Infrared Physics & Technology 67 (2014) 477–489

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Fusion method for infrared and visible images by using non-negative sparse representation

Jun Wang [a], Jinye Peng [a,b,*], Xiaoyi Feng [a], Guiqing He [a], Jianping Fan [b]

[a] School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China
[b] School of Information Science and Technology, Northwest University, Xi'an 710069, China

HIGHLIGHTS

- The non-negative sparse representation (NNSR) is introduced for image fusion.
- The activity and sparseness levels are used to describe the NNSR coefficients.
- Multiple feature extraction methods are developed to extract source image features.
- A new fusion rule is presented for non-negative sparse representation.
- The proposed method can outperform most classical and state-of-the-art methods.

ARTICLE INFO

Article history:
Received 22 April 2014
Available online 28 September 2014

Keywords:
Image fusion
Non-negative sparse representation
Infrared images
Visible images

ABSTRACT

In this paper, an interesting fusion method, named as NNSP, is developed for infrared and visible image fusion, where non-negative sparse representation is used to extract the features of source images. The characteristics of non-negative sparse representation coefficients are described according to their activity levels and sparseness levels. Multiple methods are developed to detect the salient features of the source images, which include the target and contour features in the infrared images and the texture features in the visible images. The regional consistency rule is proposed to obtain the fusion guide vector for determining the fused image automatically, where the features of the source images are seamlessly integrated into the fused image. Compared with the classical and state-of-the-art methods, our experimental results have indicated that our NNSP method has better fusion performance in both noiseless and noisy situations.

1. Introduction

Image fusion is an active research topic in optical signal processing and computer vision, which is used to combine the useful information from two or more images which are obtained by multiple sensors or one single sensor at different situations. Especially, the infrared and visible sensors are commonly used [1,2].

The infrared image records thermal radiations emitted by the objects in a scene, where the target and contour features or even the camouflaged targets could be observed easily. However, the infrared image has lower contrast and its details are usually weak. On the other hand, the visible sensor is more sensitive to the reflection of a scene with a high definition, so it contains more detail information of the scene [3]. Because of the favorable complementarity between the infrared images and the visible images, we can obtain more comprehensive, accurate and concise information by performing image fusion methods in many fields, such as military detection, public security and remote sensing [4].

It is well known that there are two key issues to address the problem of image fusion: (1) How to effectively extract the image information from the original images? (2) How to reasonably combine the information from multiple information sources into the final fused image? To solve the first question, many image representation methods have widely studied. The typical methods are pyramid transform [5], wavelet [6], Curvelet [7] and sparse representation [8]. All these existing methods except sparse representation are highly structured and can analyze the image information and extract the image features at multiple scales. Because all these existing methods for image fusion fixed their basis functions for image analysis, the edges of the images are not accurately expressed, which may seriously affect the fusion results. In contrast, the sparse representation method for image fusion learns

* Corresponding author at: School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China.

http://dx.doi.org/10.1016/j.infrared.2014.09.019
1350-4495/

a dictionary from a set of training images. The atoms, which are learned to compose the dictionary more finely, can effectively reconstruct the original images and have strong ability to limit the misleading effects of the image noises. As a result, the sparse representation method usually products better fused images than many other existing fusion methods [8–14]. The major weakness of the sparse representation method is that it does not completely analyze the features of the source images according to the activity levels of their sparse representation coefficients, and the reason is that the physical meaning of the basis atoms in the dictionary is not clear and strong. Thus the sparse representation method for image fusion may tend to lose the texture features of the visible images when we fuse the infrared images with the visible images.

To solve the second question, most existing methods focus on studying different fusion rules. The typical fusion rules are "choose" and "weighted" [13]. The "choose" rule is to select the image features from only one source image, its advantage is that the features from the selected source images are completely integrated into the fused images, and its disadvantage is that the features of other source images are completely discarded. Thus the fused images, which are obtained by using the "choose" rule, tend to be oversharp and less smooth. On the other hand, the "weighted" rule, can have better performance in some particular aspects by combining the features from multiple source images completely. But when the ratio of the weights of the features in the source images drops, the innovation features appear less in the fused image with varying degrees. In particular, for the infrared and visible image fusion, the details for some interest target regions may be weaken, which may seriously affect the quality of the fused images and their practical applications.

