Single Image Super-Resolution via Blind Blurring Estimation and Dictionary Learning

Xiaole Zhao[1], Yadong Wu[1(✉)], Jinsha Tian[1], and Hongying Zhang[2]

[1] School of Computer Science and Technology,
Southwest University of Science and Technology, Mianyang, China
wyd028@163.com
[2] School of Information Engineering,
Southwest University of Science and Technology, Mianyang, China

Abstract. Learning-based methods have been becoming the mainstream of single image super resolution (SR) technologies. It is effective to generate a high-resolution image from a single low resolution input. However, the quality of training data and the computational demand are two main problems. We propose a novel process framework for single image SR tasks aiming at these two problems, which consists of blur kernel estimation (BKE) and dictionary learning. BKE is utilized for improving the quality of training samples and compatibility between training samples and test samples, which is realized by minimizing the dissimilarity between cross-scale patches iteratively (intuitively be equivalent to maximizing the similarity of cross-scale patches). A selective patch processing (SPP) strategy is adopted in BKE and sparse recovery to reduce the number of patches needed to be processed. The fact that nature images usually contain continuity and discontinuity simultaneously ensures the feasibility of SPP. The experimental results show that our method produces more precise estimation for blurring kernel and better SR effect than several state-of-the-art SR algorithms on equal conditions, but needs much less computation time.

Keywords: Super resolution · Non-Local Self-Similarity (NLSS) · Blind Blurring Estimation (BBE) · Dictionary learning

1 Introduction

Image super-resolution (SR) is a family of technologies recovering a super-resolved image from a single image or a sequence of images of the same scene. However, it is not easy to obtain an adequate number of LR observations in many practical applications. Therefore, single image super-resolution has attracted great attentions in recent years.

Machine learning based methods are promising technologies for SR problem, and it has become the most popular topic in single image SR field. Freeman et al. [1] proposed example-based learning method firstly. The algorithm predicted HR patches from LR patches by solving Markov Random Field (MRF) model via belief-propagation algorithm. Then, Sun et al. enhanced discontinuous features such as edges and corners etc.

H. Zha et al. (Eds.): CCCV 2015, Part I, CCIS 546, pp. 22–33, 2015.
DOI: 10.1007/978-3-662-48558-3_3

by primal sketch priors, which extended example-based method further [2]. To improve execution efficiency, Chang et al. [3] proposed nearest neighbor embedding (NNE) method motivated by the philosophy of locally linear embedding (LLE) [4]. They assumed LR patches and HR patches have similar space structure. The coefficients of LR patch can be solved through least square problem, which are then applied to HR patches directly. NNE utilizes a few of training data to represent a test sample and reduces the computation time dramatically. However, fixed number of NNs may cause over-fitting and/or under-fitting phenomenon [5]. Aiming at this problem, Yang et al. [6] proposed an effective method based on sparse representation, which selects the number of NNs adaptively. However, there exist two main issues about original sparse representation methods: the compatibility between training samples and test samples, and the mapping relation between LR patch space and HR patch space. More recently, Glasner et al. [7] exploited image patch non-local self-similarity within image scale and cross-scale for single image SR tasks, which makes an effective solution for the compatibility problem between training samples and test samples. As for the mapping relation, LR patch space and HR patch space are tied by some mapping function, which could be unknown and not necessarily linear [8]. The direct mapping [6] actually could not reflect this unknown relation correctly. Yang et al. [9] proposed another joint dictionary training approach to learn the duality relation between LR/HR patch spaces. The method essentially concatenated the two patch spaces and converts the problem to the standard sparse representation in a single feature space. Further, they explicitly learned the sparse coding problem across different feature spaces. The learning problem was modeled as a bi-level optimization problem, where the optimization included an L_1 norm minimization problem in its constraints. This is the so-called coupled dictionary learning (CDL) [8] algorithm.

The precision of the training data and the computational time demanding are two main challenges of current single image SR algorithms. The performance of SR methods significantly deteriorates when the blur kernel we used deviates from the true one [10]. To ensure the precision of the training data, it is necessary to estimate the true blue kernel from the input image itself. But abundant samples and iterations make it very time-consuming. In addition, solving a L_0/L_1 norm constrained optimization problem for test samples is also computationally demanding. In this paper, we present a novel single image SR method considering both SR effect and the acceleration for the algorithm. The proposed approach estimated the true blur kernel based on minimizing the dissimilarity between cross-scale patches firstly. Then, LR/HR coupled dictionaries were trained through input image downsampled by estimated blur kernel, which improved the quality of training samples and compatibility between training set and test set. A SPP strategy is employed both in blur estimation and SR recovery stage, which is based on the observation that nature images usually consist of continuity and discontinuity concurrently, and traditional interpolation methods (such as bicubic and cubic convolution interpolation) have a higher efficiency but perform on par with learning based methods in smooth area.

