Sustainable Cities and Society 42 (2018) 1–12

Contents lists available at ScienceDirect

Sustainable Cities and Society

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

Medical image super-resolution via minimum error regression model selection using random forest

Qingyu Dou[a], Shuaifang Wei[b,*], Xiaomin Yang[b,*], Wei Wu[b], Kai Liu[c]

[a] The Center of Gerontology and Geriatrics, West China Hospital, Sichuan University, Chengdu, Sichuan, 610064, PR China
[b] College of Electronics and Information Engineering, Sichuan University, Chengdu, Sichuan, 610064, PR China
[c] College of Electrical Engineering and Information, Sichuan University, Chengdu, Sichuan, 610064, PR China

ARTICLE INFO

Keywords:
Super-resolution
Medical image
Minimum error regression
Random forest

ABSTRACT

Super-resolution is designed to construct a high-resolution version of a low-resolution for more information. Super-resolution can help doctors to get a more accurate diagnosis. In this paper, we propose a novel super-resolution method utilizing minimum error regression selection. In the training step, we partition the patches into multiple clusters through jointly learning multiple regression models. Then we train a random forest model based on the patches of multiple clusters. During the reconstruction step, we use trained random forest model to select the most suitable regression model for the reconstruction of each low-resolution patch. Several medical images are applied to test the proposed method. We compare both the objective parameters and the visual effect to other state-of-the-art example-based methods. Experiment results show that the proposed method has better performance.

1. Introduction

Medical images play a crucial role in modern medicine. High-resolution medical images can help doctors to get a more accurate diagnosis. High-resolution (HR) images provide more information than low-resolution (LR) images. Single image super- resolution (SR) is designed to construct HR version of acquired LR images. The reconstructed HR image should be as close as possible to the true high-resolution version of the low-resolution image. We can directly improve the resolution of the image sensor to obtain high-resolution images. In some special fields, such as satellite remote sensing, infrared photography, hyperspectral imaging and medical imaging, it is difficult to obtain high-resolution images with enough information. In this paper, we focus on medical imaging. Two important types of imaging in medical diagnosis are Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI). CT images are suitable for detecting bone structure while MRI images are mainly used for soft tissue. However, CT and MRI imaging equipment are very expensive. The hardware upgrade of CT and MRI relies on physics, which cause a long cycle to carry out technological innovation. For above reasons, SR is a good choice to improve the resolution of medical images.

Restoring images from a low-resolution version to a high-resolution version is an ill-posed problem. In recent years, various SR algorithms have been proposed. Existing SR methods can be broadly classified into three broad categories: interpolation-based methods, multi-frame based methods, and example-based methods.

Interpolation-based methods are generally based on classical signal processing algorithms, which explore the relationship of pixel values within a neighborhood of single image and then predict sub-pixels based on these selected pixels. Classic interpolation method like Hou and Andrews (1978) and Duchon (1979) are based on sub-pixel prediction. These traditional methods generally result in over-smoothing and failing to introduce more high-frequency and edge information to the new image.

The second category is multi-frame based methods (Farsiu, Robinson, Elad, & Milanfar, 2004; Protter, Elad, Takeda, & Milanfar, 2009; Strecha, Gool, & Fransens, 2007). These methods reconstruct HR image based on multiple LR images of the same scene. In the process of reconstruction, these methods use the complementary information of multiple LR images to restore the missing high-frequency part to obtain HR image. Multi-frame based methods perform better than the interpolation methods in many cases. However, we cannot get enough images which have information complementary in practical applications. These methods will get worse as the magnification factor increases.

The third category is example-based methods. In recent years, the examples-based methods are the most popular methods in super-resolution. These methods use the corresponding relationship between LR

* Corresponding authors.
E-mail addresses: 543909163@qq.com (S. Wei), arielyang@scu.edu.cn (X. Yang).

https://doi.org/10.1016/j.scs.2018.05.028
Received 28 February 2018; Received in revised form 20 April 2018; Accepted 17 May 2018
Available online 28 June 2018
2210-6707/

image patches and HR image patches as prior information of super-resolution. The corresponding relationship is learned from a set of examples. These prior information is put into a collection which is often called dictionary. Existing dictionary learning methods are mostly based on Baker and Kanade (2000), Freeman and Pasztor (2000). The main idea of this method is to learn the corresponding relationship between LR patches and HR patches in a group of samples and then select the most suitable HR patches in the dictionary for each input patch as a reconstruction result. These methods divide a complete image into multiple overlapping image patches. Freeman and Pasztor (2000) search several LR patches which are most similar to input LR patch in dictionary, then use the corresponding HR patches to construct a Markov random fields. HR image can be obtained by solving the Markov random fields. NE methods (Bevilacqua, Roumy, Guillemot, & Morel, 2013; Chang, Yeung, & Xiong, 2004) assume that each input LR image patch can be represented by a linear combination of a series of image patches within a neighborhood of the dictionary. This relationship is then mapped to high-resolution space to reconstruct the HR image. Yang, Wright, Huang, and Ma (2010) first introduced sparse dictionary learning methods into super-resolution. Sparse based methods learn an over-complete dictionary. The dictionary contains only a small number of atoms. Sparse based methods find the linear combination of LR dictionary atoms to represent each input LR patch. The HR patch can be reconstructed by jointly use the corresponding linear combination of HR dictionary atoms. Sparse based methods improve the performance of reconstruction while make the dictionary become concise. Zeyde, Elad, and Protter (2010) improve the efficiency of Yang's method. Optimization of includes training the low-resolution dictionary using the K-SVD, using PCA to reduce computational complexity and OMP to solve the sparse vector for fast. Tian, Wang, Zhou, and Peng (2018), Yang et al. (2017), Ying et al. (2017), Liu et al. (2018), Liu, Blasch, & John (2017a,b), Ben, Meng, Wang, and Yan (2016) improve the sparse method and get better results. Regression methods use a regression model to learn the corresponding relationship of LR and HR patches. Ridge regression methods (Dai, Timofte, & Gool, 2015; Timofte, De, & Gool, 2013) use l2-norm to regulate the process of finding sparse vector instead l1-norm in sparse based methods, thus increasing the speed of reconstruction. The ANR method (Timofte et al., 2013) searches for neighborhoods in over-complete dictionaries and computes a regression mapping matrix from LR to HR neighborhoods to enhance the reconstruction performance. Deep convolutional neural networks (Dong, Chen, He, & Tang, 2014) are also used for super-resolution by training a network model to reconstruct the HR version of the input LR image.

