Infrared Physics & Technology 76 (2016) 139–147

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Learning-based compressed sensing for infrared image super resolution

Yao Zhao, Xiubao Sui *, Qian Chen, Shaochi Wu

Jiangsu Key Laboratory of Spectral Imaging & Intelligence Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China

HIGHLIGHTS

- Toeplitz matrix is selected as the sensing matrix allows fast reconstruction.
- Proposed method use only one dictionary instead of traditional two dictionaries.
- Training samples are divided into several feature spaces by the proposed method.
- Nonlinear mapping from HR to LR space can be converted into some linear mappings.

ARTICLE INFO

Article history:
Received 19 October 2015
Available online 15 February 2016

Keywords:
Super resolution
Compressed sensing
Toeplitz matrix
Multi-sub-dictionaries

ABSTRACT

This paper presents an infrared image super-resolution method based on compressed sensing (CS). First, the reconstruction model under the CS framework is established and a Toeplitz matrix is selected as the sensing matrix. Compared with traditional learning-based methods, the proposed method uses a set of sub-dictionaries instead of two coupled dictionaries to recover high resolution (HR) images. And Toeplitz sensing matrix allows the proposed method time-efficient. Second, all training samples are divided into several feature spaces by using the proposed adaptive k-means classification method, which is more accurate than the standard k-means method. On the basis of this approach, a complex nonlinear mapping from the HR space to low resolution (LR) space can be converted into several compact linear mappings. Finally, the relationships between HR and LR image patches can be obtained by multi-sub-dictionaries and HR infrared images are reconstructed by the input LR images and multi-sub-dictionaries. The experimental results show that the proposed method is quantitatively and qualitatively more effective than other state-of-the-art methods.

1. Introduction

In recent years, super resolution (SR) has played an important role in the field of image processing. The demand for high resolution (HR) images is also increasing, particularly for HR infrared images. The methods for improving the resolution of infrared images can be divided into two categories. The first category is called micro-scanning. Fortin [1] and Armstrong [2] proposed micro-scanning imaging schemes. However, these imaging devices are always complex, thus limiting their applications. The second category pertains to the application of reconstruction algorithms. The approaches under this category can further be classified into three sub-categories, namely, interpolation- [3], reconstruction- [4], and learning-based methods [5]. Methods based on interpolation are a basic means for achieving SR but generally produce reconstructed images with unsatisfactory quality. Methods based on the reconstruction process assume that low resolution (LR) images are the product of several degradation models, such as down-sampling, motion blurring, optical distortion, and additive zero-mean white Gaussian noise of HR images, and HR images can be recovered through one or multiple LR images. Hardie et al. proposed an infrared super-resolution method that can obtain an HR infrared image from a sequence of observed frames [6]. Farsiu et al. proposed a fast and robust multi-frame super resolution method [7]. Given that SR image reconstruction is an ill-posed problem, many HR images correspond to one LR image. Therefore, prior knowledge must be considered to accurately obtain HR images. Redundancy [8] and edge-directed [9] priors are widely used in reconstruction-based methods. Nevertheless, these priors do not add sufficient new detail to the HR output, particularly at high magnification (e.g., greater than two).

Learning-based methods construct the relationship between the HR and LR images by training datasets that contain millions of co-occurring LR–HR image patches. Yang et al. propose a super-resolution method based on sparse representation [10].

* Corresponding author.
E-mail address: sxbnjust@163.com (X. Sui).

http://dx.doi.org/10.1016/j.infrared.2016.02.001
1350-4495/

There are two coupled dictionaries named D_h and D_l in their algorithm. The two coupled dictionaries make LR image patches and HR image patches share the same sparse coefficients. However, for all image patches, the method only uses one pair of fixed dictionaries to reconstruct HR image patches, which makes the method is not satisfied.

In 2015, we also propose an infrared image super-resolution method based on sparse representation [11], we select detail parts as the object of the proposed method and the method has robustness to non-uniformity. However, there are still two dictionaries in the method. In method [12], the authors propose a single image SR method using CS with a redundant dictionary, and the dictionary is also fixed for all patches. In method [5], the author proposed an efficient single image SR method that learned multiple linear mappings to directly transform LR feature into HR subspaces. However, the author uses the standard k-means algorithm to divide the LR feature into multiple feature subspaces, which is not very accurate. In addition, Yang et al. advocated learning a set of simple mapping functions from numerous image subspaces using multivariate linear regression [13]. Nevertheless, there is still a need to develop the effectiveness of the method.

