Information Sciences 367–368 (2016) 337–353

Contents lists available at ScienceDirect

Information Sciences

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

A phase congruency based patch evaluator for complexity reduction in multi-dictionary based single-image super-resolution

Yu Zhou[a,b], Sam Kwong[a,b,*], Wei Gao[a,b], Xu Wang[c]

[a] Department of Computer Science, City University of Hong Kong, Hong Kong
[b] City University of Hong Kong Shenzhen Research Institute, Shenzhen 51800, China
[c] College of Computer Science and Soft Engineering, Shenzhen University, Shenzhen, China

ARTICLE INFO

Article history:
Received 12 November 2015
Revised 23 April 2016
Accepted 22 May 2016
Available online 8 June 2016

Keywords:
Image super-resolution
Multiple dictionaries
Phase congruency
Complexity reduction
Hierarchical clustering

ABSTRACT

Single-image based super-resolution (SISR) aims to recover a high-resolution (HR) image from one of its degraded low-resolution (LR) images. To improve the quality of reconstructed HR image, many researchers attempt to adopt multiple pairs of dictionaries to sparsely represent the image patches. Conventionally, all the patches with different contents are treated equally, and each patch is coded by multiple pairs of dictionaries, which results in tremendous computational burden in the reconstruction process. In this paper, a phase congruency (PC) based patch evaluator (PCPE) is proposed to divide the LR patches into three categories: significant, less-significant and smooth based on the complexity of the contents. Thus, a flexible multi-dictionary based SISR (MDSISR) framework is proposed, which reconstructs different patches by different approaches. In this framework, multiple dictionaries are only applied to scale up the significant patches to maintain high reconstruction accuracy. Also, two simpler baseline approaches are used to reconstruct the less-significant and smooth patches, respectively. Experimental studies on benchmark database demonstrate that the proposed method can achieve competitive PSNR, SSIM, and FSIM with some state-of-the-art SISR approaches. Besides, it can reduce the computational cost in conventional MDSISR significantly without much degradation in visual and numerical results.

1. Introduction

In recent years, the demand for high-resolution (HR) images promotes the development of super-resolution techniques in multimedia-related fields [22,50]. Single-image super-resolution (SISR) applies signal processing techniques to recover HR images from one of its degraded low-resolution (LR) images. To tackle the ill-posed inverse problem, three categories of methods, including interpolation methods, regularization methods and example-based methods, are well developed. Among them, example-based methods have shown its superiority in obtaining a high-quality scaled-up image [4,7,31] by learning the relationship between LR and HR images from a given set of image samples.

* Corresponding author. Tel: +85234422907; fax: +85234420503.
E-mail addresses: yzhou57-c@my.cityu.edu.hk (Y. Zhou), cssamk@cityu.edu.hk (S. Kwong), weigao5-c@my.cityu.edu.hk (W. Gao), wangxu.cise@gmail.com (X. Wang).

http://dx.doi.org/10.1016/j.ins.2016.05.024

In example-based methods, patch-based processing is usually applied to utilize the redundancy and similarity among images adequately. For different image patches, dictionary learning (DL) based sparse coding approach provides a global and adaptive representation, which has been used in various applications, e.g. image segmentation [41], medical diagnosis [15], 3D shape estimation [21] and signal reconstruction [23,49]. SISR methods based on DL were proposed in [9,39,45], where SISR was formulated as a problem of sparse coding under a single pair of dictionaries in two spaces (one dictionary for HR and one for LR). The target HR patch and its corresponding LR patch were assumed to share the same sparse representation under two-coupled LR and HR dictionaries. Therefore, after LR patch was sparsely represented by LR dictionary, the corresponding HR patch could be obtained by HR dictionary and the sparse coefficients.

