

NEIGHBORHOOD REGRESSION FOR EDGE-PRESERVING IMAGE SUPER-RESOLUTION

Yanghao Li, Jiaying Liu*, Wenhan Yang, Zongming Guo

Institute of Computer Science and Technology, Peking University, Beijing, P.R.China, 100871

ABSTRACT

There have been many proposed works on image super-resolution via employing different priors or external databases to enhance HR results. However, most of them do not work well on the reconstruction of high-frequency details of images, which are more sensitive for human vision system. Rather than reconstructing the whole components in the image directly, we propose a novel edge-preserving super-resolution algorithm, which reconstructs low- and high-frequency components separately. In this paper, a Neighborhood Regression method is proposed to reconstruct high-frequency details on edge maps, and low-frequency part is reconstructed by the traditional bicubic method. Then, we perform an iterative combination method to obtain the estimated high resolution result, based on an energy minimization function which contains both low-frequency consistency and high-frequency adaptation. Extensive experiments evaluate the effectiveness and performance of our algorithm. It shows that our method is competitive or even better than the state-of-art methods.

Index Terms— Image Super-Resolution (SR), Edge-Preserving, Neighborhood Regression, High-frequency Details

1. INTRODUCTION

Image Super-Resolution (SR) reconstruction is currently a popular research area in signal processing. In many digital imaging applications, high-resolution images are often desired for later image processing. The SR task exactly focuses on the enhancement of image resolution. In general, given one or more low resolution (LR) images, it is responsible for mapping them to high resolution (HR) images. However, since SR image reconstruction is generally a severely ill-posed problem, many methods for SR reconstruction have been proposed these years. They can be roughly divided into three categories. Interpolation-based methods use linear or non-linear interpolation algorithms to restore a single image, such as New Directed Interpolation (NEDI) [1]; multi-frames-based methods [2, 3] aim to utilize information from a set of LR images to compose an HR image; and learning-based methods

use machine learning techniques in SR reconstruction, which are very popular in recent years.

The key part of learning-based methods is to learn the mapping between LR and HR images, and different methods are proposed to model the relationship. Following the maximum a posteriori (MAP), some proposed methods formed Markov Random Field (MRF) to connect the LR and HR images and then restored the HR images. Sparse representation based methods [4, 5] learned their own coupled LR and HR dictionaries to represent the relationship instead, based upon sparse signal representation. Moreover, assuming that two manifolds of the LR and HR image patches are locally in similar geometries, Neighbor Embedding (NE) methods [6, 7] estimated HR images by linearly combining the HR neighbors. Because NE approaches do not need to learn dictionaries or solve MRF, they dramatically reduce the execution complexity.

However, all of the above methods directly restore the whole components in the image. Since low-frequency component contains most of energy in images, restoring the whole parts together causes that the methods mainly apply on the low-frequency component. So it results in ignoring the high-frequency details of images, which are more sensitive for human vision system. In SR reconstruction, most work can do well in the low-frequency part since it is more coherent and less complex. Nevertheless, because the high-frequency part has more variations and represents details, restoration for this part is much harder and remains as a challenge of SR reconstruction. Thus, we pay more attention to the reconstruction of high-frequency details of the images in this paper.

In our work, we also consider that the SR reconstruction includes two stages: reconstruction in the low-frequency part and high-frequency details. We propose a new Neighborhood Regression method for edge-preserving Super-Resolution (NRSR), which mainly focuses on the reconstruction of high-frequency details and the combination with reconstruction of low-frequency part. Fig.1 illustrates the framework of our approach. We consider low- and high-frequency components of images separately. Then, we propose a Neighborhood Regression method for reconstruction on edge maps, which represent high-frequency details. Finally, we develop an incorporation method to combine low-frequency consistency and high-frequency adaptation to obtain the final result.

The remainder of the paper is organized as follows: we

*Corresponding author

This work was supported by National Natural Science Foundation of China under contract No. 61472011, National High-tech Technology R&D Program (863 Program) of China under Grant 2013AA013504 and Beijing Natural Science Foundation under contract No.4142021.