2 Related Work

2.1 Internal Statistics and Blur Kernel Estimation

One of the most useful internal statistical attributes of natural image patches is the patch recurrence within and cross scale, which is exploited by Glasner et al. in 2009 [7]. NLSS can be exploited as useful prior knowledge for various image restoration tasks, such as super resolution [7, 12-15], image denoising [16], deblurring [17] and inpainting [18] etc. However, it is not necessary to consider the redundancy of all patches most of the time. Maria et al. [12] quantified the property by relating it to the spatial distance from the patch and the mean gradient magnitude |grad| of the patch. Tomer Michaeli et al. [10] utilizes the non-local self-similarity to estimate the optimal blur kernel, which maximizes the cross-scale patch redundancy iteratively. For each small patch in input image, they find a few NNs in down-sampled version of input image and each NN for the small patch corresponds to a large patch in input image. All the patch pairs construct a set of linear equations which is solved using weighted least-squares to obtain an updated kernel. Fig.1 (a) shows the cross-scale blur kernel estimation process.

Actually, the method of [10] is designed to solve a convolution kernel, which makes cross-scale patch as redundant as possible. It could be interpreted as maximum a posteriori (MAP) estimation process. However, as we can see in Fig.1 (b) and Fig.1 (c), the effect of blurring in smooth area is unconspicuous. Therefore, smooth patch almost contributes nothing in blur kernel estimation with respect to structured patch.

(a) (b) (c)

Fig. 1. Descriptions of cross-scale patch redundancy and blurring effect on different patches in natural images. (a) Cross-scale patch redundancy (see context), where $\boldsymbol{k}$ is blurring kernel. (b) Structured patch and smooth patch in clean "Monarch". (c) Structured patch and smooth patch in blurred "Monarch". (c) The result of convolving (b) with a 11×11 Gaussian kernel with $\sigma = 1.5$. Obviously, Red structured patch is very different after blurring with Gaussian kernel, whereas yellow smooth patch almost do not changes visually.

2.2 Coupled Dictionary Learning

In SR tasks, HR feature space $\boldsymbol{\mathcal{F}}_H$ and LR feature space $\boldsymbol{\mathcal{F}}_L$ constitute the coupled feature space together. There exists a certain unknown relation $\mathcal{M}: \boldsymbol{\mathcal{F}}_L \to \boldsymbol{\mathcal{F}}_H$ between two feature spaces. For some specific feature $\boldsymbol{x} \in \boldsymbol{\mathcal{F}}_H$ and $\boldsymbol{y} \in \boldsymbol{\mathcal{F}}_L$, we have $\boldsymbol{x} = \mathcal{M}(\boldsymbol{y})$. The goal for coupled dictionary learning is to find a coupled dictionary

pair $\boldsymbol{D}_H$ and $\boldsymbol{D}_L$ for feature space $\mathcal{F}_H$ and $\mathcal{F}_L$ respectively, so that we can use the sparse representation of any test signal $\boldsymbol{y} \in \mathcal{F}_L$ in terms of $\boldsymbol{D}_L$ to recover the corresponding latent HR signal $\boldsymbol{x} \in \mathcal{F}_H$ in terms of $\boldsymbol{D}_H$ directly. Under the condition of L_1 norm, it can be expressed as following equations for any coupled signal pair formally:

$$\boldsymbol{\pi}_i = \arg\min_{\boldsymbol{\alpha}_i} \left\| \boldsymbol{y}_i - \boldsymbol{D}_L \boldsymbol{\alpha}_i \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha}_i \right\|_1, \quad \forall i = 1, \ldots, N_L \tag{1}$$

$$\boldsymbol{\pi}_i = \arg\min_{\boldsymbol{\alpha}_i} \left\| \boldsymbol{x}_i - \boldsymbol{D}_H \boldsymbol{\alpha}_i \right\|_2^2, \qquad \forall i = 1, \ldots, N_H \tag{2}$$

where $\boldsymbol{\pi} = \{ \boldsymbol{\pi}_i \mid i = 1, 2, \ldots, N_H \}$ is the set of sparse representation on coupled feature spaces. N_L and N_H are the numbers of atom in $\boldsymbol{D}_L$ and $\boldsymbol{D}_H$ respectively. $\boldsymbol{\alpha}_i$ is the sparse representation coefficient vector for the i th signal pair, and λ is the regularization parameter as tradeoff between the reconstruction error and the L_1 norm regularization terms. To balance the reconstruction error on observation feature space and latent feature space, CDL algorithm changes the objective function of [8]:

