Biomedical Signal Processing and Control 7 (2012) 579–590

Contents lists available at SciVerse ScienceDirect

Biomedical Signal Processing and Control

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

Structural similarity regularized and sparse coding based super-resolution for medical images

Shuyuan Yang*, Yaxin Sun, Yiguang Chen, Licheng Jiao

Key Lab of Intelligent Perception and Image Understanding of Ministry of Education, Department of Electrical Engineering, Xidian University, Xi'an 710071, China

ARTICLE INFO

Article history:
Received 31 December 2011
Received in revised form 23 July 2012
Accepted 4 August 2012

Keywords:
Medical images
Sparse coding
Co-occurrence relationship
Structural similarity

ABSTRACT

Recently the single image super-resolution reconstruction (SISR) via sparse coding has attracted increasing interests. Considering that there are obviously repetitive image structures in medical images, in this study we propose a regularized SISR method via sparse coding and structural similarity. The pixel based recovery is incorporated as a regularization term to exploit the non-local structural similarities of medical images, which is very helpful in further improving the quality of recovered medical images. An alternative variables optimization algorithm is proposed and some medical images including CT, MRI and ultrasound images are used to investigate the performance of our proposed method. The results show the superiority of our method to its counterparts.

1. Introduction

High-resolution (HR) images are desired in medical imaging. In order to increase the resolution of medical images, we can either increase the chip size of sensors or reduce the pixel size by sensor manufacturing techniques, which are both severely constrained by the physical limitation of medical imaging systems [1]. Single image super-resolution reconstruction (SISR) can be achieved through algorithmic techniques to obtain a HR image $\mathbf{X}$ from an observed degraded low-resolution (LR) image $\mathbf{Y} : \mathbf{Y} = \mathbf{HBX} + v$, where $\mathbf{H}, \mathbf{B}, v$ represent the downsampling operator, blurring operator and the additive noise, respectively [2]. Depending on the communities involved, it goes by different names including resolution enhancement, HR image reconstruction and super-resolution (SR) reconstruction, which has proved to be helpful in comprehending the medical images [3–6]. Regularization-based method [7,8] and interpolation-based method [9–11] are two popular SISR schemes. Regularization methods pose some assumptions on the ill-posed image recovery problem to obtain a determined solution. However, the performance of regularization methods degrades rapidly with the increase of the magnification factor, and interpolation methods are very limited in modeling the visual complexity of the real images.

In recent years, examples-based method has attracted increasing interests [12–16]. Its principle is that HR image can be predicted by learning the co-occurrence relationship between a set of LR example patches and their corresponding HR version. One of representative works of examples-based SISR method is Chang's method [12], where a LR image is assumed to have similar local geometry to its corresponding HR image. Sparse coding based scheme is consequently developed, which regularizes the recovery of HR image by casting on sparse prior on the local patches. One can generate dictionaries by randomly sampling raw patches from training images of similar statistical nature, which is shown to lead to state-of-the-art result [13–16]. Fig. 1 illustrates the scheme of sparse coding based SISR method.

It is well known that the preservation of details such as edges and contours is very important in computer-aided diagnosis of medical images. If the image priors such as the sparsity, smoothness, self-similarity and multiscale similarity can be incorporated into SISR, the recovery quality of medical images can be remarkably improved. Moreover, the medical images have more homogeneous regions than natural images, such as the background, bones, organs and joints in the computed tomography (CT scan) image, magnetic resonance (MR) images and ultrasound images. However, the available sparse coding based SISR methods do not make avail of other image priors except for sparse prior of patches. In this paper, we introduce structural similarity regularizer into the

* Corresponding author.
E-mail address: syyang@xidian.edu.cn (S. Yang).