Based on these observations, we try to attack the problem of image fusion from two aspects, i.e. image feature extraction and new fusion rule for the infrared and visible images. The contributions of this paper reside in three aspects:

1. The non-negative sparse representation is introduced for image fusion. The non-negative sparse representation can encode the source images efficiently by using few 'active' components and the non-negative constraints can make the representation purely additive (allowing no subtractions). Such sparse representation can achieve an easy or intuitive interpretation of the encodings of the source images.
2. An interesting method is developed to describe and analyze the non-negative sparse coefficients from two aspects, i.e. activity level and sparseness level, which may allow us to analyze the image features more comprehensive. The sparseness level is one of the important indicator in sparse representation, and it is used to describe the effectiveness of the sparse coefficients for image representation.
3. A new fusion rule is presented for non-negative sparse representation. Considering that the different features should be combined by using different rules and some important features (such as the target features for the infrared images) should be fused without loss, multiple methods are developed to extract the salient features from the source images, which include detecting the target and specific contour features from the infrared images and extracting the detail features from the visible images. The regional consistency rule is then proposed for fusing different features.

The rest of this paper is organized as follows. In Sections 2, a brief review of some most relevant work is presented. In Section 3, our proposed method for infrared and visible image fusion is introduced. We present and discuss our experimental results in Section 4. Finally, we conclude this paper in Section 5.

2. Related work

Most existing methods for image fusion can be classified into two categories: multi-scale transform based methods and sparse representation based methods.

For the multi-scale transform-based fusion method, the basic idea is that the salient information of the source images is closely related to their multi-scale decomposition coefficients. Most existing multi-scale transform-based fusion methods usually consist of three steps, including decomposing the source images into multi-scale coefficients, fusing these coefficients with one certain rule, and reconstructing the fused images by using inverse transformation. Burt et al. and Kolczynski [15] have developed the image fusion methods by performing pyramid transform, including Laplacian pyramid, gradient pyramid (GP), these transform methods obtained the fused images by combining the coefficients of the source images through "choose max" and "weighted" of image patch variance, respectively. Pajares and Cruz [16] systemically studied image fusion by performing discrete wavelets transform (DWT) methods and provided the guidelines about the use of wavelets in the process of image fusion. Shao et al. [17] introduced the focus measure operators into the Curvelet domain (denoted as Curvelet for simple), where the approximation of the source images and the detail coefficients of the source images are fused separately by the local variance weighted strategy and the fourth order correlation coefficient match strategy. These multi-scale transform-based fusion methods can decompose the source images on multi-scale and multi-direction, but they cannot represent the image details adaptively because their basis functions are fixed in advance, thus they may produce artificial or Gibbs effects in the fused images.

For the sparse representation based methods, the basic idea is that the image signals can be represented as a linear combination of a "few" atoms from a pre-leaned dictionary and the sparse coefficients are treated as the salient features of the source images. The main steps for the sparse representation based methods include: (a) dictionary learning; (b) sparse representation of the source images; (c) fusing the sparse representations by the fusion rule; (d) reconstruct the fused images from their sparse representations. Yang and Li [10] have trained a dictionary by using K-SVD method and adaptively represented the source images by using their sparse coefficients, and then combined the sparse coefficients by using the "choose max" rule, which is decided by the l_1-norm (the sum absolute value) of the sparse coefficients (i.e. it is also called as the activity level). Ding et al. [11] have showed that the fusion method, which employed the dictionary trained by using the infrared images and the fusion rule of maximum absolute of entry of sparse coefficients, has obtained almost the largest objective evaluations. Wang et al. [12] developed a dictionary learning scheme in NSCT domain and fused the salient features of the images in low- and high-frequency sub-bands respectively, and they have obtained better fusion performance than the traditional fusion method which is based on single sparse representation and DWT, NSCT (nonsubsampled contourlet transform). Yu et al. [13] have used joint sparse representation to extract the common and innovation features of the source images, and combined them according to the activity level of the innovation coefficients. Zhang et al. [14] proposed a fusion method, the local means, which are similar to the low frequency coefficients, are subtracted from the source images, where the part removed means are integrated with the fusion rule same as the Yu's method and the local means are fused by the "choose max" rule.

Because all these existing sparse representation based methods can achieve more meaningful representations of the source images and learn a dictionary with finer fittings of the original images [18],