Signal Processing: *Image Communication* 43 (2016) 68–81

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Single image super resolution using local smoothness and nonlocal self-similarity priors

Honggang Chen, Xiaohai He*, Qizhi Teng, Chao Ren

College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China

A R T I C L E I N F O

Article history:
Received 30 June 2015
Received in revised form
14 January 2016
Accepted 14 January 2016
Available online 22 January 2016

Keywords:
Single image super resolution
Reconstruction-based
Local smoothness
Nonlocal self-similarity
Split Bregman Iteration

A B S T R A C T

Single image super resolution (SISR) is an inverse problem, so an effective image prior is necessary to reconstruct a high resolution (HR) image from a single low resolution (LR) image. On the one hand, natural images satisfy the property of local smoothness; on the other hand, the patches could find some similar patches in different locations within the same image, and this property is known as nonlocal self-similarity. In this paper, we propose a SISR method by incorporating the local smoothness and nonlocal self-similarity priors in the reconstruction-based SISR framework simultaneously, and the Split Bregman Iteration (SBI) optimization algorithm is imitated to solve the L1-regularized problem. Experimental results show that, in most case, the proposed method quantitatively and qualitatively outperforms the state-of-the-art SISR algorithms.

1. Introduction

SISR aims to estimate a HR image from a LR image. Obviously, the SISR problem is inherently ill-posed as the identical LR images can be generated from different HR images. The process of degradation can be generally modeled as

$$Y = \mathbf{DH}X + n \tag{1}$$

where $X \in R^{s^2MN \times 1}$ is the lexicographically stacked representation of HR image and $Y \in R^{MN \times 1}$ is LR image in lexicographic order; the matrices $\mathbf{H} \in R^{(s^2MN)\times(s^2MN)}$ and $\mathbf{D} \in R^{(MN)\times(s^2MN)}$ represent blurring operator and down-sampling operator respectively; s stands for down-sampling factor; and n denotes noise, which is assumed as additive white Gaussian noise usually. According to the degradation model in Eq. (1), the SISR problem can be described as follows: finding the optimal estimation $\hat{X}$ of HR image X under the condition of knowing a single LR image Y. In recent years, a plenty of SISR algorithms have been proposed, and these methods can be generally grouped into four categories: interpolation-based SISR [1–7], example-based SISR [8–12], sparse representation-based SISR [13–19] and reconstruction-based SISR [20–27].

1.1. Interpolation-based SISR

Interpolation methods estimate unknown HR pixels utilizing the LR pixels around them without regarding to the blurring phase generally. Nearest, bilinear, and bicubic are conventional interpolation schemes, which are high efficiency, but the interpolated results often suffer from artifacts such as jaggies, ringing and blurring. The main reason for this problem is that each pixel is treated the same rather than considering the difference of local structures. To over come this drawback, researchers proposed many new interpolation algorithms with considering the varying structure of pixels. Li [1] proposed a new edge directed interpolation (NEDI) method using the covariance of HR image estimated from the covariance of LR image. A soft-decision interpolation (SAI) technique was presented in [2], in which the 2-D piecewise autoregressive model is used to

* Corresponding author.
E-mail address: hxh@scu.edu.cn (X. He).

http://dx.doi.org/10.1016/j.image.2016.01.007
0923-5965/

learn local structure. In addition, instead of estimating HR image pixel-by-pixel, a patch of pixels are estimated simultaneously in SAI. Based on the ideas of NEDI and nonlocal means (NLM) filter, Zhang [3] proposed a nonlocal edge directed interpolation (NLEDI) method with considering the difference of geometry structures of the samples in a local window, and the samples are assigned different weights according to the structure similarity with the pixel to be interpolated. For the above interpolation algorithms, the shape of local analysis window is fixed usually, such as a square window centered at unknown pixel, but the statistics of the samples in a shape fixed window can not adapt to local structure accurately. Based on this observation, an adaptive directional window selection strategy was proposed in [4] to choose the analysis window on the basis of local structure. Wei et al. [5] designed a contrast-guided interpolation method by considering contrast, which provides information about edge strength. Combining the sparse representation model and nonlocal self-similarity, two interpolation methods were developed in [6] and [7], and the two methods both have excellent performance. The main idea of [6] is that the sparse representation of each patch should be close to a linear combination of its similar patches sparse representations. The interpolation algorithm proposed in [7] is composed of two phases: firstly, the regions fitting nonlocal self-similarity assumption are recovered firstly, and then the recovered result is refined using nonlocal sparsity model. Although good results can be achieved by these interpolation methods, the application of this type of SISR methods is limited as the blurring phase is neglected for most of interpolation algorithms.