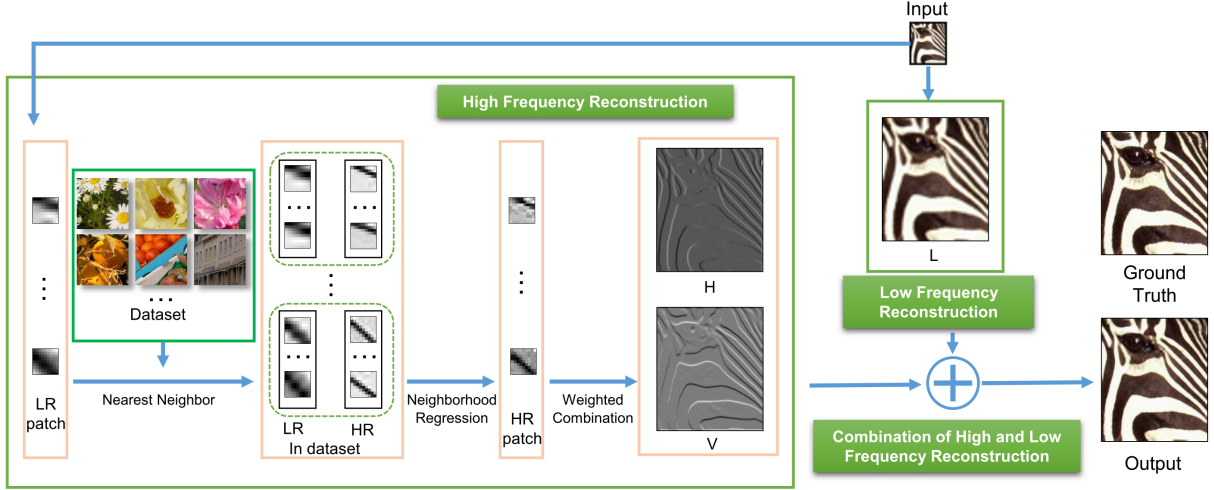


Fig. 1. An overview of the proposed method. The input LR image is reconstructed to an HR image by low- and high-frequency reconstruction parts and the final combination part. H and V are the reconstructed horizontal and vertical gradient images, and L is the reconstructed low-frequency component. Note that the patches in high-frequency reconstruction part are all in edge maps, which represent high-frequency details.

review recent approaches of neighbor embedding for super-resolution in Section 2. Then we explain our proposed method in Section 3, and show the experimental results in Section 4. Finally, the conclusions is drawn in Section 5.

2. NEIGHBOR EMBEDDING APPROACHES

Neighbor embedding approaches assume that small image patches in the low and high resolution images form manifolds with similar local geometry. Chang *et al.* [6] proposed a neighbor embedding method for image super-resolution, using manifold learning method, locally linear embedding (LLE). In LLE algorithm, local geometry is characterized by how a feature vector can be reconstructed by its neighbors in the feature space. Because the manifolds in LR and HR feature space are assumed to have similar local geometry, Chang *et al.* reconstructed HR patches as a weighted average of neighbors using the same weights as in the LR feature domain. The final result was then obtained by using reconstructed HR patches and averaging where they overlap.

The recent Anchored Neighborhood Regression (ANR) approach [7] proposed a neighbor embedding in combination with sparse learned dictionaries, which is to anchor the neighborhood embedding of a LR patch to the nearest atom in the dictionary. The ANR approach uses ridge regression to learn exemplar neighborhoods offline and precomputes corresponding embedding projection matrices to map LR patches onto the HR domain. It significantly reduces the execution time.

3. PROPOSED EDGE-PRESERVING NEIGHBORHOOD REGRESSION METHOD

In this section, we explain our NRSR method for edge-preserving image super-resolution and develop the algorithm

to obtain the final result.

