Robust Single Image Super-resolution based on Gradient Enhancement

Licheng Yu, Hongteng Xu, Yi Xu and Xiaokang Yang
Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China
Shanghai Key Lab of Digital Media Processing and Communication
E-mail: magicyu@sjtu.edu.cn Tel: +86-15821934512

***Abstract*—In this paper, we propose an image super-resolution approach based on gradient enhancement. Local constraints are established to achieve enhanced gradient map, while the global sparsity constraints are imposed on the gradient field to reduce noise effects in super-resolution results. We can then formulate the image reconstruction problem as optimizing an energy function composed of the proposed sharpness and sparsity regularization terms. The solution to this super-resolution image reconstruction is finally achieved using the well-known variable-splitting and penalty techniques. In comparison with the existing methods, the experimental results highlight our proposed method in computation efficiency and robustness to noisy scenes.**

I. Introduction

Due to increasing applications in printers, digital TV, movie restoration and video surveillance, image super-resolution (S-R) techniques are extensively studied in image processing and computer vision fields. In this paper, we focus on single image SR method.

The key objective of single image super-resolution is to reconstruct a high-resolution (HR) image based on a low-resolution (LR) image. Previous works on single-image super-resolution can be roughly divided into three categories: interpolation-based, learning-based and reconstruction-based.

The interpolation-based methods like bi-linear and bi-cubic interpolation are simple and fast but tend to blur high frequency details. Some other interpolation methods were proposed to achieve performance improvements, such as edge-adaptive NEDI [1], improved NEDI [2], iterative curvature based interpolation [3, 4] and auto-regression based interpolation [5]. These interpolation-based methods can obtain upscaled images with fewer artifacts while preserving relevant image textures. The learning based methods [6, 7, 8] can recover high frequency details from a training set of HR/LR image pairs. The relationship between HR and LR patterns can be learned from input examples. Since the similarity between the training set and the test set is variable and important, it is hard to find a general training set for any LR images with arbitrary scaling factors. To reduce the dependence on the selection of training sets, self-example-based methods were proposed in [9, 10, 11, 12] whose training set was acquired by resizing the original LR images into different scales. Following a local self-similarity assumption, SR image is achieved by extracting patches from similar regions in this set. However, this kind of approach is computationally expensive in fractional interpolation and searching algorithm.

The reconstruction-based methods enforce the similarity constraint between the original LR image and the down-sampling counterpart of the HR image. Smoothness regularization is another commonly-used constraint. Recently, some other regularization terms have been advanced as prior models: Gradient profile was introduced as a kind of prior to constrain the reconstructed HR image's gradient field [13, 14], while edge smoothness prior was introduced in [15], and Jia et al. chose the logarithmic density of gradients as another prior model [16]. Compared with the other two kinds of super-resolution approaches, the advantage of reconstruction-based method is that it can be conveniently integrated with other image processing functions in many image enhancement tasks, such as de-noising [17], de-blurring [18] and contrast enhancement [15].

In this paper, we take efforts to improve current reconstruction-based super-resolution methods by establishing rational constraints on the gradient field to achieve promising HR images. Considering edge sharpness is an important factor for image quality perception, a patch-based edge enhancement model is proposed according to human vision systems property to achieve sharpened gradient field. In addition, the sparsity regularization of gradient field is advanced to suppress noise effects during the edge enhancement procedure. As a result, we can recover a high-quality HR image and reduce its noise effects simultaneously. Meanwhile, it is encouraging that the proposed SR algorithm could achieve desirable results in a few seconds.

The rest of the paper is organized as follows. Section II introduces the proposed patch-by-patch edge enhancement prior model and the sparsity constraint of HR gradient field. Section III gives the implementation details of the super-resolution scheme. Section IV gives some experimental results. Conclusions are given in Section V.

II. Gradient Enhancement Prior Model and Noise Reduction

The assumption of low-resolution imaging process can be modeled as follows:

$$L = (f * H)\downarrow^{d} + n, \tag{1}$$

where f denotes a discrete Point Spread Function (PSF) which usually is modeled as a Gaussian filter, H is the reconstructed HR image, $*$ denotes the convolution operator,

$\downarrow^d$ is a subsampling operator with factor d and n represents noise appeared in the LR image.