Due to the complex high-frequency information and texture information present in real images, a single dictionary cannot satisfy the requirement that recovering the details in the image. Yang, Liu, Wang, Sun, and Jiao (2011), Dong, Li, Zhang, and Shi (2011) use K-means to cluster samples and train multiple dictionaries to reconstruct different classes of patches. Yang, Wang, Chen, and Sun (2012), Zhan et al. (2015) cluster the samples and train multiple dictionaries according to the geometric similarity of training samples, then select a dictionary which has similar geometry structure to the input LR patch to reconstruct the HR image patch. Yang, Wu, Liu, Chen, and Zhou (2018) use improved fuzzy clustering to cluster the samples. These methods have better performance than the single-dictionary methods, but these methods cannot find the inherent relationship of LR and HR patches.

In this paper, we propose a regression-based learning method. In our method, we train multiple regression models for SR. We assume that each input LR patch can find a most suitable model which have minimum reconstruction error in forming the HR image. The main contributions of our method in SR process are as follows:

1 To ensure that each input LR patch can find the most suitable regression model for reconstruction, the training samples are clustered iteratively based on the reconstruction error. Then each cluster is used to train a regression model. Thus, we get a representative set of regression models.
2 To allow each input LR patch to quickly select the most suitable regression model for reconstruction, we use random forest model as classifier. Random forest is pre-trained and can be applied directly to the reconstruction process. We also adjust the parameters to adjust the performance of random forests and study the effect of these parameters on the reconstruction results. The results show that the clustering results are effective.
3 The proposed method guarantees both the speed and quality of the reconstruction. The regression-based method has a closed solution while the sparse methods can only be solved by approximation. Since the proposed method is based on regression, the time consumption for reconstruction is reduced compared to the previous methods.

The rest of the article is as follows: review the basic idea of SR and introduce the regression-based method and random forest in Section 2; the proposed method is described in Section 3; Section 4 shows the experimental results and the influence of the parameters; the conclusion drawn in this paper is placed on Section 5.

2. Related works

2.1. Single image super-resolution (SISR)

The SISR algorithm attempts to recover the HR version image $\boldsymbol{X}$ from a LR version image $\boldsymbol{Y}$. Follow the assumptions in paper (Yang et al., 2010), the relationship of $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be described as:

$$\boldsymbol{Y} = OB\boldsymbol{X} + w, \tag{1}$$

where O is a down-sampling operator which compress the spatial frequency of $\boldsymbol{X}$, B is a blurring operator which decrease the high-frequency information and texture details of $\boldsymbol{X}$, w is disturbing operator which make $\boldsymbol{X}$ be disturbed by random additive noise. The task of SR algorithm can be considered as finding the HR image $\boldsymbol{X}$ from the LR image $\boldsymbol{Y}$. Mathematically, given a known condition $\boldsymbol{Y}$, the algorithm should give a solution, which satisfy the constraint (1) and have $\hat{\boldsymbol{X}} \approx \boldsymbol{X}$. There is numerous HR image $\boldsymbol{X}$ can satisfy the constraint (1). Thus, SR problem is essentially an ill-posed reverse problem. In order to stabilize the process of solving SR problem, this is, make the reconstruction error $\hat{\boldsymbol{X}} - \boldsymbol{X}_2^2$ as small as possible, we must introduce suitable prior knowledge to constrain the SR problem.

2.2. Ridge regression based super-resolution

Example-based methods always focus on image patches. In example-based methods, we always process the image patches extracted from a complete image instead of processing pixels or entire image. A single image is divided into several patches according to a certain grid. After each LR image patch is restored to an HR image patch by the SR algorithm, we combine all the HR patches into a complete HR image according to the previous grid. The training database is also use the image patch as the operating unit. Assume that we have patches training samples $\boldsymbol{P} = (\boldsymbol{P}_l, \boldsymbol{P}_h)$. Here, $\boldsymbol{P}_l$ is the LR patches and $\boldsymbol{P}_h$ is the corresponding HR patches. The regression model is learned from the training samples $\boldsymbol{P}$. The ridge regression based method assumes that each input LR patch can be represented by a coefficient vector and the training samples by minimize:

$$\min_{\theta} \boldsymbol{y} - \boldsymbol{P}_l\theta_2^2 + \lambda\theta_2, \tag{2}$$

where $\boldsymbol{y}$ is the input LR patch, θ is the coefficient vector and λ is the balance factor. The implication of this formula is that, we want to make the multiplication result of the LR samples and the coefficients $\boldsymbol{P}_l\theta$ as close as possible to the input LR patch $\boldsymbol{y}$. To reduce the computational