional SR problem [6-8], the LR image is modeled as the Gaussian blurred and down sampled one of its HR version. Namely, given the HR image I_h, the LR one I_l is generated by

$$I_l = (I_h * G) \downarrow_s \tag{1}$$

where $*$ is a convolution operator, $\downarrow$ is a down-sampling operation, s is a scaling factor, and G is a blur kernel which is commonly approximated as a Gaussian function. The edge directed single image super resolution methods [7,11] usually model the SR problem as Equation (2):

$$I_h^* = \arg\min_{I_h} \left[E\left(I_h|I_l, \nabla I_h^T\right)\right] = \arg\min_{I_h} \left[E_i\left(I_h|I_l\right) + \beta E_g\left(\nabla I_h|\nabla I_h^T\right)\right] \tag{2}$$

where $E_i\left(I_h|I_l\right)$ is the reconstruction constraint in image domain, $E_g\left(\nabla I_h|\nabla I_h^T\right)$ is the gradient constraint in gradient domain, ∇I_h^T is the estimated HR gradient field, and parameter β is a weighting constant to balance these two constraints as a trade-off. In experiments, a larger β imposes more importance on the gradient domain, which contributes to producing sharp edges with little artifacts. Conversely, a smaller β places much importance on the image domain, resulting in better image color and contrast, yet with ringing or jaggy artifacts along edges. The reconstruction constraint measures the difference between the LR image I_l and the smoothed and down-sampled version of the HR image, i.e.,

$$E_i\left(I_h|I_l\right) = \left|(I_h * G) \downarrow_s - I_l\right|^2 \tag{3}$$

The gradient constraint requires that the gradient ∇I_h of the recovered HR image should be close to the estimated gradient ∇I_h^T as Equation (4). This paper mainly focuses on the estimation of ∇I_h^T, which will be presented in Section 3.1.

$$E_g\left(\nabla I_h | \nabla I_h^T\right) = \left|\nabla I_h - \nabla I_h^T\right|^2 \tag{4}$$

3. Proposed Edge Directed Super Resolution Method.

3.1. Sample training.

(1) Clustering of sample sets

Since natural images are abundant and easily acquired, we can assume that there are sufficient exemplar patches available for each cluster. Only meaningful image patches whose variance is greater than TH_1 are selected.

Firstly, as shown in Figure 1, for a scaling factor 3, only the pixels on position 1 in I_{bic} are from LR image directly. While other pixels on positions 2 to 9 are interpolated by the surrounding pixels on position 1. Therefore, overlapped patches in I_{bic} can be classified into 9 classes centered by the 9 kinds of pixels. From each I_h and the corresponding I_{bic}, a large set of patches $\hat{P}_c$ $(c = 1, 2, \ldots, 9)$ of I_{bic} and corresponding HR gradient patches $\partial_x I_h$ and $\partial_y I_h$ can be cropped. For a patch P_c in I_{bic}, we compute its mean value as μ and extract the feature $\hat{P}_c$ as the intensity P_c minus μ. Denote the gradient field $\nabla I_h = (\partial_x I_h, \partial_y I_h)$, and the image I_h is convolved respectively by discrete gradient operator $k_1 = [-(1/2), 0, (1/2)]$ and $k_2 = k_1^T$ to obtain $\partial_x I_h$ and $\partial_y I_h$.

Secondly, we adopt the K-means method to partition $\hat{P}_c$ into K clusters $\left\{\hat{P}_{c1}, \hat{P}_{c2}, \ldots, \hat{P}_{cK}\right\}$ and denote by C_{ck} the center of $\hat{P}_{ck}$. For given image patch, the most suitable cluster can be selected to estimate its HR gradient. Here K is usually set as integer power of 2 so that classification can be implemented with binary tree to improve efficiency.