1746-8094/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bspc.2012.08.001

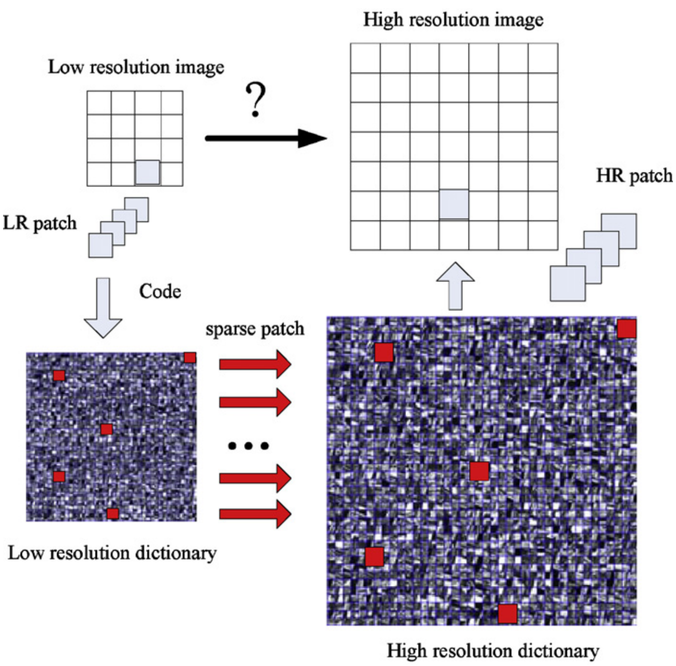

Fig. 1. The scheme of sparse coding based SISR.

recovery algorithm to improve the quality of recovered images, because the non-local structural self-similarity of image has successfully improved the performance of many image processing methods [17,18]. Using a weighted recovery constraint in a non-local region, one can enhance the detail information in the similar patches. By formulating the SISR as a multiple variables optimization problem, we can alternately optimize the variables to obtain the HR image. Some experiments are taken on comparing the performance of our proposed method with its counterparts on some CT images, MRI images and ultrasound images, and the superiorities of our proposed method to its counterparts can be observed in both the visual result and some numerical guidelines.

The rest of this paper is organized as follows. In Section 2, we depicted the structural similarity and sparse coding based SISR approach for medical images. In Section 3, some simulation experiments are taken to illustrate the efficiency and superiority of our proposed method to its counterparts. Finally some conclusions are drawn in Section 4.

2. Structural similarity and sparse coding based SISR

2.1. Sparse coding based SISR

According to [15,16], the LR image patches can be regarded as the degraded version of HR images, and two prototype dictionaries $\mathbf{D}_l$ (LR dictionary) and $\mathbf{D}_h$ (HR dictionary) are constructed, where $\mathbf{D}_h$ is formed by some HR prototype patches sampled from training HR images and $\mathbf{D}_l$ is formed by the corresponding prototype LR patches. Assuming the number of columns of the dictionaries is K and $\mathbf{x}_i \in \Re^n$ is the ith $\sqrt{n} \times \sqrt{n}$ local patch vector extracted from a HR image $\mathbf{X}$ at the spatial location i: $\mathbf{x}_i = \mathbf{R}_i\mathbf{X}$. $\mathbf{R}_i$ denotes a rectangular windowing operator and the overlapping is allowed. Let $\mathbf{y}_i \in \Re^m$ denotes the $\sqrt{m} \times \sqrt{m}$ LR observation vector of $\mathbf{x}_i$: $\mathbf{y}_i = \mathbf{HB}\mathbf{x}_i + \upsilon$ where $\mathbf{H}$, $\mathbf{B}$, υ represent the block downsampling operator, block blurring operator and the additive noise on the patch, respectively [2]. The recovery of $\mathbf{x}_i$ from $\mathbf{y}_i$ under the sparse coding prior can be mathematically written as [16],

$$\begin{cases} \min\limits_{\boldsymbol{\alpha}_i,\mathbf{x}_i} ||\boldsymbol{\alpha}_i||_0 \\ s.t. ||F\mathbf{y}_i - F\mathbf{D}_l\boldsymbol{\alpha}_i||_2^2 \le \varepsilon_1; \\ \quad ||\mathbf{y}_i - \mathbf{HB}\mathbf{x}_i|| \le \varepsilon_2; \\ \quad ||\mathbf{x}_i - \mathbf{D}_h\boldsymbol{\alpha}_i|| \le \varepsilon_3; \end{cases} \tag{1}$$