For designing a good super-resolution scheme, the essential issue is how to apply an effective prior or constraint on the HR image because of the ill-posedness of getting H. Currently, the gradient-based constraint has been widely used. It makes the reconstruction equation as follows,

$$E(H) = E_{data} + \lambda E(\nabla H), \tag{2}$$

where

$$E_{data} = \left\| f * H - \widehat{H} \right\|_2^2, \widehat{H} = L \uparrow^d . \tag{3}$$

Here $E(\cdot)$ is the expression of energy function, H is the reconstructed HR image which is our goal, ∇H denotes its gradient field, $\widehat{H}$ is the upsampled version of input LR image and λ controls weight. So, the term E_{data} is the reconstruction constraint in the image domain for making the smoothed and down-sampled HR image consistent with the LR image by minimizing $E(H)$, while $E(\nabla H)$ is a regularization term.

Most works focus on how to formulate $E(\nabla H)$. The gradient distribution prior [16, 18] and the gradient profile prior [13] have been proven useful. However, all of these methods focus merely on either the global feature or the local feature of image's gradient, which may not lead to the optimal results. What's more, noise effects are often ignored as an independent problem and would deteriorate HR images visual quality during image detail enhancement. To overcome these flaws, we present a fast patch-by-patch edge enhancement model, regrading $E(\nabla H)$ as a combination of a local term $E_{local}(\nabla H)$ and a global term $E_{global}(\nabla H)$, where the local term is for gradient enhancement and the global term is for suppressing noise effects.

A. Local Regularization Based On Patch-by-Patch Gradient Enhancement

Considering HR image reconstruction as an inverse problem from degradation process (1), previous works [13, 15, 16] exploited the gradient regularity constraint to make HR reconstruction problem well-posed. Although it is hard to generalize a specific equation to describe gradient field degradation process, we can roughly assume such process should make gradient field flatter and smoother, as shown in Fig.1. So the first regularization consists of a gradient-enhanced map for constraining local gradient features.

Generally, the gradient profile of an image can be modeled as Generalized-Gaussian-distribution. Based on this assumption, gradient profile prior [13] is established as the prior model of natural image gradients, which achieves image super-resolution by enhancing the sharpness of image gradient. In spite of this work, we propose a prior model on gradient patch to achieve sharper image gradient. The patch-by-patch gradient transformation is defined as follows,

$$G_{new} = \sqrt{\frac{\|G_0\|_2^2}{\|G_0^\alpha\|_2^2}} \cdot G_0^\alpha, \tag{4}$$

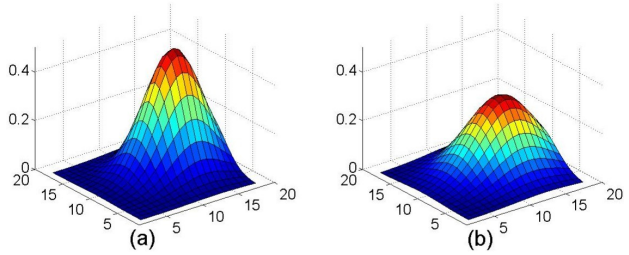


Fig. 1: Illustration of gradient degradation. (a) Gradient field of HR image. (b) Gradient field of corresponding LR image.

where G_0 represents each patch of $\nabla\widehat{H}$ and G_{new} is its transformed version; α, in our case, represents sharpness which measures the degree of enhancement; the coefficient $\sqrt{\frac{\|G_0\|_2^2}{\|G_0^\alpha\|_2^2}}$ is used to keep energy conservation during transformation procedure. It can be observed that this transformation would enhance edge's sharpness and smooth homogeneous regions.

Differing from the work in [13], we perform gradient sharpness enhancement patch by patch instead of pixel-wise operation. Compared with searching for each pixel's belonged directional gradient profile in [13], our patch-based method greatly reduces computational complexity. Another improvement of our method is that we achieve joint super-resolution and contrast enhancement by adding a multiplication coefficient β to the transformation (4), that is, the reconstructed HR image would have larger dynamic range of gradient than its LR image. The patch-by-patch gradient transformation model is further refined as,

$$G_{new} = \beta \cdot \sqrt{\frac{\|G_0\|_2^2}{\|G_0^\alpha\|_2^2}} \cdot G_0^\alpha, \tag{5}$$

where β is usually greater than 1. Such gradient enhancement is conducted on overlapped patches to prevent block artifacts in HR reconstruction results. The patch size is typically chosen as 8×8 or 16×16 pixels. Fig. 2 gives an illustration of the gradient enhancement effect.