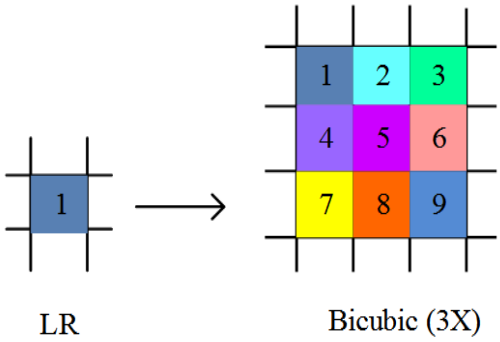

FIGURE 1. Patches classified into 9 classes by central pixel position for a scaling factor 3

(2) Regression coefficients

Supposing there are l LR exemplar patches belonging to the same cluster, let $(\nabla I_c)_i = \left[(\nabla_x I_c)_i, (\nabla_y I_c)_i\right]$ and $(\nabla I_h)_i = \left[(\nabla_x I_h)_i, (\nabla_y I_h)_i\right]$ $(i = 1, \ldots, l)$ be gradient vector of bicubic and HR patches respectively, in dimension m. Here we only calculate the gradient in central region. We learn a set of linear regression functions to individually predict the values of HR gradient. Let $\nabla_x I_c \in R^{m \times l}$ and $\nabla_x I_h \in R^{m \times l}$ be the matrix of $(\nabla_x I_c)_i$ and $(\nabla_x I_h)_i$, and $\nabla_y I_c \in R^{m \times l}$ and $\nabla_y I_h \in R^{m \times l}$ be the matrix of $(\nabla_y I_c)_i$ and $(\nabla_y I_h)_i$, and we compute the regression coefficients $C_x^*, C_y^* \in R^{m \times (m+1)}$ by

$$C_x^* = \arg\min_{C_x} \left\| \nabla_x I_h - C_x \begin{pmatrix} \nabla_x I_c \\ \mathbf{1} \end{pmatrix} \right\|; \quad C_y^* = \arg\min_{C_y} \left\| \nabla_y I_h - C_y \begin{pmatrix} \nabla_y I_c \\ \mathbf{1} \end{pmatrix} \right\| \tag{5}$$

where $\mathbf{1}$ is a $1 \times l$ vector with all values as 1. A set of functions is learned to map a bicubic gradient patch to the corresponding HR gradient patch in central region.

3.2. **Proposed gradient estimation of super resolution image.** Based on the fact that gradients of similar samples in one cluster are alike, we can estimate gradient of a patch by HR gradient of samples in the same cluster with their feature regression coefficients. Given the LR image I_l, we upsample it to obtain bicubic interpolated one I_{bic}. For patches whose variance is larger than a threshold TH_1 in I_{bic}, we perform our gradient estimation method to reduce computational complexity. And for the other patches, we use gradient of bicubic images instead. For each cropped image patch $\mathbf{p}$ with size $n \times n$ in I_{bic}, $\hat{\mathbf{p}}$ is the high-frequency component as $\mathbf{p}$ minus its mean value u. Then we find the cluster that $\hat{\mathbf{p}}$ belongs to by the standard of minimum Euclidean distance as Equation (6)

$$k^* = \arg\min_k \|\hat{\mathbf{p}} - \ \mathrm{C}_{ck}\|_2^2, \quad k = 1, \ldots, K \tag{6}$$

where C_{ck} is the cluster center of the cluster labeled by c and k. According to the cluster center, we apply the learned coefficients to compute the HR gradient of the patch by

$$\nabla_x \mathbf{p}_h = \mathrm{C}_x^* \begin{pmatrix} \nabla_x \mathbf{p} \\ 1 \end{pmatrix}; \quad \nabla_y \mathbf{p}_h = \mathrm{C}_y^* \begin{pmatrix} \nabla_y \mathbf{p} \\ 1 \end{pmatrix} \tag{7}$$

where $\nabla_x \mathbf{p}$ and $\nabla_y \mathbf{p}$ are the gradient vectors of the central region of $\mathbf{p}$, and $\nabla_x \mathbf{p}_h$ and $\nabla_y \mathbf{p}_h$ are the estimated gradients of it. We estimate gradient of each patch independently. Then to obtain the estimated image gradient, we average each pixel as it appears in the different patches.