where F can be an identity matrix or a (linear) feature extraction operator; $\boldsymbol{\alpha}_i$ is the sparse coding coefficient of $\mathbf{x}_i$ under the HR dictionary $\mathbf{D}_h$ and LR dictionary $\mathbf{D}_l$; and $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are the admissible errors. For a LR test image, one can firstly use a feature transformation F to ensure that the computed coefficients fit the most relevant part of the LR image. Considering the recovery of the whole image $\mathbf{X}$, we can reformulate (1) as,

$$\begin{cases} \min\limits_{\{\boldsymbol{\alpha}_i\},X} \sum\limits_i ||\boldsymbol{\alpha}_i||_0 \\ s.t \sum\limits_i ||F\mathbf{y}_i - F\mathbf{D}_l\boldsymbol{\alpha}_i||_2^2 \le \varepsilon_1; \\ \sum\limits_i \left\| \mathbf{y}_i - \mathbf{HB}\mathbf{R}_i\mathbf{X} \right\|_2^2 \le \varepsilon_2; \\ \sum\limits_i \left\| \mathbf{R}_i\mathbf{X} - \mathbf{D}_h\boldsymbol{\alpha}_i \right\|_2^2 \le \varepsilon_3; \end{cases} \tag{2}$$

$\left\|\boldsymbol{\alpha}_i\right\|_0$ is the number of nonzero coefficients of $\boldsymbol{\alpha}_i$ and l_0-norm is nonconvex, and the l_0-norm optimization problem in (1) and (2) is NP-hard in general. In order to solve (2), three approaches can be used: (1) using greedy pursuit algorithms to calculate an approximated minimum l_0-norm; (2) changing the nonconvex l_0-norm to convex l_1-norm and solve it using basis pursuit algorithm and inner points algorithm; and (3) relaxing the nonconvex l_0-norm to nonconvex l_p-norm $(0 < p < 1)$ and solve it by Focuss algorithm. Because l_1-norm is a convex function that is nearest to the l_0-norm, it is often used to replace l_0-norm in the optimization problem. Recent result indicates that as long as the desired coefficients $\boldsymbol{\alpha}_i$ is sufficiently sparse, they can be efficiently recovered by instead minimizing the l_1-norm [19].

2.2. Similarity regularizer

It is well known that there are often many repetitive image structures (or self-similarity) in the medical images. Such nonlocal redundancy is very helpful to improve the quality of reconstructed images [17,18]. In this section, we introduce a non-local structural similarity constraint into patch aggregation to improve the quality of images. For each local patch in an image especially medical image, we can find its similar patches in the whole image according to Gaussian Neighborhood (GN). Therefore the value of the pixel at position (n_1,n_2) in $\mathbf{X}$, $\mathbf{X}(n_1,n_2)$ can then be regarded as a mean of the values of all points whose Gaussian neighborhood looks like the neighborhood of $\mathbf{X}(n_1,n_2)$, that is,

$$\mathbf{X}(n_1, n_2) = \sum_{q,l \in \mathrm{GN}} \mathbf{X}(q, l) w(\boldsymbol{z}^{n_1,n_2}, \boldsymbol{z}^{q,l}) \tag{3}$$

where $\mathbf{X}(q,l)$ is the pixel value at the position (q,l) that belongs to its neighborhood, and $w_{q,l}$ is the connected weights between $\mathbf{X}(q,l)$ and $\mathbf{X}(n_1,n_2)$, which is calculated by

$$w(\boldsymbol{z}^{n_1,n_2}, \boldsymbol{z}^{q,l}) = \frac{\exp(-||\boldsymbol{z}^{n_1,n_2} - \boldsymbol{z}^{q,l}||/h)}{\sum_{n_1,n_2 \in \mathrm{GN}} \exp(-||\boldsymbol{z}^{n_1,n_2} - \boldsymbol{z}^{q,l}||/h)} \tag{4}$$