Signal Processing 123 (2016) 53–63

Contents lists available at ScienceDirect

Signal Processing

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

Single image super-resolution using regularization of non-local steering kernel regression ☆

Kaibing Zhang [a,b], Xinbo Gao [b,*], Jie Li [b], Hongxing Xia [a]

[a] School of Computer and Information Science, Hubei Engineering University, Xiaogan 432000, China
[b] School of Electronic Engineering, Xidian University, Xi'an 710071, China

ARTICLE INFO

Article history:
Received 17 July 2015
Received in revised form
15 October 2015
Accepted 28 November 2015
Available online 28 December 2015

Keywords:
Image super-resolution
Non-local similarity
Steering kernel regression
Regularization
Gradient decent

ABSTRACT

One promising technique for single image super-resolution (SR) is reconstruction-based framework, where the key issue is to apply reasonable prior knowledge to well pose the solution to upsampled images. In this paper, we employ the non-local steering kernel regression (NLSKR) model to devise an effective regularization term for solving single image SR problem. The proposed regularization term is based on the complementary properties of local structural regularity and non-local self-similarity existing in natural images, aiming at preserving sharp edges and producing fine details in the resultant image. By integrating the regularization term into the standard back-projection framework, we solve a least squares minimization problem to seek the desired high-resolution (HR) image. Extensive experimental results on several public databases indicate that the proposed method produces promising results in terms of both objective and subjective quality assessments.

1. Introduction

The objective of single image super-resolution (SISR) is to produce a high-resolution (HR) image with more details using only one low-resolution (LR) observation [1]. The SISR technique has attracted much attention in recent two decades due to its wide range of practical applications such as computer vision, medical and remote sensing imaging, video surveillance, and mobile multimedia devices. Up to now, a large number of methods have been proposed to yield a higher quality image which cannot be directly captured from imaging sensors. Generally speaking, existing SISR methods can be divided into three categories: interpolation-based methods, reconstruction-based methods, and example learning-based methods.

Interpolation-based methods [2,3] usually apply a base function or an interpolation kernel to predict the unknown pixels in the HR grids, so it is probably an efficient way to yield an HR image for real time applications. However, because of the discontinuities in natural images, these methods tend to generate blurring details and ringing artifacts in the result, so the SR capability of this kind cannot meet the application requirement in practice.

Reconstruction-based methods achieve an SR estimate by imposing a certain prior knowledge on the reconstructed image. These approaches also require that the reconstructed HR image should be consistent with the LR input image via back-projection. A variety of edge-directed priors (e.g., edge prior [4,5], gradient profile prior [6], and total variation prior [7]) have been proposed to preserve edges. Another popular prior is self-similarity redundancies within the input image that can be found in [8]. In addition, many

☆ Fully documented templates are available in the elsarticle package on CTAN.

* Corresponding author.
E-mail address: xbgao@mail.xidian.edu.cn (X. Gao).

http://dx.doi.org/10.1016/j.sigpro.2015.11.025
0165-1684/

Bayesian prior models, including Markov random fields [9] and the Gibbs distribution with a Huber potential function [10], have also been shown to be particularly effective for the reconstruction-based SR approaches. Nevertheless, these previous priors are beneficial only to suppress artifacts and preserve edges, but are clumsy at adding novel details lost in the LR image.

The third is example-based SR methods [11–20], which presume that the high frequency details missing in an LR image can be predicted from a database containing a large number of LR and HR image pairs. According to how the input and the output are formulated and how the mapping relationship is established, the representative learning techniques include *k*-nearest neighbor (*k*-NN) learning [11,12], manifold learning [13–16], sparse coding [17–19], and regression methods [20]. The *k*-NN learning methods are simple and intuitional and perform detail synthesis by similarity search among an enormous database, so it is time-consuming to perform SR reconstruction. Manifold learning-based methods are good at generalization in the case of smaller training database. Nonetheless, with a fixed number of nearest neighbors, they are prone to produce blurring results due to over- or under-fitting. Sparse representation methods use sparse prior to adaptively choose the most related neighbors to synthesize the details, so they enable to achieve better results than the previous example-based methods. However, one of the most important issues of them is the difficulty to determine an optimal dictionary for a generic natural image. Kim et al. [20] utilized kernel ridge regression to learn the correspondence between the LR and HR images for SR. However, this method cannot effectively control blurring effects along dominant edges. The example-based SR approaches exceed the reconstruction-based methods in that they are capable of generating new details which cannot be found in the LR input. However, the SR quality remains largely affected by the quality of training examples or applied dictionary.