3.3. **Reconstruction of super resolution image.** The estimated ∇I_h^T above is regarded as the gradient constraint in edge directed SR model as Equation (2). The objective energy function in Equation (2) is a quadratic function with respect to I_h; therefore, it is convex and the global minimum can be obtained by the standard gradient descent by solving the gradient flow equation. In our implementation, we use the following iterative schemes to optimize it:

$$\begin{aligned} I_h^{t+1} &= I_h^t - \tau \cdot \frac{\partial E\left(I_h | I_l, \nabla I_h^T\right)}{\partial I_h} \\ &= I_h^t - \tau \cdot \left[\left(\left(I_h * G\right) \downarrow_s - I_l\right) \uparrow_s * G - \beta \left(div\left(\nabla I_h\right) - div\left(\nabla I_h^T\right)\right)\right] \end{aligned} \tag{8}$$

where $div\left(\nabla \cdot\right) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator, $div\left(\nabla I_h\right) = \frac{\partial(\partial_x I_h)}{\partial x} + \frac{\partial(\partial_y I_h)}{\partial x}$. It can be carried out using standard finite difference scheme and τ is the step size.

4. **Experiments.** We select the Berkeley Segmentation Datasets [14] as our training set and test our method on a variety of test examples. 9 test images are shown in Figure 2. Note that, for color images, we transform them from RGB color space to YUV space. As human vision is more sensitive to luminance information, we only apply the proposed edge directed method on luminance channel (Y) and up-sample chrominance channels (UV) by bicubic interpolation. Finally, YUV are transformed into RGB as the final SR result.

FIGURE 2. 9 test images

Because of memory limitation (8GB) of the computer, we randomly collected 10^6 patches for 9 classes from the Berkeley Segmentation Datasets. The number of clusters is a trade-off between result quality and training computational complexity. With more clusters, the most suitable cluster can be selected to estimate HR gradient for the given patch, so that the estimated gradient is close to HR gradient and meanwhile high-frequency regions of the reconstructed image are better with less jaggy artifacts along edge. In our experiments, K is set as 2048.

For comparison with other methods, we use the bicubic down-sampled version as the low-resolution image. For a scaling factor 3, the patch size n is set as 11 and $\sqrt{m} = 7$. We set TH_1 as 10 in sample training. In construction, in terms of the objective indicator and visual effect, β and τ are respectively set as 0.05 and 1.9 with number of iterations set as 30.

MSE between estimated gradient and HR gradient is calculated to evaluate their similarity and the fidelity of estimated gradient. The gradient MSE results of bicubic, GPP [7], GMSI [11] and our method are listed in Table 1 and part images of estimated gradient maps are shown in Figure 3. From Table 1, our error of estimated gradient is less than that of GMSI and GPP. From Figure 3, we can see the gradient of our method is sharper and much closer to the HR gradient compared to GMSI and GPP, because we estimate patch gradient with HR gradient of similar samples in the same cluster.

TABLE 1. MSE of gradient on nine examples

Test images	Bicubic	GPP [7]	GMSI [11]	Our method
Zebra	9.952	10.327	12.050	7.232
baby	4.102	4.533	4.618	3.564
barbara	11.984	12.284	12.383	11.372
bird	5.177	5.639	5.882	4.001
foreman	8.378	8.731	8.983	7.394
head	5.066	5.428	5.275	4.737
lady	7.706	8.020	8.798	5.448
lenna	5.408	5.689	5.926	4.338
monarch	7.682	7.861	9.420	5.547
Mean value	7.273	7.612	8.148	5.959