After applying the transformation to each patch of $\nabla\widehat{H}$, we get a new estimation of the gradient of HR image, $(\nabla H)_{new}$. So, the $E_{local}(\nabla H)$ is defined as following,

$$E_{local}(\nabla H) = \|\nabla H - (\nabla H)_{new}\|_2^2. \tag{6}$$

For getting proper values of α and β, 200 HR natural images are collected. We denote the gradient patch set of these images as $\{G_{HR}\}$. Then these images are down-sampled with different factors, and we perform the transformation in (5) to their up-sampled versions with corresponding factor. We denote the gradient patch set of such versions as $\{G_{new}\}$. Using sum-of-squares, we can measure the difference between the two patch sets,

$$Error(\{G_{HR}\}, \{G_{new}\}) = \sum_i \|G_{HR_i} - G_{new_i}\|_2^2, \tag{7}$$

where i denotes corresponding patch index for $\{G_{HR}\}$ and $\{G_{new}\}$. Because β controls contrast, usually it cannot be too

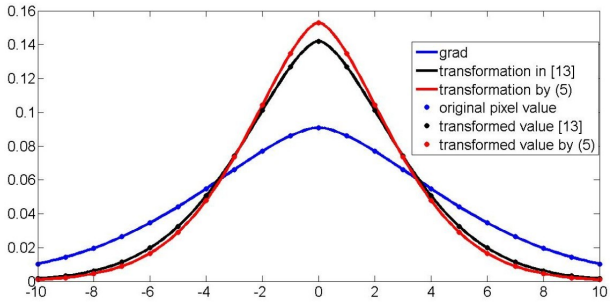


Fig. 2: Illustration of transformation in (5). The blue line represents the original 1-D gradient magnitude. Blue dots are corresponding pixels. The black line shows the transformed 1-D gradient magnitude by gradient profile transformation in [13]. The red line shows the transformation using our method.

large, in our test, we set β to 1.0∼1.3. Under this assumption, Fig. 3. illustrates that the error is minimized when α=3∼4 for down-sampling factor 4 and α=2∼2.5 for factor 2.

The whole transformation effects on natural image are shown in Fig. 4. It can be seen that our method has comparable visual result with the method in [13], both presenting sharp edges with rare ringing artifacts. However, our patch-based method does not need to conduct gradient sharpness enhancement pixel by pixel, which saves much computation time. Moreover, contrast enhancement is meanwhile achieved. As a result, considering the whole complexity, we choose patch-based gradient enhancement in our SR reconstruction scheme.

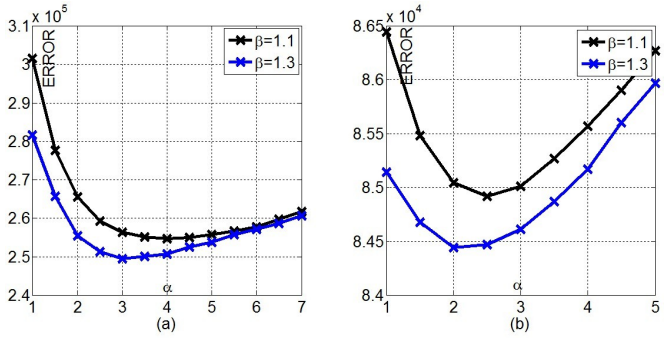


Fig. 3: Error computed for different α with β fixed. (a) 4X factor. (b) 2X factor.

B. Global Regularization based on Gradient Sparsity and Robustness to Noise

It is notable that noise magnification during super-resolution is often ignored in previous works. For noisy LR images, noises might produce local maximums in the gradient domain, in which case gradient enhancement would make them magnified as edges, especially in the smooth regions.

