J. Vis. Commun. Image R. 35 (2016) 91–102

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Single image super-resolution via internal gradient similarity ☆

Yang Xian, Yingli Tian *

The Graduate Center and the City College of New York, City University of New York, New York, NY 10016, USA

ARTICLE INFO

Article history:
Received 11 June 2015
Accepted 30 November 2015
Available online 17 December 2015

Keywords:
Image super-resolution
Image quality enhancement
Patch similarity
Self-redundancy
Across scale
Gradient similarity
Optimization
Gradient descent algorithm

ABSTRACT

Image super-resolution aims to reconstruct a high-resolution image from one or multiple low-resolution images which is an essential operation in a variety of applications. Due to the inherent ambiguity for super-resolution, it is a challenging task to reconstruct clear, artifacts-free edges while still preserving rich and natural textures. In this paper, we propose a novel, straightforward, and effective single image super-resolution method based on internal across-scale gradient similarity. The low-resolution gradients are first upsampled and then fed into an optimization framework to construct the final high-resolution output. The proposed approach is able to synthesize natural high-frequency texture details and maintain clean edges even under large scaling factors. Experimental results demonstrate that out method outperforms existing single image super-resolution techniques. We further evaluate the super-resolution performance when both internal statistics and external statistics are adopted. It is demonstrated that generally, internal statistics are sufficient for single image super-resolution.

1. Introduction

Image super-resolution (SR) is to predict a fine-resolution image from one or multiple coarse-resolution image(s). It is a fundamental operation in image editing software and plays a crucial role in a variety of applications such as video surveillance, desktop publishing, movie restoration, and object tracking in satellite images.

Image SR techniques can be applied to regular 2D images and depth images [5,18,20,21,25]. Broadly speaking, within the scope of 2D images, SR tasks can be divided into two categories: multi-image SR and single-image SR. Multi-image SR methods [1,3,8,10,17,24,27] utilize the non-redundant information of multiple frames of the same scene to reconstruct one fine-resolution image. Single image SR, on the other hand, only has one low-resolution input image at disposal which leads to a numerically ill-posed problem. It is possible to generate numerous high-resolution images given the same low-resolution image. Therefore, single image SR task relies on strong image priors or assumptions to eliminate the ambiguities and to finalize a visually pleasing output image among all the possible candidate solutions.

Single image SR approaches can be categorized into three classes: interpolation-based, example-based and reconstruction-based methods. Interpolation-based methods are based on data invariant linear filtering. Commonly used interpolation-based methods such as bilinear and bicubic interpolations are simple and efficient to obtain upsampled images and thus are widely used in the commercial software. These linear interpolation-based SR methods hinge on the inaccurate assumption that natural images are always smooth. However, with discontinuities existing in natural images, the generated results are over-smoothed and with obvious visual artifacts such as blurring, aliasing and jaggies. More sophisticated interpolation-based methods were proposed [22,28] to suppress artifacts and restore sharper edges compared with interpolation with simple linear filters.

Example-based image upscaling methods were proposed by Freeman et al. [12,13] to learn the relationship between high-resolution images and their corresponding low-resolution versions through an external image dataset. In general, with lack of relevance between testing images and a universal training dataset, the produced results are noisy with irregularities along curved edges. The performance may be improved with an increment in the size of the external image dataset. However, it would lead to heavy computational cost as well as the increasing ambiguity among patch correspondences [39]. Various methods have been proposed later after Freeman to improve the SR performance and the computational speed. Coupled high-resolution and low-resolution dictionaries [15,31,35,36,38,39] are popular representations for the external patch exemplars where optimization techniques could apply. Yang et al. [35,36] learnt a compact

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author at: Department of Electrical Engineering, The City College of New York, Convent Avenue at 140th Street, New York, NY 10031, USA. Fax: +1 (212) 650 8249.

E-mail addresses: yxian@gradcenter.cuny.edu (Y. Xian), ytian@ccny.cuny.edu (Y. Tian).

http://dx.doi.org/10.1016/j.jvcir.2015.11.015
1047-3203/

dictionary based on sparse signal representation which allows the possibility to adaptively choose the most relevant reconstruction neighbors. Built upon Yang's work, Zeyde et al. [38] introduced several modifications to further improve the execution speed. Timofte et al. [31] temped to combine the benefits of both neighbor embedding [2,4] and sparse coding. Zhu et al. [39] proposed a method based on deformable patches which lead to a more "expressive" dictionary without increasing the size of the dictionary. Deep network has recently been explored to structure external example-based learning [6,7].