3.1. Modeling Low- and High-frequency Components of Images

For single image super-resolution, the LR image Y is a blurred and downsampled version of the HR image X :

$$Y = DBX + n, \quad (1)$$

where D is the downsampling operator, B is the blurring filter, and n is the noise term. Then the low- and high-frequency components of X are denoted by $F_{low}(X)$ and $F_{high}(X)$. In the spatial domain, high-frequency components represent abrupt spatial changes in the image, such as edges that give details. Low-frequency components, on the other hand, represent global information which is in smooth variations. In our algorithm, we consider edge maps of the image to stand for the high-frequency details. Specifically, we use horizontal and vertical gradients to represent them:

$$F_{high}(X) = \{g_h(X), g_v(X)\}, \quad (2)$$

where g_h, g_v are the horizontal and vertical gradient operators. For $F_{low}(X)$, since it has smooth variation and less complexity, most simple methods perform well. We directly use bicubic result of Y as our reconstruction of $F_{low}(X)$. Thus, we use bicubic downsampling for DB and bicubic up-sampling for $B^T D^T$ to define $F_{low}(X)$:

$$F_{low}(X) = B^T D^T (DBX). \quad (3)$$

The goal of our algorithm is to reconstruct $F_{low}(X)$, $F_{high}(X)$ respectively and then combine them. Since $F_{low}(X)$ is reconstructed by bicubic method, the next two subsections focus on the reconstruction of high-frequency details and how to combine the two components.

3.2. Neighborhood Regression for Reconstruction on Edge Maps

As the previous subsection explains, $F_{high}(X)$ contains $H = g_h(X)$ and $V = g_v(X)$. We can reconstruct them respectively by the neighborhood embedding method. Least squares problems are usually more time-consuming and less efficient when regularized by the l_1 -norm of the coefficients. Thus, we formulate this problem as a least squares regression problem regularized by the l_2 -norm of the coefficients. We adopt Ridge Regression, which is known as common used in regularization of ill-posed problems, to solve it:

$$\min_{\{\alpha_h\}} \|y_h - N_h \alpha_h\|_2^2 + \lambda \|\alpha_h\|_2^2, \quad (4)$$

where y_h corresponds to the horizontal gradient feature of the given edge map patch of the LR image Y , and N_h corresponds to the neighborhood of y_h in the LR space, which is obtained by nearest neighbors searching in the dataset. The parameter λ is the regularization term coefficient. The coefficient α_h can be given by an explicit solution of Ridge Regression:

$$\alpha_h = (N_h^T N_h + \lambda I)^{-1} N_h^T y_h. \quad (5)$$

The corresponding HR's horizontal gradient patch then can be calculated by the same coefficient α_h :

$$h = N_H \alpha_h, \quad (6)$$

where h is the HR horizontal edge map patch, and N_H is the neighborhood in the HR space corresponding to N_h .

Finally, we need to use these computed patches to construct H . For the overlap portions of patches, most methods average them among different patches. However, in our edge images, since a point value represents variation around it, different confidences inside a patch should be considered. As a central point of a patch, we believe that it is constructed better than a point in the boundary of the patch. Thus, we use different weights inside a patch, and when combining different patches, the overlap parts can be weighted averaged. We adopt Gaussian function as the weighting coefficients.

With this regression method, we can also reconstruct V , another part of $F_{high}(X)$, in the same way.

3.3. Combination of Low-frequency Consistency and High-frequency Adaptation

This subsection illustrates how to combine the reconstructions of high-frequency details (H, V) and low-frequency part (L), which can be obtained by the neighbor regression method as shown in Sec. 3.2 and bicubic method respectively. Considering these constraints, low-frequency consistency and high-frequency adaptation, we propose to minimize the energy function below to calculate the desired HR image X :

$$\min_{\{X\}} E_{high}(F_{high}(X), \{H, V\}) + E_{low}(F_{low}(X), L), \quad (7)$$

where E_{high} and E_{low} are the high and low-frequency terms to represent the data description errors. As the definitions in previous subsections, Eq.(7) is reformulated as follow:

$$\min_{\{X\}} \|g_h(X) - H\|_2^2 + \|g_v(X) - V\|_2^2 + \lambda_L \|B^T D^T (DBX) - L\|_2^2, \quad (8)$$

where λ_L is a parameter used to balance high and low description error terms.

