

Adaptive Blurring Estimation for Learning-Based Super Resolution

Y.W. Chen

College of Computer Science and Information Technology
Central South Univ. of Forestry and Technology
Hunan, China
Graduate School of Information Science and Engineering
Ritsumeikan University
Shiga, Japan

K. Taniguchi

Graduate School of Information Science and Engineering
Ritsumeikan University
Shiga, Japan

X.H. Han

Graduate School of Information Science and Engineering
Ritsumeikan University
Shiga, Japan

Abstract-In this paper, we address the problem of generating high-resolution (HR) image from a single low-resolution (LR) image, which is called image super-resolution (SR). Recently learning-based SR with sparse coding (SC), locality-constraint linear coding (LLC) and so on has been explored, and achieve acceptable performance. However, the conventional learning based methods cannot directly deal with a blurred LR input, which is usually considered as another research line of de-blurring, and extremely difficult to implement for real applications. This paper proposes to firstly estimate the blurring degree of an input image, and then generate the adaptive codebook (dictionary) for learning-based SR, which can simultaneously achieve the de-blurring and high-resolution image in the learning framework. We integrate our previous LLC based SR with the adaptive de-blurring procedure. Experimental results show that our proposed strategy can reconstruct the HR images more accurately than conventional methods, and its processing time is much faster.

Keywords-image super-resolution; locality-constrained linear coding; k-means clustering; de-blurring

I. INTRODUCTION

Image super-resolution (SR) referring to the problem of generating a high-resolution (HR) image from one or several low-resolution (LR) images encompasses a wide range of applications such as biometric authentication and medical imaging. The classical approach for reconstructing an HR image is to fuse several LR images of the same scene forming the multi-frame SR, which requires registration with sub-pixel accuracy [1]. In addition, generally, multi-frame SR leads to the quality degradation of the recovered HR image if the magnification factor is large or if the number of LR images is not sufficient [2].

Another research line for SR is to generate HR images from only one LR image, called single-frame SR, where the basic idea is to deduce the lost information by learning from the training samples that consist of corresponding LR and HR image local patches. Example-based SR [3] is a typical single-frame SR method, which approximates any input LR patch using only one training LR patch in prepared database, and then the unknown HR patch can be recovered by the

corresponding training HR patch, which necessarily requires a large-scale database to generate high-quality HR images. SR through neighbor embedding (NE-based SR) [4] was proposed to represent a local patch by a weighted linear combination of several training patches using locally linear embedding (LLE) [5], which can achieve acceptable results with a medium-scale database. However, the above methods easily produce noise and artifacts in the recovered HR image if there is no training patch similar to the input patch. A recent research investigated a sparse coding strategy for image SR (ScSR), which reconstructs a local patch by a sparse linear combination of atoms from an overcomplete dictionary [6]. ScSR can recover high-quality HR images by using only the learned overcomplete dictionary as the database. medical imaging, the image resolution comes limited by a number of factors such as performance of hardware, time limitations or patient's comfort. In many cases, the acquired image resolution is too low to medical diagnosis. Super resolution image reconstruction (SR) is one of the techniques to solve the problem [1]. SR is to generate a high resolution (HR) image from input low resolution (LR) image. SR can be broadly classified into two of methods.(1)the multi-frame super resolution [2,3] and (2)the single-frame super resolution [4,5]. Multi-frame SR is to reconstruct a HR image by combining multiple sub-pixel-warped LR images such as movie. Single-frame SR is to reconstruct a HR image from a LR image and low and high resolution image patches set from a database.

Our previous work (LLCSR) [7] proposed to represent each patch by a weighted linear combination of its nearer codes (also called anchor points) in the codebook. The codebook is generated by using k-means clustering. First, the HR codebook can be generated by the standard k-means clustering with the training HR database, and then the precise corresponding LR codebook can be obtained according to the membership information of the k-means clustering for the HR training patches. However, the conventional learning-based SR cannot directly deal with the blurred LR input, which is usually considered as another research line of de-blurring, and extremely difficult to implement for real applications.

Therefore, we propose a novel learning-based image super-resolution by integrating our previous LLCSR and adaptive estimation for blurring degree in the input images. Firstly we utilize a normalized sparsity measure [9] for estimating the parameter of blur kernel in an input image, and then generate adaptive codebook with similar blurring degree to the input. Experimental results show our proposed learning method can improve the quality of images for blurred input images, and the processing time is also fast.