$$\arg\min_{\boldsymbol{D}_H, \boldsymbol{D}_L, \boldsymbol{\pi}} \sum_{i=1}^{N_H} \left(\left\| \boldsymbol{D}_H \boldsymbol{\alpha}_i - \boldsymbol{x}_i \right\|_2^2 + \left\| \boldsymbol{D}_L \boldsymbol{\alpha}_i - \boldsymbol{y}_i \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha}_i \right\|_1 \right) \quad \text{s.t. } \left\| \boldsymbol{D}_H(:, i) \right\|_2 \le 1, \ \left\| \boldsymbol{D}_L(:, i) \right\|_2 \le 1 \tag{3}$$

to the following form:

$$\arg\min_{\boldsymbol{D}_H, \boldsymbol{D}_L, \boldsymbol{\pi}} \sum_{i=1}^{N_H} \frac{1}{2} \left(\gamma \left\| \boldsymbol{D}_H \boldsymbol{\alpha}_i - \boldsymbol{x}_i \right\|_2^2 + (1-\gamma) \left\| \boldsymbol{D}_L \boldsymbol{\alpha}_i - \boldsymbol{y}_i \right\|_2^2 \right) \quad \text{s.t. } \left\| \boldsymbol{D}_H(:, i) \right\|_2 \le 1, \ \left\| \boldsymbol{D}_L(:, i) \right\|_2 \le 1 \tag{4}$$

where γ is the parameter balancing the reconstruction error terms on both feature spaces. Eqn. (4) is solved by alternatively optimizing over $\boldsymbol{D}_L$ and $\boldsymbol{D}_H$ while keeping the other fixed. Stochastic gradient descent using implicit differentiation is employed to solve $\boldsymbol{D}_L$ during the iterative process.

3 Proposed Approach

3.1 Blind Kernel Estimation with SPP

We use $\boldsymbol{Y}$ to represent input LR image, and $\boldsymbol{X}$ to be latent HR image. For each small patch $\boldsymbol{p}_i$ in the input image $\boldsymbol{Y}$, we can find a few NNs $\boldsymbol{q}_{ij}^s$ for it in down-sampled version $\boldsymbol{y}^s$. The "parent" patches $\boldsymbol{q}_{ij}$ right above $\boldsymbol{q}_{ij}^s$ are viewed as the candidate parent patches of $\boldsymbol{p}_i$. The patch pairs $\{ \boldsymbol{p}_i, \boldsymbol{q}_{ij} \}$ are used to construct a set of linear equations, which is solved by weighted least square method. According to [10], the weight of each "parent" patch is calculated by the following formula, so that good NNs contribute more than their poor counterparts:

$$w_{ij} = \frac{\exp\left(-\left\|\boldsymbol{p}_i - \boldsymbol{q}_{ij}^s\right\|^2 / \sigma^2\right)}{\sum_{j=1}^{M_i} \exp\left(-\left\|\boldsymbol{p}_i - \boldsymbol{q}_{ij}^s\right\|^2 / \sigma^2\right)} \tag{5}$$

where M_i is the number of NNs of each small patch $\boldsymbol{p}_i$ in $\boldsymbol{Y}$. σ is the standard deviation of noise added on $\boldsymbol{p}_i$, and s is scale factor. Note that we apply the same symbol to express column vector corresponding to the patch. Maximizing the cross-scale NLSS of $\boldsymbol{Y}$ with respect to scale factor s is equivalent to minimizing the dissimilarity between cross-scale patches. Therefore, we solve the following L_2 norm optimization problem for BKE:

$$\arg\min_{\boldsymbol{k}} \sum_{i=1}^{N} \left\| \boldsymbol{p}_i - \sum_{j=1}^{M_i} w_{ij} \boldsymbol{R}_{ij} \boldsymbol{k} \right\|_2^2 + \lambda \left\| \boldsymbol{C}\boldsymbol{k} \right\|_2^2 \tag{6}$$

where N is the number of small patches in $\boldsymbol{Y}$. Matrix $\boldsymbol{R}_{ij}$ corresponds to the operation of convolving with $\boldsymbol{q}_{ij}$ and down sampling by s. $\boldsymbol{C}$ is a matrix used as the penalty of non-smooth kernel. The second term of Eqn. (6) is kernel prior and λ is the balance parameter as the tradeoff between the error term and kernel prior. By setting the gradient of the objective function in Eqn. (6) to zero, we can get the update formula of $\boldsymbol{k}$:

$$\hat{\boldsymbol{k}} = \left(\sum_{i=1}^{N} \sum_{j=1}^{M_i} w_{ij} \boldsymbol{R}_{ij}^T \boldsymbol{R}_{ij} + \lambda \boldsymbol{C}^T \boldsymbol{C} \right)^{-1} \cdot \sum_{i=1}^{N} \sum_{j=1}^{M_i} w_{ij} \boldsymbol{R}_{ij}^T \boldsymbol{p}_i \tag{7}$$