To deal with noisy input LR images, current works usually divide the reconstruction process into two disjoint steps: firstly

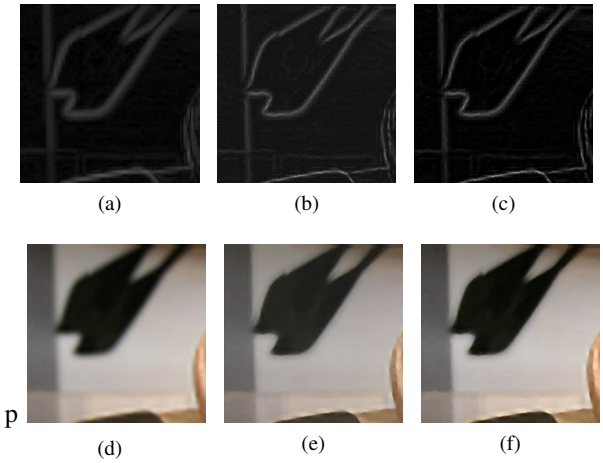


Fig. 4: Gradient field transformation on patches and its effects. (a) Gradient map of the observed LR image. (b) Transformed gradient map using [13]. (c) Our transformed gradient map. (d) The observed LR image. (e) The reconstructed image using [13]. (f) Our reconstructed image.

denoising and then super-resolution. However, any artifacts during denoising on the LR image will be kept or even magnified in the latter super-resolution process. For example, over-smoothing will destroy LR image's high-frequencies, which in turn increases the difficulties of super-resolution. Here we introduce sparsity constraint to achieve promising denoising results.

Considering gradient magnitudes' abeyance to a heavy tailed distribution, the hyper-Laplacian distribution is often used as a gradient prior in super-resolution [16], denoising [18], and de-convolution [19], etc. By applying Maximum-a-Posterior rule, the form of Laplacian distribution can be converted to L-1 norm, i.e. sparsity term [19]. Thus, we propose to employ such sparsity as another regularization in our framework for suppressing noise.

According to the analysis above, it is reasonable to rewrite reconstruction equation (2) by adding the sparsity term.

$$E(H) = E_{data} + \lambda_1 E_{local}(\nabla H) + \lambda_2 E_{global}(\nabla H), \quad (8)$$

where

$$E_{local}\left(\nabla H\right) = \left\|\nabla H - (\nabla H)_{new}\right\|_2^2, \quad (9)$$

$$E_{global}\left(\nabla H\right) = \left\|\nabla H\right\|_1. \quad (10)$$

Here λ_1 and λ_2 measure the weight of each regularization term.

As (8) shows, E_{local} constrains ∇H to be close to the transformed gradient field $(\nabla H)_{new}$ and E_{global} constrains marginal distribution of ∇H. Using the two extra constraints, the reconstructed image's gradient domain would be sharper, while noises are suppressed by the sparsity term.

III. Solution to HR image reconstruction

In section II, we enforce the constraints in both image domain and gradient domain, and then formulate the HR image

reconstruction as minimization of the energy function (8), it can be further expanded as:

$$H = argmin \left\| f * H - \widehat{H} \right\|_2^2 + \lambda_1 \cdot \left(\|\partial_x H - G_x\|_2^2 + \|\partial_y H - G_y\|_2^2 \right) + \lambda_2 \cdot \left(\|\partial_x H\|_1 + \|\partial_y H\|_1 \right). \quad (11)$$

Here G_x, G_y and $\partial_x H$, $\partial_y H$ respectively represent the directional components of $(\nabla H)_{new}$ and (∇H) along x-axis and y-axis.

Solving this problem involves the half-quadratic penalty method[16, 18, 20], so we introduce auxiliary variables $\mu = (\mu_x, \mu_y)$ to separate the optimization of (11) into two steps. Each step involves a sub-minimization problem which is a quadratic function easy to solve. The whole algorithm is in Table I. The convergence is speeded up by increasing λ_3 in each step and there are in all only $3\times n$+1 FFTs for n iterations. So the optimization in Table I can be conducted in seconds, experimental results are given in the next section. Also note, when the input LR image is noise-free, we can simply set $\lambda 2$ equal to 0.