II. LOCALITY-CONSTRAINED LINEAR CODING BASED SR

Locality-constrained Linear Coding based SR (LLCSR), proposed in our previous work [7], applies LLC [8] for SR instead of sparse coding. LLC is first proposed for classification problem and it is a fast implementation of Local Coordinate Coding (LCC) [9]. LCC shows that locality is more essential than sparsity under certain assumption for non-linear learning. Then, LLC utilizes the locality-constraint to project each sample into its local-coordinate system.

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ be a set of reshaped vector from image patches, and $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_M]$ be a codebook with M entries. Then, we can calculate the codebook for representation by following equation:

$$\begin{aligned} \min_{\mathbf{B}, \mathbf{c}} \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{B}\mathbf{c}_i\|^2 + \lambda \|\mathbf{d}_i \odot \mathbf{c}_i\|^2 \\ \text{s. t. } \mathbf{1}^\top \mathbf{c}_i = 1, \forall i \end{aligned} \quad (1)$$

Where \odot denotes the element-wise multiplication, and \mathbf{d} is the locality adaptor. Specifically,

$$\mathbf{d}_i = \exp\left(\frac{\text{dist}(\mathbf{x}_i, \mathbf{B})}{\sigma}\right) \quad (2)$$

Where $\text{dist}()$ means the distance between \mathbf{x}_i and \mathbf{B} . σ is used for adjusting the weight decay speed for locality adaptor. Usually we further normalize \mathbf{d}_i to be between (0, 1). The locality-constraint has three good properties. The first is better reconstruction for image representation, the second is local smooth sparsity because similar patches will have similar codes in codebook, and the third is that it can solve analytically and very fast.

For LLCSR[7], in training step, we use k-means clustering for codebook generation, which can guarantee the correspondence between the learned LR and HR codes (atoms) as unsolved problem in ScSR [6], [7] uses the standard k-means clustering to generate only HR codebook $\mathbf{C}_h = [\mathbf{c}_h^1, \mathbf{c}_h^2, \dots, \mathbf{c}_h^K]$ using the HR training samples $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$, following which the membership information belonging to all the clusters can also be obtained, such as the HR training samples $\mathbf{X}^i = [\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_{n_i}^i]$ in the i th cluster $\mathbf{c}_h^i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_j^i$. n_i is the number of sample in i th cluster. With the corresponding LR training samples $\mathbf{Y}^i = [\mathbf{y}_1^i, \mathbf{y}_2^i, \dots, \mathbf{y}_{n_i}^i]$, the i th code in the LR codebook can also be achieved as $\mathbf{c}_l^i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{y}_j^i$. Since the LR and HR training samples have a perfect correspondence, it is obvious that the generated LR and HR codes will also have a perfect relation.

In test step, we calculate the coefficient \mathbf{c} for each input LR patches. As [8] indicates, LLC can be approximated by

using M ($M \ll N$) nearest neighbors of \mathbf{x}_i as the local bases \mathbf{B}_i . Therefore, eq.1 can be rewrite as,

$$\begin{aligned} \min_{\mathbf{c}} \sum_{i=1}^M \|\mathbf{x}_i - \mathbf{B}\mathbf{c}_i\|^2 \\ \text{s. t. } \mathbf{1}^\top \mathbf{c}_i = 1, \forall i \end{aligned} \quad (3)$$

It is similar to LLE [5]. The difference between LLC and LLE is whether there is a codebook generation or not. Any local LR patch can be reconstructed by a weighted linear combination of its nearer codes in the LR codebook. The calculated weight coefficient significantly affects the reconstruction errors of local patches, which in turn plays a key role in assuring the quality of the recovered HR image. This study calculates the reconstruction weights as follows.

Any LR patch \mathbf{y}_L can be represented by a weighted linear combination of its M nearest neighbors in codebook \mathbf{B}_L .

$$\mathbf{y}_L = c_1 \mathbf{b}_L^1 + c_2 \mathbf{b}_L^2 + \dots + c_M \mathbf{b}_L^M = \sum_{i=1}^M c_i \mathbf{b}_L^i \quad (4)$$

The reconstruction error of eq.7 for the local patch \mathbf{y}_L is:

$$\varepsilon(\mathbf{c}) = \|\mathbf{y}_L - \sum_{i=1}^M c_i \mathbf{b}_L^i\| \quad (5)$$

To achieve the weight c_i , we minimize eq.6 with a constraint:

$$\min_{\mathbf{c}} \varepsilon(\mathbf{c}) \text{ s. t. } \sum_{i=1}^M c_i = 1 \quad (6)$$

By mathematical deduction, the reconstruction weight can be obtained:

$$\mathbf{c} = [c_1, c_2, \dots, c_M] = \frac{\mathbf{G}^{-1} \mathbf{1}}{\mathbf{1}^\top \mathbf{G}^{-1} \mathbf{1}} \quad (7)$$