This is similar to the result of [10], which can be interpreted as maximum a posteriori (MAP) estimation on $\boldsymbol{k}$. However, our blur kernel estimation has two essential differentials with respect to [10]. Firstly, our method is driven by the idea of minimizing the dissimilarity between cross-scale patches while [10] tends to maximize the similarity directly. Secondly, the number of NNs of each small patch is not fixed which provides more flexibility during solving least square problem. Therefore, the terminal criterion cannot be the totality of NNs. We use the average patch dissimilarity (APD) as terminal condition of iteration:

$$\mathrm{APD} = \left(\sum_{i=1}^{N} \sum_{j=1}^{M_i} \left\| \boldsymbol{p}_i - \boldsymbol{q}_{ij}^s \right\|_2^2 \right) \cdot \left(\sum_{i=1}^{N} M_i \right)^{-1} \tag{8}$$

To eliminate the effect on BKE caused by smooth patches, we selectively employ structured patches to calculate blur kernel (see Fig.1) according to the average gradient magnitude |grad| of a small patch. Specifically, with removing DC component, if the average gradient magnitude |grad| of each query patch is smaller than a threshold, then we abandon it. Otherwise, we use it to estimate blur kernel according to Eqn. (7). According to the research of [12], finding NNs for structured patches requires larger search region. We need to perform search in entire image here.

3.2 SR Recovery

We assume $\boldsymbol{y}_p$ is a LR patch which contains original image data, $\boldsymbol{y}$ is LR patch feature extracting from $\boldsymbol{y}_p$ .We solve L_1 norm constrained optimization problem:

$$\arg\min_{\boldsymbol{\alpha}} \left\| \boldsymbol{y} - \boldsymbol{D}_L \boldsymbol{\alpha} \right\|_2^2 + \mu \left\| \boldsymbol{\alpha} \right\|_1 \tag{9}$$

where μ allows alleviating the ill-posed problems and stabilizes the solution. $\boldsymbol{y}$ corresponds to a test LR patch extracted from enhanced interpolation version of input image. $\boldsymbol{D}_L$ is the LR dictionary trained by CDL. The solution of Eqn. (9) is found by solving a L_1-regularized Lasso problem. This can be done by Sparse Learning with Efficient Projections (SLEP) [20]. After we get the coefficient vector $\boldsymbol{\alpha}$ of LR patch feature vector $\boldsymbol{y}$, HR patch feature then can be computed directly by:

$$\boldsymbol{x} = \frac{\boldsymbol{D}_H \boldsymbol{\alpha}}{\left\| \boldsymbol{D}_H \boldsymbol{\alpha} \right\|_2} \tag{10}$$

where $\boldsymbol{D}_H$ is the HR dictionary trained by CDL. Then, Recovery HR image patch $\boldsymbol{x}_p$ can be obtained by:

$$\boldsymbol{x}_p = \left(c \times u \right) \cdot \boldsymbol{x} + v \tag{11}$$

where

$$u = \text{Mean}\left(\boldsymbol{y}_p \right), \text{ and } \quad v = \left\| \boldsymbol{y}_p - u \right\|_2$$

Here c is a constant. Generally, there are two directions to accelerate the SR process: reducing the number of patches to process and finding a fast solver for L_1 norm minimization problem Eqn. (1). A selective patch process strategy is adopted in our method as [8]. However, the criterion of selecting patches is the gradient magnitude |grad| instead of the variance of a patch for the spatial gradient magnitude of a patch is more expressive than variance when removing mean according to [12]. Thus, if the spatial gradient magnitude of a patch is smaller than a threshold, we simply apply bicubic interpolation for SR recovery. Otherwise, we employ sparse representation method to estimate HR image patch.