1.2. Example-based SISR

Example-based SISR approaches require a training database generally, which is composed of LR and HR pairs and used to provide extra information in SR phase. Inspired by locally linear embedding, Chang [8] presented a SR method named neighbor embedding (NE). For each patch of input LR image, finding the K-nearest LR samples in a prepared database firstly, and the corresponding HR samples in the database are used to estimate HR patches. Based on the property that patches in a natural image would appear repeatedly both within the same scale and across different scales, Glasner et al. constructed a unified SISR framework assembling the multi-image SR and example-based SR [9]. The basic idea of the unified framework is that the similar patches within the same scale can be seen as multi-observations generated from the same HR patches, while the similar patches across different scales can be viewed as neighbors. Timofte [10] proposed a fast example-based SR method called Anchored Neighbor Regression (ANR), which maps the LR patches onto HR domain using the projections learned from neighborhoods. Combining the SR methods of learning from self-examples and learning from a prepared database, a fast SR method was introduced by Yang [11]. In [11], the in-place examples are used to train a function to map the LR patches to HR patches, and the problem of lacking self-examples can be overcome by utilizing an external database. In order to minimize the overall reconstruction error of all training examples, a collection of local regressors are optimized jointly in [12], and each test LR patch is reconstructed by the most appropriate regressor. The drawback of the example-based methods is that the performance depends heavily on the correlation between test images and training database. Besides, the completeness of database has significant impact on SR performance.

1.3. Sparse representation-based SISR

In recent years, the sparsity of image has been used in many image processing fields, such as denoising, deblurring and SR. In [13], sparse representation was introduced to SR work firstly by Yang based on the fact that the representation of HR image with respect to a proper dictionary can be recovered from the corresponding LR image. The dictionaries were constructed by using the randomly sampled patches of LR and HR training images in Yang's algorithm. This work was improved by themselves two years later with an advanced dictionary training strategy, where the LR and HR dictionaries are trained jointly to obtain more compact dictionaries [14]. The K-SVD method was used in dictionary training stage and the sparse coding was performed by using the OMP algorithm in Zeyde's SR framework [15], which is developed from Yang's method, but significantly reduces the computational complexity. Different from the previous sparse representation based SR methods, Dong et al. [16] proposed a multi-dictionaries based SR algorithm. In the dictionary training phase, a series of compact sub-dictionaries are trained based on the classified training samples; in the reconstruction phase, each patch is assigned a sub-dictionary adaptively to obtain more accurate sparse representation coefficients and reduce the computation of sparse coding phase. Furthermore, two regularization terms are incorporated into sparse representation model in Dong's work. Zhu et al. presented a self-learning based SR method [17], in which the patches selected from the upsampled LR image are used to train a dictionary. The SR method proposed in [17] is efficient and practical, nevertheless, compared with [14,15], the performance of this approach degrades slightly as the insufficiency of training patches. In [18], the low resolution features were transformed into high resolution subspaces using the multiple linear mapping (MLM) learned from classified training patches, thus the high resolution image can be obtained effectively and efficiently. Moreover, a nonlocal means based enhancement method was adopted to reduce the artifacts of SR results. Dong [19] proposed a deep convolutional neural network (CNN) based SISR method and showed that the traditional sparse-coding-based algorithm can also be seen as a kind of deep convolutional network. The end-to-end mapping between LR images and HR images was optimized in Dong's SR method, which achieves excellent reconstruction performance. Similar to the example-based methods, most of sparse representation-based methods need an external training database, thus the SR performance is related to database. Besides, the training phase should be consistent with the observation model of test images, otherwise, some unwanted artifacts would be produced in SR results.