As many works have shown that the nonlocal redundancies existing in natural images are very useful for image restoration [8] and incorporating nonlocal method can enhance the performance of image super-resolution [5], [9], [10]. We also integrate the nonlocal similarity to our algorithm framework. For each local patch x_i , we search for its similar patches x_i^l , and then expect the prediction error $\|x_i - \sum_{l=1}^L b_i^l x_i^l\|_2^2$ to be small. The nonlocal weight b_i^l is defined in [8]. Thus, Eq.(8) can be now reformulated by:

$$\min_{\{X\}} \|g_h(X) - H\|_2^2 + \|g_v(X) - V\|_2^2 + \lambda_L \|B^T D^T (DBX) - L\|_2^2 + \eta \sum_{x_i \in X} \|x_i - \sum_{l=1}^L b_i^l x_i^l\|_2^2, \quad (9)$$

where η is a constant balancing the contribution of nonlocal regularization.

To tackle the optimal minimization problem in Eq.(9), we design an iterative algorithm to optimize X . For the first two terms in Eq.(9), we use gradient descent method to update it. Then we adopt nonlocal mean method to deal with the fourth term. Finally, we use back projection [4] to handle the third term. These updates should be processed alternatively.

3.4. Summary of the Algorithm

The proposed NRSR method has been explained, and now the entire process is summarized in Algorithm 1.

Algorithm 1 NRSR

Input: Dataset for high-resolution and low-resolution images, a low-resolution image Y .

• **For** each patch y_h in edge maps of Y :

- use the neighborhood regression method described in Sec. 3.2 to reconstruct horizontal and vertical gradient patches.
- use weighted average to reconstruct H and V .

• Use bicubic upsampling method to reconstruct the low-frequency component L .

• Use $X = L$ for the initial value.

• **For** each iteration **Until** convergence:

- update X by gradient descent method for the first two terms in Eq.(9),
- update X by nonlocal mean method for the fourth term in Eq.(9),
- update X by back projection for the third term in Eq.(9).

Output: high-resolution image X

Table 1. PSNR(dB) results on image super-resolution (scaling factor = 3)

Images	Bicubic	ScSR[4]	SCDL[5]	ANR[7]	BPJDL[11]	DPSR[10]	Proposed	Gain vs. DPSR
<i>Cameraman</i>	25.20	25.84	26.19	25.54	26.18	26.22	26.47	0.25
<i>Child</i>	29.35	30.36	30.74	29.81	30.94	30.95	31.10	0.15
<i>Lena</i>	30.11	31.08	31.59	30.35	31.54	31.75	31.97	0.22
<i>Oldman</i>	29.84	30.25	30.41	29.28	30.45	30.67	30.71	0.04
<i>Zebra</i>	23.64	25.68	27.41	25.98	27.35	27.79	28.27	0.48
<i>Average</i>	27.56	28.64	29.27	28.19	29.29	29.48	29.70	0.22

4. EXPERIMENTAL RESULTS

In this section, we evaluate our method via the reconstruction precision and visual quality. For a fair comparison, our training dataset adopts the same dataset in [10], which consists of 28 logo images and 34 natural images. In our experiments, the patch size is 9×9 . We choose 9 nearest neighbors for Neighborhood Regression, and the regularization parameter λ in Eq.(4) is set to be 0.15. The representative state-of-the-art image super-resolution methods, including Bicubic, ScSR [4], SCDL [5], ANR [7], BPJDL [11] and DPSR [10], are employed to compare with the proposed NRSR method.