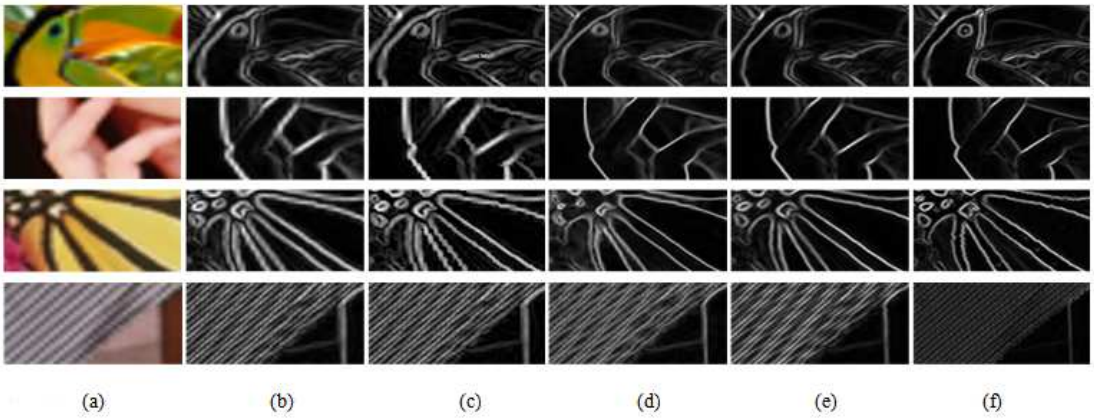

FIGURE 3. Comparisons of estimated gradient with bicubic, GPP [7] and GMSI [11]: (a) part of bicubic interpolated image; (b) the gradient field of images in (a); (c) transformed gradient of GPP [7]; (d) transformed gradient of GMSI [11]; (e) estimated gradient of our method; (f) ground truth gradient

TABLE 2. Performance in PSNR and SSIM on the 9 color images

Test images	IBP [6]		GPP [7]		GMSI [11]		SRCNN [12]		Our method	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Zebra	27.80	0.8370	27.89	0.8390	28.10	0.8458	28.88	0.8485	**29.19**	**0.8558**
baby	34.83	0.9210	34.88	0.9214	35.13	0.9239	35.04	0.9215	**35.27**	**0.9250**
barbara	26.68	0.7836	26.71	0.7843	26.71	0.7877	26.59	0.7828	**26.74**	**0.7885**
bird	33.59	0.9401	33.67	0.9412	34.33	0.9487	34.71	0.9491	**34.82**	**0.9508**
foreman	30.51	0.9141	30.72	0.9171	30.73	0.9295	30.77	0.9288	**31.13**	**0.9324**
head	33.27	0.8227	33.31	0.8231	33.47	0.8258	33.40	0.8224	**33.52**	**0.8260**
lady	30.27	0.9086	30.36	0.9102	30.86	0.9232	31.73	0.9244	**32.37**	**0.9309**
lenna	33.05	0.8975	33.12	0.8982	33.69	0.9035	**33.99**	0.9044	33.98	**0.9053**
monarch	30.35	0.9300	30.54	0.9321	30.72	0.9368	**32.43**	0.9454	32.12	**0.9459**
Mean value	31.15	0.8838	31.24	0.8852	31.53	0.8917	31.95	0.8919	**32.13**	**0.8956**

Note: IBP, our method is implemented on the same parameters: $\tau = 1.9$, $\beta = 0.05$, and iterations $= 30$. Parameter of GMSI is $\tau = 1.9$, $\beta = 0.01$ and iterations $= 30$ (Bold: best, underline: second best).