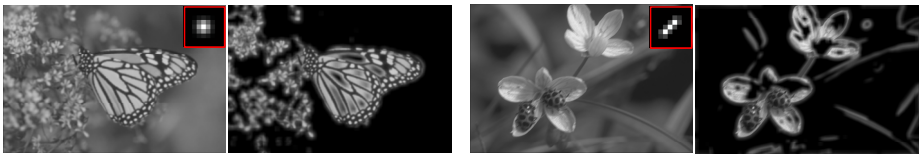
4 Experimental Results

In this section, we give some experimental results about blur kernel estimation and SR recovery. All the experiments are performed on a Philips PC with 8.0 GB memory and running a single core of Intel Xeon 2.53 GHz CPU. We mainly compare our blur kernel estimation method with [10] in terms of kernel accuracy and efficiency, while SR performance is compared with several state-of-the-art SR algorithms.

4.1 Comparisons for BKE

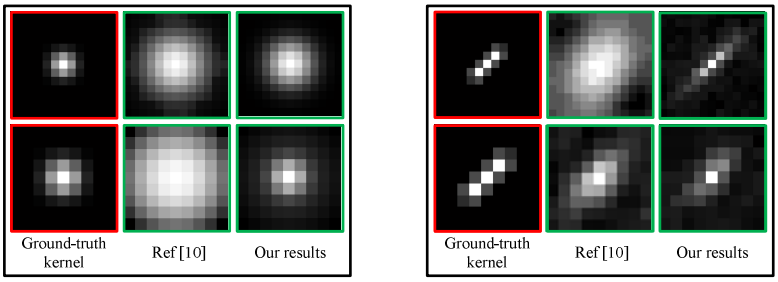
We performed $\times 2$ and $\times 3$ SR in our experiments on BKE. When scale factor $s = 2$, we set the size of small query patches $\boldsymbol{p}_i$ and candidate patches $\boldsymbol{q}_{ij}^s$ of NNs to 5×5, while the size of "parent" patches $\boldsymbol{q}_{ij}$ are set to 9×9. Both $\boldsymbol{p}_i$ and $\boldsymbol{q}_{ij}^s$ do not change size, but "parent" patches are set to be 13×13 patches when perform $\times 3$ SR. Noise standard deviation σ is assumed to be 5. Parameter λ in equation (6) is set to 0.25, and matrix $\boldsymbol{C}$ is chosen to be the derivative matrix corresponding to x and y directions of "parent" patches. The threshold of |grad| for selecting query patches is set to 10. Fig. 2 shows the structured parts used to perform BKE, and the blur kernels of "Monarch" and "Flower". It can be cleanly seen that input data are seriously blurred and inaccurate. Besides, there is only a few part of original image data that can used for BKE due to selective patch processing.

Firstly, we present the BKE results estimated from "Monarch" and "Flower" in Fig. 3 to illustrate the accuracy of the recovered kernel qualitatively. As shown, both algorithms can estimate the rough shape of the ground-truth kernel, but our approach gives more accurate results on both kernel size and shape. This is true especially when the ground-truth one is motion kernel. On the one hand, our blurring kernel estimation does not take fixed number of NNs for each query patch in LR images. This makes



(a) (b)

Fig. 2. Blurring on original images and selective patch process in BKE. (a) The first one is gray "Monarch" image blurred by a 9×9 Gaussian kernel with standard deviation $\sigma = 1.5$, the second one is structured content selected to estimate blur kernel. (b) The first image is gray "Flower" image blurred by a 9×9 "motion" kernel with len = 5 and theta = 45, the second one is same as top row.

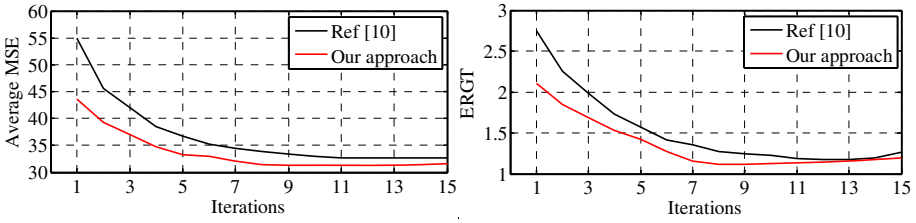


(a) Gaussian estimation (b) Motion estimation

Fig. 3. BKE Comparisons with [10]. (a) Gaussian kernel estimation with "Monarch" image showed in Fig.2. (b) Motion kernel estimation with "Flower" image showed in Fig.2. The size of kernel in first row is 13×13, the second row is 9×9.

BKE more flexible, and reduces the adverse impact caused by over-fitting or/and under-fitting. On the other hand, Selective patch processing strategy measured by the average gradient amplitude |grad| abandoned useless smooth patches and kept more structured non-smooth patches, which relieved the computational demanding for CPUs significantly.