It is often the case that an image patch may contain the pixels from different structures, such as line segments, textures, abundant edges, corners, smooth regions or the combination of these structures. Although some regularization terms are added into SISR model to enhance the sparse representation capability in [20,43], it is still insufficient and inaccurate to use one single dictionary to sparsely represent the LR patches [40]. To overcome the shortage, multi-dictionary based SISR (MDSISR) which adopts multiple pairs of dictionaries to reconstruct one HR patch is proposed for various types of images, such as natural image [42,46], remote sensing image [11], textual image [34] and depth image [47]. However, multiple dictionaries usually result in huge computational burdens in the reconstruction process, which is not expected in electronic devices or imaging systems. In addition, for some patches that belong to a single structure, such as textures, over-smoothness may be caused by using multiple dictionaries, which even degrades the reconstruction quality. In fact, due to the redundancy of dictionary atoms and compactness of sparse representation [1,38], a single dictionary performs rather competitively in reconstructing the patches with a single type of structure. Therefore, to reduce the computational cost and maintain the reconstruction quality simultaneously, it is desirable to adaptively reconstruct the LR patches based on the complexity of structures they contain.

To measure the complexity of the structures, it is useful to extract certain features. Phase congruency (PC), a local energy based indicator, is proved to be effective in distinguishing the informative structures, such as line segments, singularities, textures, edges and corners from smooth regions [10,16]. Recently, exploiting PC features has shown its great potential in image or video processing applications, such as verification and identification [26], image registration [5] and foreground extraction [6]. However, it is difficult to measure the complexity of the patches directly based on PC values as one patch may consist of the pixels from multiple informative structures mentioned above.

In this paper, a PC based patch evaluator (PCPE) is proposed to classify the patches into three categories: significant, less-significant and smooth patches. The significant patch contains the pixels from more than one types of informative structures, the less-significant patch consists of only one single informative structure and the smooth patch contains the smooth region. Different from using binary PC map in [48], PCPE employs a hierarchical structure, where the first level divides the patches into smooth patches and non-smooth patches and in the second level, clustering is applied to partition the non-smooth patches into the significant patches and less-significant patches.

We integrate PCPE into the conventional framework of MDSISR. For significant LR patches, multiple dictionaries are applied to reconstruct the HR patches to maintain high reconstruction accuracy. While for the less-significant ones, the faster approach, single dictionary is used to recover the corresponding HR patches more efficiently without much deterioration in quality. In addition, bicubic interpolation, which performs fast and effective in scaling up the smooth region of the images is applied to restore the HR patch of a smooth LR patch. Experimental studies on the benchmark database demonstrate that our proposed PCPE-MDSISR can achieve competitive reconstruction quality without much deterioration compared with conventional MDSISR and save much time in reconstruction process. Particularly, by using Zeyde's method [45] as a baseline, PCPE-MDSISR also outperforms some state-of-the-art SISR methods in PSNR, SSIM and FSIM.

The rest of the paper is organized as follows. Section 2 introduces the related works and the background followed by our proposed PCPE-MDSISR in Section 3. The experimental studies are given in Section 4. And the conclusion is finally made in Section 5.

2. Related works and background

2.1. Related works

Image super-resolution has attracted extensive attention from the researchers and practitioners in image processing area. To reduce the computational complexity, one fast single super-resolution approach, anchored neighborhood regression (ANR) [31] was proposed, where sparse dictionaries and regressors were learned to be anchored to the atoms. In [30], an improved version of the method in [31] was developed, which combined the advantages of anchored neighborhood regression and simple functions. Experimental results demonstrated it achieved the state-of-the-art performances both in quality and efficiency. Other sparse representation based methods included 2D sparse representation [24], sparse support regression [12] and local rank representation [7]. Recently, varieties of machine learning techniques were also applied in image super-resolution to achieve competitive results, such as extreme learning [2] and deep learning [3]. In addition, the efficiency of SISR methods could be improved by using GPU acceleration [13] and simple mapping functions [37].

Sparse coding technology is widely applied in multimedia-related areas. In [44], a semantic-contents based distance metric was learned for image clustering, where the dissimilarity between two patches was evaluated based on the semantic relationship and sparse coding. In [28], a rank preserving sparse learning based dimension reduction method was proposed