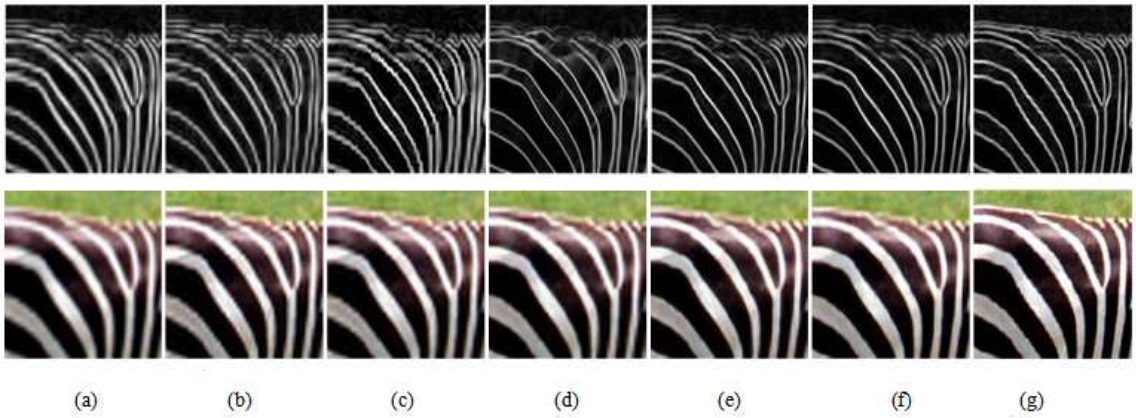

FIGURE 4. Part super resolution results of Zebra with comparison (3×) of other edge directed methods: (a) bicubic upsample (26.64dB/0.7961); (b) back-projection (27.80dB/0.8370) [6]; (c) GPP (27.89dB/0.8390) [7]; (d) GSMI (28.10dB/0.8458) [11]; (e) SRCNN (28.88dB/0.8485) [12]; (f) our method (29.19dB/0.8558); (g) ground truth. The first line is gradient domain. Among them, (a)-(f) are estimated gradient and (g) is the ground truth gradient.

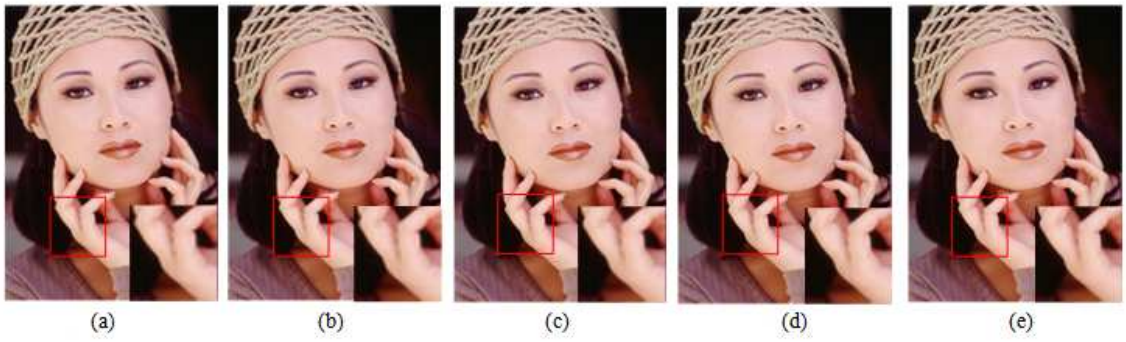

FIGURE 5. Super resolution results of Lady with comparison (3×) of other methods: (a) IBP (30.27dB/0.9086) [6]; (b) GPP (30.36dB/0.9102) [7]; (c) GMSI (30.86dB/0.9232) [11]; (d) SRCNN (31.73dB/0.9244) [12]; (e) our method (32.37dB/0.9309)

Moreover, the PSNR and SSIM [15] results are just calculated on Y channel to measure the SR results qualitatively, which are listed in Table 2. We compared our algorithm with IBP [6], GPP [7], GMSI [11], and SRCNN [12]. Figure 4 presents two comparisons

of our method with these methods in image domain and gradient domain. As shown in the figure, images are blurred by bicubic interpolation and jagged along edges by back-projection. GMSI method estimates a much sharper gradient domain, leading to edges of the reconstructed image sharper yet unnatural and artificial (refer to close-ups). The main reason is that the goal of GMSI is to obtain gradient domain which is sharper but not close to HR gradient. Figure 5 presents one example with these methods in image domain. As shown, the results of GMSI are very sharp, with rare ringing and blurring. However, unreal parts begin to appear and small scale edges are not well recovered. For example, eye area of Lady face seems to be very unreal.

In implementation, the computational load is only linear multiplication and linear addition for calculating the gradient of a patch. The total computation complexity is linearly dependent on the number of high frequency patches in the bicubic interpolated image.

5. **Conclusions.** In this paper, a novel edge directed image SR method by learning based gradient estimation has been presented. In proposed method, the gradient of HR image is estimated by using regression model of simple function. Considering the fact that the training samples of the given sub-set for regression should have similar local geometric structure based on clustering, we employed bicubic interpolated image patches with removing the mean value to perform clustering. Moreover, the simple mapping function reduced the computational complexity further. In reconstruction, the estimated gradient was regarded as the gradient constraint to guarantee that the resulted HR image preserves sharpness and refrains from artifacts. Experimental results show that estimated gradient of our proposed method is much close to the ground truth and recovered image can preserve sharper edge compared with other SR methods.