We collect the average MSE of each NN during iterations as shown in Fig4 (a). Though we selectively use structured patches to estimate blur kernels, the average MSE of our algorithm still smaller than [10]. This is mainly attributed to that we employ unfixed number of NNs to construct the set of linear equations. As we can see in Fig. 4 (a), our BKE method converges about 8th iteration, whereas the algorithm of [10] converges at almost 15th iteration. After the convergence of algorithms, the average MSE between cross-scale patches of our algorithm is also smaller than Michaeli's algorithm. Therefore, the selective patch process and adaptive number of NNs produce more accurate BKE results, but need less iteration times.



(a) (b)

Fig. 4. BKE efficiency Comparisons with [10]. SR algorithm is original sparse representation proposed by Yang et al. [6]. (a) The attenuation of average MSE. (b) The change of ERGT as presented in [10].

We apply the same measures with [10] to compare the effect of estimated kernels on SR algorithms: The Error Ratio to Ground Truth (ERGT), which measures the ratio between the SR reconstruction error with estimated kernels and the SR reconstruction error with the ground-truth one

$$\text{ERGT}=\frac{\| \boldsymbol{X}-\overline{\boldsymbol{X}}_{\hat{k}} \|_2}{\| \boldsymbol{X}-\overline{\boldsymbol{X}}_{k} \|_2} \tag{12}$$

where $\overline{\boldsymbol{X}}_{\hat{k}}$ and $\overline{\boldsymbol{X}}_{k}$ are the recovered HR images with estimated kernel $\hat{\boldsymbol{k}}$ and ground-truth kernel $\boldsymbol{k}$ respectively. If ERGT is close to 1, it indicates that the estimated kernel is nearly as good as the ground-truth kernel. Fig. 4 (b) shows the results of applying both estimated kernels in the SR algorithm proposed by Yang et al. [6]. Here, we super resolve all the five test images with scale factor $s = 2$, and collect the average ERGT at each iteration. In the whole iterative process, the ERGT of our algorithm is always closer to 1 than [10].

Table 1 presents the effects of thresholding on SR recovery accuracy and processing time with sparse representation solved by L_1 norm minimization problem. Instead of simply using the variance of a small patch as the criterion of patch

selection, we adopted average gradient magnitude |grad| presented in [12] to filter smooth patches. As showed in table 1, with the increase of the threshold of |grad|, the consumed time decreases rapidly whereas SSIM reduces slightly. Table 2 shows the time statistics of several methods consumed in reconstruction. We use bicubic interpolation as reference and assume its time consumption is 0 for each case. Because of selective patch processing, our method provides the least amount of time consumption.

Table 3 shows PSNR and SSIM comparisons between the proposed method and several state-of-the-art SR algorithms on equal conditions. The most important equivalent condition is all input images for each algorithm are obtained by clean images blurred with default blur kernels, which is also very different from current general experimental methodologies. Our approach brings big improvement over PSNR and SSIM. Actually, this is mainly due to that our BKE process aimed directly at counteracting the blurring effect during degeneration of natural images. Moreover, the coupled dictionary learning [8] gives us accurate mapping relation between LR and HR feature spaces, which reduces reconstruction error further.

Table 1. Thresholding effects on SR recovery accuracy and processing time with sparse representation solved by L_1 norm optimization problem

Thresholds on |grad|		0	10	20	30
Lena (×2)	SSIM	0.9538	0.9527	0.9514	0.9507
	Time (s)	53.56	44.25	27.74	15.67
Monarch (×2)	SSIM	0.9428	0.9415	0.9405	0.9386
	Time (s)	63.78	41.13	22.87	11.86

Table 2. Computation comparisons between several typical methods and proposed algorithm (threshold on |grad| is 15)