In this paper, we propose to develop a novel SISR algorithm based on the reconstruction-based model. We utilize an effective regression model in [21], i.e., the non-local steering kernel regression (NLSKR), to design a unified regularization term for solving reconstruction-based SISR problem. The major characteristics of newly proposed method are three-fold.

1. We jointly use two complementary characteristics of the local structural regularity and non-local self-similarity in natural images to develop a unified regularization term for ill-posed SR problem. The local structural regularity is helpful to preserve sharp edges. While non-local self-similarity is beneficial to yield a reliable result.
2. We frame the proposed regularization term under the standard reconstruction-based SR framework and solve a local optimal solution to the desired HR image using gradient descent rule.
3. Our approach that employs the regularization term constructed from the input LR image itself can yield state-of-the-art SR performance on a large public database.

The rest of this paper is organized as follows. Section 2 briefly reviews related works for SISR in the literature. Section 3 describes our NLSKR regularization based SISR method. Experimental results are presented in Section 4. Finally, Section 5 concludes the paper.

2. Related work

It is widely acknowledged that the performance of reconstruction-based methods largely depends upon the imposed prior knowledge. To produce an edge sharpened result, edge-directed prior is a very popular method applied in the reconstruction-based methods. For example, Fattal proposed to use edge-statistics [5] to conduct image upsampling. Sun et al. [6] utilized a gradient profile prior for SR reconstruction. A soft edge smoothness prior applied in the alpha-channel of salient edges was found in [4]. Morse and Schwartzwald presented a level-set based method for image magnification [7]. The advantage of these previous methods is good at producing sharp edges. The weakness, however, is insufficient to restore plausible details in the textural regions.

By contrast, example learning-based approaches predict the missing high frequency details in LR images by learning the mapping relationship from a set of LR images and the corresponding HR ones. The seminal work of this class of methods was first proposed by Freeman et al. [11,12], which uses the Markov random model to build the mapping relationship between the related pixels at two different scales. Chang et al. [13] introduced locally linear embedding (LLE) [22] from manifold learning, assuming that two manifolds of the LR image patches and the corresponding HR ones are in similar local geometries. Although the above methods can generate many novel details into an input LR image, one of the most challenging problems is their intensive computational cost. Particularly, Yang et al. [17] proposed sparse coding (SC) to adaptively choose the most related neighbors to synthesize the HR patches based upon a jointly learnt LR-HR dictionary pair, leading to promising results with fine details. Under a similar framework with Yang's method, Zeyde et al. [23] improved the SR efficiency using PCA-based dimensionality reduction and Orthogonal Matching Pursuit (OMP) [24] for sparse coding. More example-based methods can be seen in [18,19]. Another class is called regression-based method [20], which directly estimates the HR pixels via a set of learnt mapping functions. Example learning-based methods have shown an effective way to hallucinate missing details but difficult to suppress noticeable artifacts along edges.

Self-similarity is particularly useful for various image restoration tasks. One of the most representative work that thoroughly exploits the self-similar repetitive patterns within the same scale as well as across different scales for SISR was proposed by Glasner et al. [18]. This particular method combines both multi-frame reconstruction- and example learning-based methods by utilizing repetitive patterns found in the LR input. Nevertheless, the success of this method heavily relied on whether there exist sufficient repetitive patterns across different scales. Dong et al.