In the future, we want to speed up the proposed super-resolution algorithm, and then extend the proposed method to video super-resolution and enhancement. We are also interested in applying the nonlinear regression functions to estimate the gradient of HR image.

Acknowledgment. This work is partially supported by the National Natural Science Foundation, P. R. China (No. 61071166, 61172118, 61071091, 61471201, 61471162), Jiangsu Province Universities Natural Science Research Key Grant Project (No. 13KJA510004), Natural Science Foundation of Jiangsu Province (BK20130867), the Six Kinds Peak Talents Plan Project of Jiangsu Province (2014-DZXX-008), Jiangsu Province Postgraduate Innovative Research Plan under Grant (No. CXLX11_0409), Program of International Science and Technology Cooperation (2015DFA10940), NUPTSF (NY214039, NY213136) and the "1311" Talent Plan of NUPT.

REFERENCES

[1] S. Farsiu, D. Robinson, M. Elad and P. Milanfar, Fast and robust multiframe super resolution, *IEEE Trans. Image Process.*, vol.13, no.10, pp.1327-1344, 2004.

[2] K. Guo, X. Yang, H Zha, W. Lin and S. Yu, Multiscale semilocal interpolation with antialiasing, *IEEE Trans. Image Processing*, vol.21, no.2, pp.615-625, 2012.

[3] M. Bevilacqua, A. Roumy, C. Guillemot and M.-L. A. Morel, Low-complexity single-image super-resolution based on nonnegative neighbor embedding, *BMVC*, pp.1-10, 2012.

[4] R. Timofte, V. De Smet and L. Van Gool, Anchored neighborhood regression for fast example-based super-resolution, *ICCV*, pp.1920-1927, 2013.

[5] G. Freedman and R. Fattal, Image and video upscaling from local self-examples, *ACM Trans. Graph.*, vol.28, no.3, pp.1-10, 2010.

[6] M. Irani and S. Peleg, Motion analysis for image enhancement: Resolution, occlusion and transparency, *Journal of Visual Communication and Image Representation*, vol.4, no.4, pp.324-335, 1993.

[7] J. Sun, Z. Xu and H.-Y. Shum, Gradient profile prior and its application in image super-resolution and enhancement, *IEEE Trans. Image Process.*, vol.20, no.6, pp.1529-1542, 2011.

[8] H. Zhang, Y. Zhang, H. Li and T. S. Huang, Generative bayesian image super resolution with natural image prior, *IEEE Trans. Image Proc.*, vol.21, no.9, pp.4054-4067, 2012.
[9] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong and A. K. Katsaggelos, SoftCuts: A soft edge smoothness prior for color image super-resolution, *IEEE Trans. Image Process.*, vol.18, no.5, pp.969-981, 2009.
[10] K. I. Kim and Y. Kwon, Single-image super-resolution using sparse regression and natural image prior, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.32, no.6, pp.1127-1133, 2010.
[11] L. Wang, S. Xiang, G. Meng et al., Edge-directed single-image super resolution via adaptive gradient magnitude self-interpolation, *IEEE Trans. Circuits and Systems For Video Technology*, vol.23, no.8, pp.1289-1299, 2013.
[12] C. Dong, C. L. Chen, K. He et al., Learning a deep convolutional network for image super-resolution, *Computer Vision – ECCV*, pp.184-199, 2014.
[13] C. Y. Yang and M. H. Yang, Fast direct super-resolution by simple functions, *IEEE International Conference on Computer Vision*, pp.561-568, 2013.
[14] D. Martin, C. Fowlkes, D. Tal and J. Malik, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, *ICCV*, vol.2, pp.416-423, 2001.
[15] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image quality assessment: From error measurement to structural similarity, *IEEE Trans. Image Process.*, vol.13, no.4, pp.600-612, 2004.