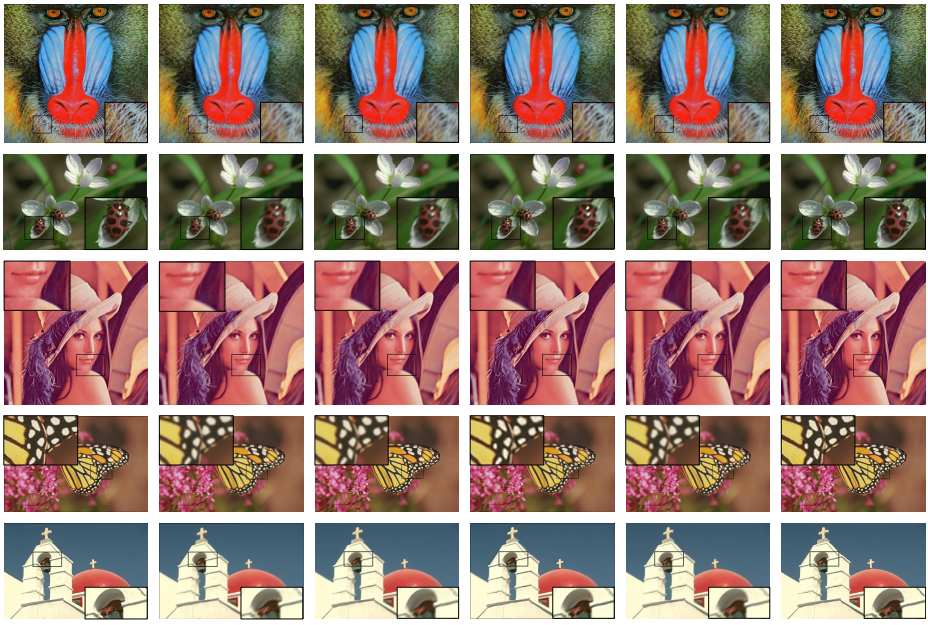
Test images	Bicubic (s)	Yang et al. [6] (s)	Zeyde et al. [19] (s)	Ours (s)
Baboon	0	86.49	56.81	36.54
Flower	0	34.60	38.62	23.03
Lena	0	79.04	53.72	32.82
Monarch	0	38.91	57.82	24.71
Tower	0	41.19	24.46	17.05

Table 3. SR comparisons between several state-of-the-art methods and proposed algorithm (threshold on |grad| is 15)

No.	scale	Yang et al. [6]		Zeyde et al. [19]		Timofte et al. [11]		Our results	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
1	×2	30.2146	0.9083	30.3194	0.9152	30.3491	0.9127	**30.5430**	**0.9335**
2	×2	32.7452	0.9361	32.7342	0.9368	32.7703	0.9376	**33.1741**	**0.9496**
3	×3	28.9768	0.8935	29.3156	0.8994	29.3036	0.8967	**29.8251**	**0.9163**
4	×3	28.3827	0.8921	28.3294	0.8954	28.3278	0.8917	**29.0495**	**0.9083**
5	×3	27.1774	0.8618	27.3329	0.8574	27.3863	0.8613	**28.6172**	**0.8853**

The number of test images: 1. Baboon; 2. Flower; 3. Lena; 4. Monarch; 5. Tower

Fig.5 shows some visual comparisons of SR reconstruction between our algorithm and several state-of-the-art SR algorithms proposed recently. For layout purpose, all images are diminished when inserted in the paper. Note that the input images of all algorithms are obtained through reference images blurring with several blur kernel. Namely, input image data is set to be of low quality via some blur kernels in our experiments for the sake of simulating many actual application scenarios. When down-sampling blurred input images with kernels before SR reconstruction, all of the algorithms adopted bicubic kernel to simulate the degradation of input images, whereas our method employed the results of BKE. Though it is well known that the three SR algorithms presented in Fig.5 are efficient for many SR tasks, they fail to reconstruct distinct details when test images are degenerated severely. This is mainly because of the great deviation between the blurring kernel used by these algorithms and the ground-truth one.



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

Fig. 5. SR reconstruction comparisons. (a). reference images. (b). blurred input with default kernels. (c). Yang et al. [6]. (d). Zeyde et al. [19]. (e). Timofte et al. [11]. (f). our results. Note that input images are obtained by blurring original clean images with different blur kernels for test purpose. (c), (d) and (e) adopted bicubic kernel to down-sample input images while our method employed estimated kernel. Ground-truth kernels: top row ($\times 2$) is 9×9 Gaussian kernel with $\sigma = 1.5$; the second row ($\times 2$) is 9×9 motion kernel with len = 5 and theta = 45; the third row ($\times 3$) is 13×13 Gaussian kernel with $\sigma = 1.5$; the fourth row ($\times 3$) is 11×11 Gaussian kernel with $\sigma = 1.5$ and the last row ($\times 3$) is 13×13 motion kernel with len = 5 and theta = 45.

5 Conclusion

We proposed a novel single image SR framework aiming at improving the SR effect and time performance in this paper. The algorithm framework mainly contains blur kernel estimation and SR recovery based on CDL. The former is realized by minimizing dissimilarity of cross-scale image patches, which is slightly similar to the MAP_k estimation approach proposed by Michaeli et al. [10]. The latter relies on a coupled dictionary learning process [8]. The selective patch processing utilized in these two stages is dependent on the criterion of average gradient amplitude. The SR effect of our method is guaranteed by improving the quality of training samples. The SPP strategy has ensured the improvement of time performance via reducing the number of query patches. All above-mentioned processes make our SR algorithm could achieve the better level of performance than several typical SR approaches. However, we can find out, from Table 1 and Table 2, that solving L_1 norm optimization problem when reconstructing is extraordinarily time-consuming. SPP should not be the single and final solution for efficiency improvement. Therefore, our further research work will be finding out other better methods for improving the SR effect and reducing the time consumption of the algorithm further.

