Infrared Physics & Technology 71 (2015) 506–513

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

A novel infrared image super-resolution method based on sparse representation

Yao Zhao *, Qian Chen, Xiubao Sui, Guohua Gu

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

HIGHLIGHTS

- Instability of the sparse decomposition and self-similarity of images are considered in the training stage.
- Detail parts are selected as the object of the proposed method.
- Our method has robustness to non-uniformity.
- Proposed method greatly reduces the artifacts caused by the instability of the sparse decomposition.

ARTICLE INFO

Article history:
Received 10 February 2015
Available online 2 July 2015

Keywords:
Super-resolution
Infrared images
Sparse representation
Multi-dictionary
Reconstruction artifacts
Infrared imaging system

ABSTRACT

This paper presents a novel method for the reconstruction of super-resolution infrared images based on sparse representation. Assume we can get a pair of dictionaries which makes the low-resolution image patches share the same sparse representation with high-resolution image patches, then the high-resolution image patches can be reconstructed through the sparse representation of low-resolution image patches. Firstly, considering the instability of the sparse decomposition and the self-similarity of the image patches, the stable multi-dictionary pairs can be obtained by training the classified samples twice. Then, in order to get more accurate sparse coefficients, detail patches are selected as the object of our method. Finally, high-resolution infrared images can be reconstructed. In addition, some non-uniform images with fixed pattern are added into training samples, allowing our algorithm robust to infrared imaging system. The proposed method can greatly reduce the reconstruction artifacts caused by the instability of the sparse representation and have a high reconstruction precision.

1. Introduction

With the rapid development of image processing technology, super-resolution has become the research hotspot. The demands for high-resolution (HR) images are becoming more and more urgent, especially for HR infrared images. However, affected by the fabrication technique and material properties, it is difficult to manufacture high density infrared focal plane array (IRFPA). Meanwhile, infrared imaging devices need certain quantum efficiency. If pixel's physics size is too small, weak signals will not be detected. Besides, restricted by the diffraction limit, even if the detector array is big enough, it does not help to improve spatial resolution of infrared images. Therefore, it is useless to improve spatial resolution by increasing the number of pixels on per unit area. One approach to obtain HR images is through the fixed displacement of the controlled devices, which is known as micro-scanning in the field of infrared. Armstrong [1] put forward a micro-scanning imaging scheme based on infrared optical system in 1996. In order to achieve the purpose of translation images, there are four parallel-plate-refractors with different angles in the device. However, due to the complexity of the equipment, its application is limited. Another way to obtain HR infrared images is by algorithm entirely. That is to say, we can recover an HR image through one or multiple low-resolution (LR) images. Many researchers have done a lot of theoretical research on how to reconstruct HR images and got the degradation models from HR to LR images [2], including motion blur, optical jitter and down-sampling. Many algorithms have been proposed to obtain HR images [3–8]. In general, these methods can be divided into three types, which are interpolation-based methods, regularization-based methods, and learning-based methods. The interpolation-based methods construct an HR image by projecting LR images to the reference image. This process usually consists of three steps, registration stage, non-uniformity interpolation stage, and deblurring stage. Ur and Gross [3] performed a non-uniformity

* Corresponding author.
E-mail address: zhaoyaosunshine@163.com (Y. Zhao).

http://dx.doi.org/10.1016/j.infrared.2015.06.017
1350-4495/

interpolation of a set of spatially shifted LR images by utilizing the generalized multichannel sampling theorem of Papoulis [4] and Brown [5]. However, degradation models are limited in interpolation-based method and the optimality of the whole reconstruction algorithm is not guaranteed. The basic idea of regularization-based method is to use regularization strategy to incorporate prior knowledge of the HR images. A maximum likelihood (ML) and maximum a posterior (MAP) estimation have been applied to super-resolution reconstruction [6–8]. Bahy [9] proposed an adaptive regularization-based super-resolution reconstruction method in 2014. However, image sequences with sub-pixel displacement are needed in the two methods mentioned above.

In recent years, the learning-based methods offer a new way to acquire HR images [10–12]. These algorithms obtain the relationship between HR images and LR images by learning the training samples, and restore the HR images based on the relationship. Freeman et al. [10] got the relationship by Markov random field, and restored the HR image successfully. However, massive training samples are needed in this method. Chakrabarti et al. [13] proposed a kernel principal component analysis based on the prior model to generate super-resolution face images. Chang et al. adopt the philosophy of locally linear embedding (LLE) [14,15] from manifold learning, assuming similarity between the two manifolds in the HR and the LR patch spaces. Yang et al. [16,17] proposed a super-resolution reconstruction method based on sparse representation, and it is applied to visible light and specific human face images restoration. However, the method will result in noticeable reconstruction artifacts due to the instability of sparse decomposition.