TABLE I: Algorithm

Introduce $\mu = (\mu_x, \mu_y)$ to separate the optimization of (11) into:

$$H = argmin \left\| f * H - \widehat{H} \right\|_2^2 + \lambda_1 \left(\|\mu_x - G_x\|_2^2 + \|\mu_y - G_y\|_2^2 \right) + \lambda_2 (\|\mu_x\|_1 + \|\mu_y\|_1) + \lambda 3 (\|\mu_x - \partial_x H\|_2^2 + \|\mu_y - \partial_y H\|_2^2)$$

Iteration begins:
1) μ sub-problem, fix H and extract all the terms involving μ:

$$E(\mu) = \lambda_1 (\|\mu_x - G_x\|_2^2 + \|\mu_y - G_y\|_2^2) + \lambda_2 \left(\|\mu_x\|_1 + \|\mu_y\|_1 \right) + \lambda_3 (\|\mu_x - \partial_x H\|_2^2 + \|\mu_y - \partial_y H\|_2^2).$$

Minimize $E(\mu_x)$ and $E(\mu_y)$ on each pixel separately to get $\mu = (\mu_x, \mu_y)$.
2) H sub-problem, fix $\mu = (\mu_x, \mu_y)$ and minimize energy involving H:

$$E(H) = \left\| f * H - \widehat{H} \right\|_2^2 + \lambda_3 (\|\mu_x - \partial_x H\|_2^2 + \|\mu_y - \partial_y H\|_2^2).$$

Applying FFT to the above derivation to remove convolution operator, and then solve a quadratic minimizing problem to get Fourier transform of H, the final H is computed as $FFT^{-1}(H)$. Meanwhile, increase λ_3.
Iteration ends.

IV. Experimental results and discussion

We have described an image super-resolution scheme that is assumed to be conceptually simple, fast and noise-resistant. In this section, we prove its high-quality results by experiments.

Firstly, we compare our approach with the state-of-the-art methods to show the performance of super-resolution. In this experiment, we choose ICBI[3] for interpolation result $\widehat{H}$ in (3) and set $\lambda_1 = 1.5$,$\lambda_2 = 0$. The initial value of λ_3 is $\frac{1}{2000}$, for each iteration λ_3 is multiplied by $2\sqrt{2}$ and the iteration number n is set to be 6. The Gaussian kernel f is empirically chosen according to Table II. Also according to the statistical analysis in section II, we set $\alpha = 3$, $\beta = 1.2$ for an upscaling factor of 4 in this experiment. In Fig. 5, Bi-cubic, Glasner2009[12] and Fattal2007[15] are chosen for comparisons. From the experimental results, our proposed method is quite comparable in quality to these algorithms.

TABLE II: Empirical Kernel size and deviation for different upscaling factor

	Factor=2	Factor=4	Factor=8
Gaussian Kernel Size	5*5	11*11	15*15
Deviation	1.0	1.4	2.0

For the child face in Fig. 6, we show the comparison results with Bi-cubic, Shan08[16] and Sun11[13], our proposed method presents a better visual result. What's more, the root-mean-square (RMS) values are computed for most state-of-art methods using the ground truth of the child face. They are listed in Table III, and our method has the lowest RMS value.

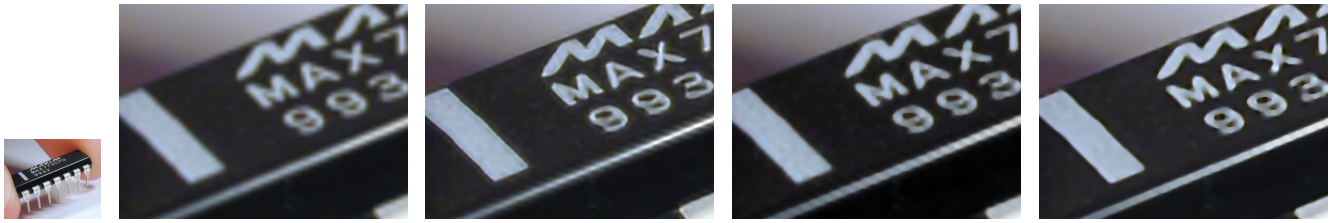
Secondly, we show the ability of noise resistance. By tuning λ_1 and λ_2 in (11), where the first weight controls how sharp the output edges would be and the second weight decides sparsity degree, we can produce a high-quality output with greatly reduced noisy effects. In Fig. 7, the input image "Lena" (128×128) is added with noise of deviation $\sigma = 0.02$. For the upsampling factor of 4, we empirically set $\lambda_1 = 2$ and $\lambda_2 = 0.15$. As shown in Fig. 7, the sparsity term avoids noise magnification during gradient enhancement step, so the result acquired using the sparsity term provides more pleasant visualization.