Acknowledgment. This work is partially supported by National Natural Science Foundation of China (Grant No. 61303127), Western Light Talent Culture Project of Chinese Academy of Sciences (Grant No.13ZS0106), Project of Science and Technology Department of Sichuan Province (Grant Nos.2014SZ0223, 2015GZ0212), Key Program of Education Department of Sichuan Province (Grant Nos.11ZA130, 13ZA0169), and the innovation funds of Southwest University of Science and Technology (Grant No. 15ycx053).

References

1. Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-Based Super-Resolution. J. IEEE Computer Graphics and Applications **22**(2), 56–65 (2002)
2. Sun, J., Zheng, N.N., Tao, H., Shum, H.Y.: Image hallucination with primal sketch priors. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 729–736. IEEE Press, Monona (2003)
3. Chang, H., Yeung, D.Y., Xiong, Y.: Super-resolution through neighbor embedding. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 275–282. IEEE Press, Washington DC (2004)
4. Roweis, S., Saul, L.: Nonlinear Dimensionality Reduction by Locally Linear Embedding. J. Science **290**(5500), 2323–2326 (2000)
5. Bevilacqua, M., Roumy, A., Guillemot, C., Alberi-Morel, M.: Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In: 23rd British Machine Vision Conference (BMVC), Guildford, pp. 1–10 (2012)
6. Yang, J.C., Wright, J., Huang, T., Ma, Y.: Image super-resolution as sparse representation of raw image patches. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8. IEEE Press, Anchorage (2008)
7. Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: IEEE Conference on Computer Vision, pp. 349–356. IEEE Press, Kyoto (2009)

8. Yang, J.C., Wang, Z.W., Lin, Z., Cohen, S., Huang, T.: Coupled Dictionary Training for Image Super-Resolution. J. IEEE Transaction on Image Processing **21**(8), 3467–3478 (2012)
9. Yang, J.C., Wright, J., Huang, T.S., Ma, Y.: Image Super-Resolution via Sparse Representation. J. IEEE Transactions on Image Processing **19**(11), 2861–2873 (2010)
10. Michaeli, T., Irani, M.: Nonparametric blind super-resolution. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp. 945–952. IEEE Press, Sydney (2013)
11. Timofte, R., Smet, V.D., Gool, L.V.: Anchored neighborhood regression for fast example-based super-resolution. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp. 1920–1927. IEEE Press, Sydney (2013)
12. Zontak, M., Irani, M.: Internal statistics of a single natural image. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 977–984. IEEE Press, Providence (2011)
13. Yang, C.Y., Huang, J.B., Yang, M.H.: Exploiting self-similarities for single frame super-resolution. In: The 10th Asian Conference on Computer Vision, pp. 497–510. IEEE Press, Queenstown (2010)
14. Zoran, D., Weiss, Y.: from learning models of natural image patches to whole image restoration. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 479–486. IEEE Press, Barcelona (2011)
15. Hu, J., Luo, Y.P.: Single-Image Super-Resolution Based on Local Regression and Non-Local Self-Similarity. Journal of Electronic Imaging **23**(3), 033014 (2014)
16. Zhang, Y.Q., Liu, J.Y., Yang, S., Guo, Z.M.: Joint image denoising using self-similarity based low-rank approximations. In: Visual Communications and Image Processing (VCIP), pp. 1–6. IEEE Press, Kuhing (2013)
17. Michaeli, T., Irani, M.: Blind deblurring using Internal patch recurrence. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part III. LNCS, vol. 8691, pp. 783–798. Springer, Heidelberg (2014)
18. Guillemot, C., Meur, O.L.: Image Inpainting: Overview and Recent Advances. J. Signal Processing Magazine **31**(1), 127–144 (2014)
19. Zeyde, R., Elad, M., Protter, M.: On single image scale-up using sparse-representations. In: Boissonnat, J.-D., Chenin, P., Cohen, A., Gout, C., Lyche, T., Mazure, M.-L., Schumaker, L. (eds.) Curves and Surfaces 2011. LNCS, vol. 6920, pp. 711–730. Springer, Heidelberg (2012)
20. Liu, J., Ji, S., Ye, J.: SLEP: Sparse Learning with Efficient Projections. http://www.yelab.net/software/SLEP/