In this paper, we carry out a research on infrared image super-resolution and propose a novel infrared image super-resolution method based on sparse representation. The proposed method can greatly reduce the reconstruction artifacts caused by the instability of the sparse representation while have a high reconstruction precision at the same time. In order to adapt to infrared imaging system, some LR images with fixed non-uniformity pattern are added into training samples. Hence, we can obtain a pair of dictionaries with robustness to non-uniformity. The experimental results show that the proposed algorithm can reconstruct HR infrared images accurately.

The paper is organized as follows. Section 2 describes the principle of the proposed algorithm and the innovations of the article; Section 3 describes the method of how to obtain a pair of stable dictionaries; Section 4 briefly describes the way how to get sparser detail images and verifies the sparsity between extracted detail images and original images. HR detail images and base images can be obtained in Section 5; Section 6 is the simulation results of the algorithm; Section 7 concludes the paper.

2. Principle of the proposed algorithm

Any signal in R^N can be represented in terms of a basis of $N \times 1$ vectors. Generally, the basis D is orthonormal. The signal x can be represented as

$$x = D\alpha \tag{1}$$

where α is the $N \times 1$ column vector made up of α_i. If non-zeros elements in α is less than N, the signal is considered sparse.

LR image Y is a blurred and down-sampled version of the HR image X. Y can be sparse represented as α_l under the basis of D_l, X can be sparse represented as α_h under the basis of D_h. We call D_l and D_h as dictionaries in the following part. If $\alpha_l = \alpha_h$, then HR image can be restored by α_l and D_h. The key of the method is to obtain a pair of dictionaries satisfy HR image patches share the same sparse representation with LR image patches.

It seems to be a perfect way to obtain HR images. Nevertheless, due to the instability of the sparse decomposition, those image patches with repetitive structures and patterns will share different sparse coefficients, which will result in noticeable reconstruction artifacts and lose the faithful details.

Compressed sensing [18,19] (CS) provides the theoretical basis for sparse representation. If image patch y_1 is similar to y_2, then their sparse coefficient α_1 and α_2 will also have a high similarity theoretically. Due to the instability of the sparse decomposition, if α_1 is largely different from α_2, which will result in reconstruction artifacts inevitably. To solve this problem and get HR images accurately, the following two methods will be introduced in the next two sections.

(i) The performance of the algorithm depends on the stability of the dictionaries. Therefore, we take the instability of the sparse representation into consideration in the process of training dictionaries, dividing training process into two steps. Firstly, training samples are classified into several clusters, and then two dictionaries can be obtained by training them. Next, continue training two dictionaries under the constraint condition. The constraint condition will be introduced in Section 3. Finally, a pair of stable dictionaries can be obtained.

(ii) In Section 4, we demonstrate the extracted detail images have a better sparsity than the original infrared images. Based on this consideration, LR detail patches are used to restore HR detail patches, which make our method more accurate than others.

The intact principle scheme of proposed algorithm is clearly illustrated in Fig. 1.

3. Obtain a pair of stable dictionaries

The stability of the dictionary is critical for sparse decomposition. In order to get a pair of stable dictionaries, some HR images and LR images patches are needed as the training samples. In our experiments, human body images, human face images, building images, car images and outdoor images, etc. are included in our training samples. Theoretically, the more sample types the experiments include, the more stable the dictionaries are. Meanwhile, the number of the samples is also important to the performance of proposed method. If the number of samples is too few, the dictionaries cannot capture all the features, and the performance of the method will not be very well. In fact, as long as the samples include enough features, then they are acceptable and can obtain stable dictionaries. And then, training samples are classified into several clusters. Each image patches in the same cluster share a very high similarity. The method how to classify these image patches will be illustrated below. Schematic diagram is shown in Fig. 2.

Step (i) Choose one image patch from LR training samples arbitrary. Assume the selected image patch is y_1, and then calculate the error value between y_1 and others image patches. Error function is as follows

$$ERR_k = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(y_k(i,j) - y_1(i,j))^2}{m \times n} \tag{2}$$

where $y_k(i,j)$ is the (i,j)th pixel's value of the y_k, m and n are the column and row numbers of the image patch. k from 2 to N, N is the total number of image patches. If $ERR_k < \xi$, y_1 and y_k belong to