Where,

$$\begin{aligned} \mathbf{G} &= (\mathbf{y}^\top \mathbf{1}^\top - \mathbf{B}_L)^\top (\mathbf{y}^\top \mathbf{1}^\top - \mathbf{B}_L) \\ \mathbf{B}_L &= [\mathbf{b}_L^1, \mathbf{b}_L^2, \dots, \mathbf{b}_L^M] \\ \mathbf{1} &: \text{a column vector} \end{aligned}$$

III. PROPOSED METHOD

In the following subsection, we propose to integrate the LLCSR strategy for adaptively dealing with the blurred LR input. First, we explain the proposed learning method, and then we introduce a sparsity measure-based blurring kernel estimation method, which is used in our proposed SR strategy.

A. Blurring-Adaptive Learning-Based SR

In order to deal with blurred LR input, we propose to adapt the blurring parameter by kernel estimation from the input image for generating adaptive LR codebook in the test step of learning-based SR. Like in LLC-based learning SR, we use $\mathbf{Y} = \mathbf{L}\mathbf{H}\mathbf{X}$ and \mathbf{X} as the training non-blurred LR and the corresponding HR images, where \mathbf{L} is down-sampling operator and \mathbf{H} is the blurring kernel corresponding to down-sample factor. The HR codebook can be generated by K-means method as the cluster centers, and at the same time, the membership information of training HR data can also be achieved. As in LLCSR, one candidate of the corresponding LR codebook can be obtained with the membership information and the non-blurred training LR images \mathbf{Y} , which only can deal with a non-blurred LR input for SR. However, since the LR inputs are not always non-blurred in real SR

application, we must deal with the blurred LR input Y_t , which cannot be corresponding to the training LR images, to generate a non-blurred HR image. One intuitive strategy is to recover a non-blurred image Y'_t , by deconvolution directly from Y_t , as the following:

$$Y'_t = \text{Deconv}(Y_t) \quad (8)$$

However, even with the known real blurring parameter, the deconvolution procedure is very difficult and usually cannot recover acceptable non-blurred image Y'_t due to uncontrolled imaging condition noise.

Therefore, we propose the blurring adaptive learning-based SR by adapting the training LR images Y to the blurred LR input Y_t , which is much easier to be implemented than deconvolution as:

$$Y' = BY \quad (9)$$

Where B is the blurring kernel estimated from a LR input Y_t . In order to recover the corresponding HR image X_t of the LR Y_t , the obtained images Y' , which has same blurring parameter with Y_t , is used as training LR image for generating LR codebook using the membership information in HR clustering.

To cluster training HR patches, and calculate the means of each cluster as HR codebook. For generating the adaptive LR codebook, we firstly estimate the blurring kernel of the LR test input to produce the corresponding training LR with same blurring degree of the test one, and then use the obtained membership information in HR cluster to yield the LR means as codebook, which doesn't require any extra processing time for clustering again. Finally, the un-observed HR image can be generated using the LLC SR with the previous learned HR and adaptive LR codebook, which can produce non-blurred HR image in spite of blurred LR input.

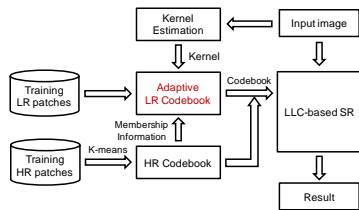


FIGURE I. THE FLOWCHART OF PROPOSED LEARNING STRATEGY.

B. Kernel Estimation

For obtaining the blurring kernel of any LR input, we use a normalized sparsity measure with a regularization term [10], which is used for blind deconvolution problem. Normalized sparsity means the measure function: the L1 norm divided by L2 norm, and can give the lower cost for sharper (non-blurred) images. Let's assume the blurred image g as following equation,

$$g = Ku + N \quad (10)$$

Where K is a blur kernel, u is a sharp (non-blurred) image, and N is gaussian noise. Then, the cost function can be defined as the following;

$$\min_{x,k} \lambda \|x \otimes k - y\| + \frac{\|x\|_1}{\|x\|_2} + \psi \|k\|_1 \quad (11)$$

y is a high frequency component of g and x is the high frequency component in the unknown sharp image. After solving eq.14, we can get blur kernel K .

IV. EXPERIMENTAL RESULTS

In this section, we compare the experimental results using our proposed method and conventional learning-based SR algorithms. In experiments, we use 5×5 patches in the interpolated LR (magnified to the same size as the HR image) and HR images, and magnification is $2 \times$. The original training LR and HR patch data sets with 100,000 random LR and HR samples are constructed. The number of codes in the codebook is set to 512, which are generated by the standard k-means clustering using the original HR training data set. The number of nearest neighbors (codes) for reconstructing a local patch is set to 5.