In this paper, we propose an infrared image SR model based on compressed sensing (CS). A Toeplitz matrix is selected as the measurement matrix because it allows the significantly fast CS reconstruction [14]. Under the framework of proposed method, a HR image patch can be reconstructed by only one dictionary instead of traditional two coupled dictionaries. Moreover, LR feature spaces are divided into several sub-spaces by our proposed adaptive k-means classification method to construct multi-sub-dictionaries. Such multi-sub-dictionaries contain the relationships between HR image patches and LR image patches in each sub-space. Compared with k-means classification algorithm, the proposed classification method does not need to estimate k in advance. The experimental results show that the proposed method is highly computationally efficient and maintains the quality of the reconstructed images.

The rest of the paper is organized as follows. Section 2 describes the principle of the proposed algorithm and the proposed classification method. The SR model and multi-sub-dictionaries are also established in this part. Section 3 discusses the reconstruction of an HR infrared image after the SR model is developed under the CS framework. Section 4 presents a comparison of the proposed method with other state-of-the-art methods from a subjective perspective and with objective data. Reconstruction time is also compared in this section. Finally, Section 5 concludes.

2. Principle of the proposed method

In this study, we establish a SR model under the theory of CS and use a Toeplitz matrix as the measurement matrix, thus making the proposed method faster than other learning-based approaches. In the training phase, we divide the whole training samples into several sub-spaces according to their texture structure. Each subspace has its own mapping from LR to HR feature space. Correspondingly, the feature space in which an input LR image patch belongs to is identified. For an input LR image patch, after the corresponding sub-space is determined, the sparse coefficient can be obtained by the input image, measurement matrix and the dictionary (mapping function). Finally, an HR image is reconstructed by using the dictionary and sparse coefficient. The schematic of the proposed method is demonstrated in Fig. 1.

2.1. Model of SR

CS was proposed by Donoho, Candes, and Tao in 2006 [15,16]. The theory of CS provides a possible means of reconstructing a high-dimensional sparse signal from low-dimensional samplings. Fig. 2 displays the measurement process of CS, where x is the high-dimensional signal, s is the sparse coefficient of signal x under the basis Ψ, and y is the low-dimensional sampling signal. If Θ satisfy restricted isometry property (RIP) principle, then sparse signal s can be well reconstructed by orthogonal matching pursuit (OMP) [17] or total variation (TV) regularization [18] method. In normal conditions, Ψ is selected as the orthogonal basis. Inspired by the theory of CS, we reconstruct HR images under the CS framework.

In this event, s and y are used to depict an HR image under the dictionary Ψ and an LR image, respectively. Here, a Toeplitz matrix is selected as the measurement matrix Φ.

Toeplitz matrix can be expressed as follows:

$$T = \begin{bmatrix} t_n & t_{n-1} & \dots & t_1 \\ t_{n+1} & t_n & \dots & t_2 \\ . & . & . & . \\ t_{2n-1} & t_{2n-2} & \dots & t_n \end{bmatrix} \tag{1}$$

where every left-to-right descending diagonal is a constant, i.e., $T_{i,j} = T_{i+1,j+1}$. T' denotes the first m rows of T.

where m is the number of rows, and $\sqrt{n/m}$ is the amplify factor.

On the basis of the above theory, the proposed model can be formulated as follows:

$$Y_k = T'X_k + n = T'D_k\alpha + n \tag{2}$$

where Y_k is the LR image patch that belongs to k_{th} feature space, T' is the measurement matrix under the CS framework, D_k is the k_{th} dictionary containing the relationships between the LR and HR images from the k_{th} feature space, α is the sparse coefficient of the HR infrared image under D_k, and n is the noise vector. Here, different feature spaces contain different texture structure image patches.

The advantages of the SR model are that there is only one dictionary in each feature space instead of traditional two coupled dictionaries. And different dictionaries exist in the different feature space will improve the restoration results of the HR infrared images. Besides, Toeplitz matrix is selected as the measuring matrix allows the significantly fast CS reconstruction.

2.2. Establishment of image patch library

The library of the image patch is important to the performance of the proposed method. In particular, the selection of image patches is remarkably important. In theory, the dictionaries are more stable when experiments include more sample types. In our experiments, images of the human body, human face, buildings, cars, outdoors, and so on are included in the image patch library. The number of image patches is also relevant to the performance of the proposed method. If the number of patches is extremely few, the mappings cannot capture all the texture features and the performance of the method will not be satisfied. As long as the library includes enough features, then the library is acceptable and can obtain stable sub-dictionaries.

2.3. Classification of image patches

In this section, the image patches are classified into several feature spaces. Each space represents one specific feature structure. In Ref. [5], the author adopted the standard k-means algorithm [19] to divide the samples into multiple feature sub-spaces. However, by using an ordinary approach, k can hardly be estimated. Furthermore, selecting the initial point is important to the algorithm. In our method, we propose an accurate approach for classifying image patches. This technique is called the adaptive k-means algorithm. Adaptive signifies that the number of k is decided by the