Table 1 compares the final results on the five testing images, while some examples are shown in Fig.2. For color images, we only calculate PSNR measures for the luminance channel. From Table 1, we can see that our proposed method outperforms the state-of-arts in the five testing images, and its PSNR is in average 0.22dB higher than DPSR[10], which is the second best result among the competing methods. In particular, from Fig.2 we can see that our method can preserve edges better than the state-of-art methods in visual quality. The edges are less blurred and sharper in our results, while reconstructions in other methods also have more ringing effects and artifacts. These experiments demonstrate that our method performs better than the state-of-art methods.

5. CONCLUSION

In this paper, we proposed a new edge-preserving image super-resolution method using Neighborhood Regression. By considering low- and high-frequency components separately, our method reconstructs them respectively, and finally uses energy minimization to obtain the result. In contrast to the state-of-art methods, our method preserves edges better, which are sharper and more natural. It achieves better results in both reconstruction precision and visual quality.

6. REFERENCES

- [1] Xin Li and Michael T. Orchard, "New edge-directed interpolation," *IEEE Transactions on Image Processing*, 2001.
- [2] Sina Farsiu, M. Dirk Robinson, Michael Elad, and Peyman Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, 2004.
- [3] Xuelong Li, Yanting Hu, Xinbo Gao, Dacheng Tao, and Beijia Ning, "A multi-frame image super-resolution method," *Signal Processing*, 2010.

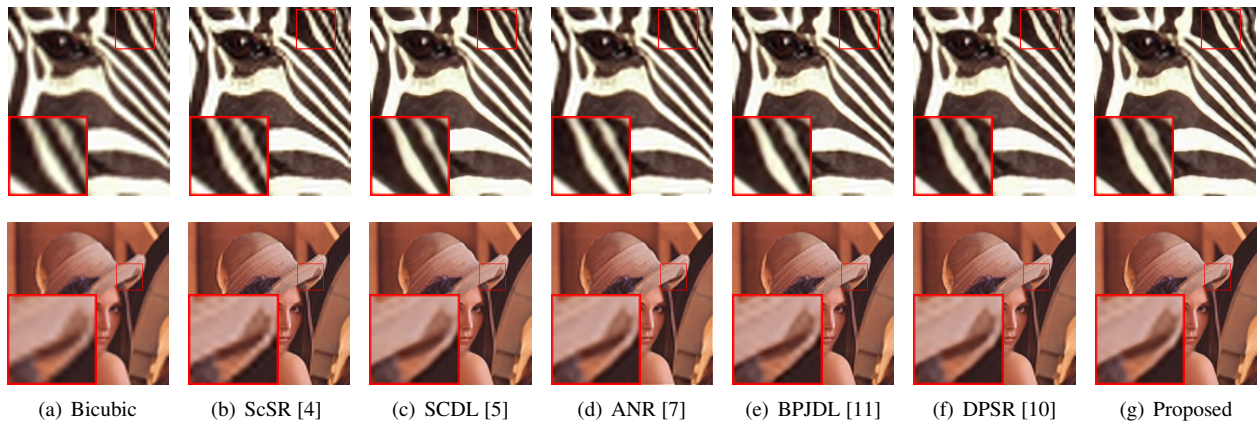


Fig. 2. Super-resolution results by $3 \times$ on the (*Zebra*, *Lena*) images. The red block with its corresponding magnification on the left-bottom corner of each image shows the reconstruction details.

- [4] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, 2010.
- [5] Shenlong Wang, Lei Zhang, Yan Liang, and Quan Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, 2012.
- [6] Hong Chang, Dit-Yan Yeung, and Yimin Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, 2004.
- [7] Radu Timofte, Vincent De Smet, and Luc Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int'l Conf. Computer Vision*, 2013.
- [8] Antoni Buades, Bartomeu Coll, and J-M Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, 2005.
- [9] Weisheng Dong, Lei Zhang, Guangming Shi, and Xiaolin Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, 2011.
- [10] Yu Zhu, Yanning Zhang, and Alan L. Yuille, "Single image super-resolution using deformable patches," in *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, 2014.
- [11] Li He, Hairong Qi, and Russell Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, 2013.