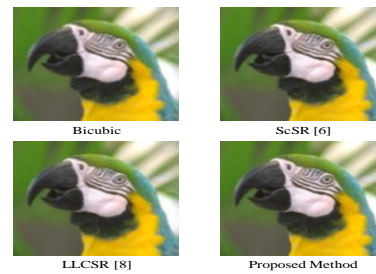


FIGURE II. EXPERIMENTAL RESULTS (PARROT).

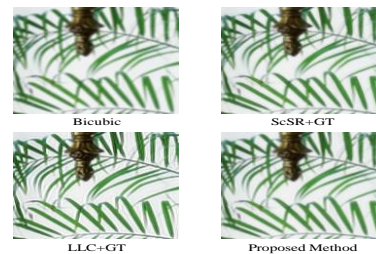


FIGURE III. EXPERIMENTAL RESULTS (LEAF).

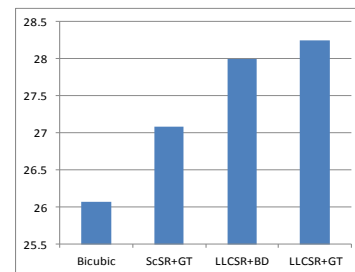


FIGURE IV. THE AVERAGED PSNR OF 7 RESULTS.

The experimental results are shown in Fig.2 with a blurred LR input. From Fig.2, it is obvious that our proposed method can achieve much clearer HR images than conventional bicubic interpolation, ScSR [8] and LLC SR. In addition, in order to validate the efficient of the proposed strategy the recovered HR images from the blurred LR input is compared with the recovered ones by the conventional Bicubic, ScSR LLC SR from the non-blurred LR in Fig.3, denoted as Bicubic, ScSR+GT, and LLC SR+GT, respectively. Furthermore, the quantitative evaluations in mean PSNR values of 7 test images

are also shown in Fig.4, which proves that our proposed strategy can achieve the HR image with much better performance than conventional learning based SR, even better than ScSR with the non-blurred LR input.

V. CONCLUSIONS

In this paper, we proposed a blurring-adaptive learning based SR. Our learning method can utilize the blur kernel \mathbf{H} , which is estimated from the LR input using a normalized sparsity measure, for generating adaptive codebook. The proposed strategy substitute de-blurring procedure of the blurred LR input, which is extremely difficult for obtaining acceptable performance, with a blurring step on the training LR data. Therefore, it can be easily to deal with the blurred LR input by integrating the blurring procedure to the learning based SR.

Finally, we apply our proposed method to restore the HR image from a blurred LR one. the proposed blurring-adaptive learning method is more robust to blur than the conventional one. Experimental results show our proposed method can get more accurate HR image than the conventional ScSR and LLCSSR, and even much better than the ScSR from non-blurred LR image. In future work, we are going to develop the real application tool for easily being used as image restoration by any user.

This work is supported in part by the MEXT Support Program for the Strategic Research Foundation at Private Universities (2013-2017).

REFERENCES

- [1] Michal Irani, Shmuel Peleg, "Improving Resolution by Image Registration", *GVGIP: Graphical Models and Image Processing*, vol.53, Issue 3, 1991.
- [2] S. Baker, T. Kanade, "Limits on Super-Resolution and how to break them", *CVPR*, vol.2, pp.372-379, 2002.
- [3] W. T. Freeman, T. R. Jones, E. C. Pasztor, "Example-based super resolution", *IEEE Computer Graphics and Applications*, vol.22, no.2, pp.56-65, 2002.
- [4] H. Chang, D. Y. Yeung, Y. Xiong, "Super-resolution through neighbor embedding", *CVPR*, vol.1, pp.275-282, 2004.
- [5] Sam T. Roweis, Lawrence K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding", *SCIENCE*, vol.290, pp.232, 2000.
- [6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no.11, pp. 2861-2873, 2010.
- [7] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, Y. Guo, "Locality-constrained linear coding for image classification". *CVPR 2010*, pp.3360-3367, 2010.
- [8] Kazuki Taniguchi, Xian-Hua Han, Yutaro Iwamoto, So Sasatani, Yen-Wei Chen, "Image Super-resolution based on Locality-constrained Linear Coding", *ICPR2012*
- [9] K. Yu, T. Zhang, Y. Gong. "Nonlinear learning using local coordinate coding", *NIPS'09*, pp.1-9, 2009.
- [10] Krishnan, D. et al. "Blind deconvolution using a normalized sparsity measure", *CVPR 2011*, pp233-240, 2011