TABLE III: RMS values of different methods for the child face in Fig. 6.

Method	RMS	Method	RMS
Youv10 [11]	23.467	Shan08 [16]	15.392
ICBI [4]	18.397	Glasener09 [12]	20.137
Freeman02 [10]	20.146	Fattal07 [15]	20.099
NEDI [1]	17.840	Sun11 [13]	16.920
Fattal11 [9]	20.480	Our Method	**15.079**

In Table IV, we add noise with different deviation to the 128×128 "Lena" input, and compute PSNR values for different SR methods using the ground truth. The results indicate that our method is resistant to small noise. We show the visual details in Fig. 8. Some other examples for noisy input LR images are presented in Fig. 9, we can see the robustness of our method.

Thirdly, we show the speed of the proposed method. As both the patch-based gradient enhancement and the convergence of (11) can be conducted in seconds, our method is time-saving. In Table V, we list the time for upsampling the 128*128 "Lena" with factor 2 and 4 using different methods. All the experiments of NEDI[1], Sun11[13] and our methods run on 3.10GHz CPU using matlab, and the results of Shan08[16]



(a) Input (b) Bi-cubic (c) Glasner (d) Fattal (e) Ours

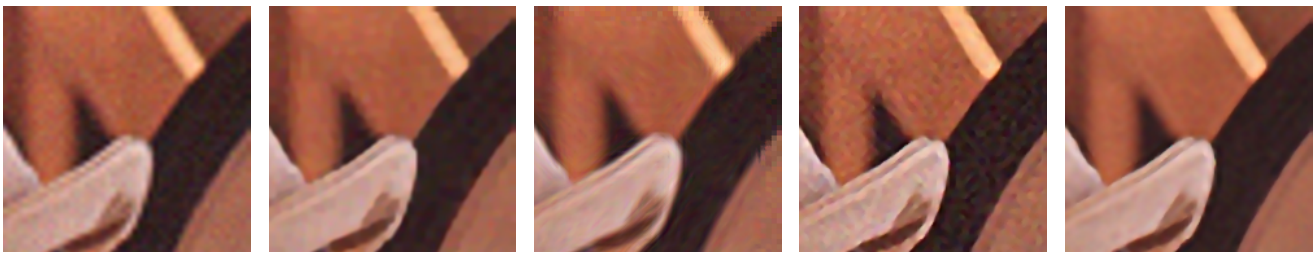
Fig. 5: Image upscaled with a factor of 4. (a) Input 128×128 LR image. (b) Bi-cubic interpolation. (c) Glasner09[12]. (d) Fattal07[15]. (e) our method.

(a) Input (b) Bi-cubic (c) Shan (d) Sun (e) Ours

Fig. 6: Image upscaled with a factor of 4. (a) Input 128×128 LR image. (b) Bi-cubic interpolation. (c) Shan08[16]. (d) Sun11[13]. (e) Our method.

(a) (b) (c) (d)

Fig. 7: Noised "Lena" (σ=0.02) upscaled with a factor of 4. (a) Our method without sparsity term. (c) Our method with sparsity term. (b) and (d) are gradient maps of (a) and (c).



(a) (b) (c) (d) (e)

Fig. 8: Close-up of noised "Lena" (σ=0.02) upscaled with a factor of 4 using different methods. (a) Bi-cubic. (b) Shan08[16]. (c) NEDI[1]. (d) Our method without sparsity constraint. (e) Our method with sparsity constraint.

TABLE IV: PSNR values for different methods in noisy scenes

	Bi-cubic (dB)	NEDI (dB)	Shan08[16] (dB)	Our SR (dB)	SR+sparsity (dB)
σ=0.01	27.53	25.07	27.18	27.27	**27.59**
σ=0.015	27.32	24.98	27.04	27.23	**27.43**
σ=0.02	27.15	24.87	26.84	26.62	**27.18**
σ=0.05	**24.80**	23.67	23.84	23.49	24.54

TABLE V: Computation time for up-scaling input image of 128×128 pixels

	NEDI	Shan08[16]	Sun11[13]	Our SR
2X(256×256)	16.8 s	3.1 s	6.67s	**1.7s**
4X(512×512)	77.42 s	10.2 s	25.65s	**6.24s**

Fig. 9: More SR results with the up-scaling factor of 4. All input images are polluted by Gaussian noises with σ=0.02.

run on their own software[1]. As we can see, our method outperforms the other three with much less time.