With all the attempts made in external example-based image SR methods, the inadequate relevance between certain testing images and the universal training dataset remains unsolved. Self-example-based image SR provides a solution to build a tailored 'training dataset' for each input low-resolution image. Self-example-based image SR methods [11,14,37] have been proposed based on the observation that for small image patches in a natural image, self-similarities exist within the image itself and across different resolutions. Glasner et al. [14] proposed a patch searching scheme based on a patch pool formed with internal patches collected through a pyramid structure with only the input image at different resolutions. Freedman et al. [11] utilized a real time multi-step coarse-to-fine algorithm which adopts a local search instead of a global search. Yang et al. [37] combined learning from self-examples as well as an external dataset into a regression model based on in-place examples. These patch-based approaches are capable of generating natural-looking textures estimated from across-scale self-similarities with/without the assistance of external statistics under small magnification factors. However, it is difficult to handle the visual artifacts introduced during the estimation process especially when the upscaling factor is relatively large due to the lack of a robust global constraint.

Reconstruction-based image SR methods tend to form global constraints to enforce the fidelity between the predicted high-resolution image and the provided low-resolution image. In [26], Shan et al. built a feedback-control framework that enforces the output image to be consistent with the input image when down-scaling to the input resolution. Recently, reconstructed gradient profile has been a popular prior utilized during the restoration of the target HR image [9,29] due to its heavy-tailed distribution [16]. Fattal [9] proposed an image SR method which generates the gradient field of the target HR image other than determining its pixel intensities directly. Reconstruction of the high-resolution gradient field is based on a statistical edge dependency relating certain edge features of two different resolutions. Sun et al. [29] modeled the image gradient by a parametric profile model. A gradient field transformation was learnt to constrain the gradients of the HR image given the low-resolution gradients. These reconstruction-based approaches are often referred to as "edge-directed SR" [30] and are capable of creating sharp edges even under large magnification factors. However, they tend to produce "unnatural" or "unrealistic" patterns within detailed texture regions due to their emphasis on preserving sharp edges. Moreover, it is extremely difficult to capture the complicated local features of natural images with a limited number of parameters.

In this paper, we propose a novel and accurate single image SR method based on internal across-scale gradient similarity. Given an input low-resolution image, its gradients in horizontal and vertical directions are first calculated and upscaled through the proposed internal gradient similarity-based upsampling algorithm. The target high-resolution image is then reconstructed based on the two upsampled gradients and the input low-resolution image. Our proposed approach combines the advantages of both self-similarity-based methods and the gradient-based techniques.

Internal across-scale gradient similarity is based on the observation that small patches in a natural image tend to recur redundantly across different resolutions. Therefore, we should expect the similar redundancy for gradient patches. In this paper, we refer to the usage of image patch redundancy as "image similarity" and the gradient patch redundancy as "gradient similarity".

As shown in Fig. 1, our proposed method successfully generates visually pleasing result in both edges and textures. From the zoom-in comparisons between the input low-resolution image (the input image is magnified to the target resolution by nearest-neighbor interpolation for better illustration) and the generated high-resolution result, it is clearly demonstrated that our approach is capable of restoring clear and natural face and eyes contours while still preserving realistic knit textures. As later shown in Fig. 7, it is a challenging task to restore the textures within the hat region for most of the recent state-of-the-art SR algorithms due to the complicated patterns. However, our generated result rebuilds this region naturally with minimal artifacts closest to the ground-truth.

Compared with the image similarity-based SR approaches, our proposed method has the following advantages:

- The proposed upscaling scheme is robust with stable SR performance by ensuring both local and global fidelity between the input low-resolution and the output high-resolution correspondences in both gradient level and image level.
- An easily-optimized energy function is utilized to combine constraints from different domains into one uniformed framework.
- Gradients of natural images have been modeled by a heavy-tailed distribution [16]. Computational cost can be reduced by upscaling the patches with small variance directly through bicubic interpolation.

The rest of the paper is organized as follows: Section 2 provides a detailed description of the proposed image SR method. Section 3 demonstrates the feasibility of reconstructing a high-quality image from its gradients and the advantages of the proposed gradient similarity-based SR algorithm. Section 4 discusses why only the internal statistics is adopted in our approach and the role that external statistics could play in image SR tasks. Experimental results and discussions are presented in Section 5. The conclusions are drawn in Section 6.

2. Internal gradient similarity-based super-resolution

In this section, we introduce the proposed image SR method based on internal across-scale gradient similarity. Fig. 2 illustrates the schematic pipeline of our approach. Since human eyes are more sensitive to brightness changes than color changes, therefore, same as the majority of other SR approaches [2,4,6,7,9,11,14,15,26,29,31,33,35–37,39], for a given input low-resolution color image, our proposed SR algorithm is only performed in the luminance channel of the YUV color space while the other two channels are upsampled through bicubic interpolation.

Given a grayscale low-resolution image L, to upsample L by a magnification factor s, we first calculate the gradients L_x and L_y of L in horizontal and vertical directions. After that, based upon internal across-scale gradient similarity, L_x and L_y are upsampled individually by s to obtain the gradients in high-resolution as H_x and H_y; the second step is to reconstruct the target high-resolution image H from L, H_x and H_y through optimizing a uniformed energy function which incorporates the constraints in both image and gradient levels. Details of these two steps will be presented in the following subsections respectively.

2.1. Upscale low-resolution gradients

In order to upscale the low-resolution gradient in x direction by factor s, L_x is firstly decomposed into a set of overlapping patches at