V. Conclusions

In this paper, a single image super-resolution based on gradient regularizations is proposed. By performing Patch-by-Patch Gradient Enhancement, we get a sharpened gradient map using a local regularization procedure. Also, considering that input images might be noisy, we introduce the sparsity constraint as a global regularization term for suppressing noise effects. Combining these two constraints, we can get a promising SR result in a few seconds. For achieving better super-resolution results, our future work may focus on how to make gradient transformation adaptive to different image contents and trying to extend such model to video applications.

[1]The software can be downloaded from http://www.cse.cuhk.edu.hk/~leojia/projects/upsampling/index.html and we choose the CPU mode for comparison.

Acknowledgment

This paper is supported in part by NSFC (60932006, 60902073), Qualcomm Cooperation Project "Video Super-resolution" and the 111 project (B07022) and STCSM (12DZ2272600).

References

[1] Xin Li, and M. T. Orchard, "New edge-directed interpolation," IEEE Trans. on Image Processing, vol. 10, pp. 1521C1527, Oct. 2001.
[2] N. Asuni and A. Giachetti. "Accuracy improvements and artifacts removal in edge based image interpolation," In Proc. 3rd Int. Conf. Computer Vision Theory and Applications (VISAPP08), 2008.
[3] Andrea Giachetti and Nicola Asuni, "Fast artifact-free image interpolation," BMVC 2008, pages 13.1-13.10.
[4] Giachetti, A. and Asuni, N, "Real time artifact-free image upscaling," IEEE TIP, vol. 20, pp. 2760-2768, Oct. 2011.
[5] Xiangjun Zhang, Xiaolin Wu, "Image Interpolation by Adaptive 2-D Autoregressive Modeling and Soft-Decision Estimation," IEEE TIP, Vol 17, pages 887-896, 2008.
[6] Jianchao Yang, Wright, J, Huang, T and Yi Ma, "image super-resolution as sparse representation of raw image patches," CVPR 2008, pages 1-8, 2008.
[7] Jianchao Yang, Wright, J, Huang, T and Yi Ma, "Image Super-Resolution Via Sparse Representation," TIP, vol. 19, pp. 2861-2783, Nov. 2010.
[8] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," CVPR 2004.
[9] Freeman and Fattal, "image and video upscaling from local self-examples," ACM Transactions on Graphics, Vol. 30, Article No. 12 , April 2011.
[10] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," IEEE Comput. Graph. Appl., 22(2):56-65, March 2002.
[11] HaCohen, Y. Fattal, Lischinski, "Image upsampling via texture hallucination," ICCP 2010, page 1, 29-30 March 2010.
[12] Daniel Glasner, Shai Bagon and Michal Irani, "Super-Resolution from a Single Image," ICCV 2009, pages 349 - 356 Oct. 2009.
[13] Jian Sun, Jian Sun and Heung-Yeung "Gradient Profile Prior and Its Applications in Image Super-Resolution and Enhancement," IEEE TIP, Vol. 20, pp 1529-1542, 2011.
[14] Jian Sun, Zongben Xu, Heung-Yeung Shum, "Image super-resolution using gradient profile prior," CVPR08, Pages 1-8, 2008
[15] Fattal. R, "Image upsampling via imposed edges statistics," ACM Transactions on Graphics (Proceedings of SIGGRAPH 2007) 26, 95, 2007.
[16] Qi Shan, Zhaorong Li, Jiaya Jia and Chi-Keung Tang, "Fast Image/Video Upsampling," ACM Transactions on Graphics (SIGGRAPH ASIA), 2008
[17] Antonio Marquina and Stanley J. Osher, "Image Super-Resolution by TV-Regularization and Bregman Iteration," Vol. 37, 2008.
[18] Dilip Krishnan and Rob Fergus, "Fast Image Deconvolution using Hyper-Laplacian Priors," NIPS(2009)1033-1041.
[19] Krishnan D, Tay T, Fergus, "Blind Deconvolution Using a Normalized Sparsity Measure," CVPR 2011, Page(s): 233 - 240.
[20] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," SIAM J. Imaging Sciences, 1(3):248-272, 2008