G Model
IJLEO-58353; No. of Pages 11
ARTICLE IN PRESS

Optik xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Texture clear multi-modal image fusion with joint sparsity model

Zhisheng Gao [a,b,*], Chengfang Zhang [b]

[a] The State Key Laboratory of Astronautic Dynamics, Xian 710038, China
[b] Center for Radio Administration & Technology Development, Xihua University, Chengdu 610039, China

ARTICLE INFO

Article history:
Received 9 June 2016
Accepted 13 September 2016

MSC:
00-01
99-00

Keywords:
Multi-modal image fusion
Joint sparsity model
Fusion rules
Texture clear

ABSTRACT

Multi-modal image fusion is necessary for describing a target abundantly. With the consideration of the correlations between multi-source signals and the sparse characteristics of image, this paper proposed a novel fusion rule of multi-modal image fusion scheme based on the joint sparsity model. First, the source image was represented as a shared sparse component and an exclusive sparse component with an over-complete dictionary. Second, the designed novel fusion rule acts on the shared and exclusive sparse coefficients to obtain the fused sparse coefficients. Finally, fused image was reconstructed by the fused sparse coefficients and the dictionary. The proposed approach was tested on the infrared and visual images, medical images. The results were compared with those of traditional methods, such as the multi-scale transform based methods, sparse representation based methods and joint sparsity representation based methods. Experimental results demonstrated that the proposed method outperforms the existing state-of-the-art methods, in terms of better texture clarity. Moreover, the fused image shows better edge consistence and visual effect.

1. Introduction

Currently, numerous imaging sensors have been widely applied to various fields. Multiple images on a same scene such as medical images, infrared and visual images can be obtained by different sensors. It is necessary for multi-sensor image fusion to carry out computerized processing, artificial observation and monitoring, which makes fused image more suitable for human visual perception. The main idea of image fusion is that the same scene images from the different sensors are carried out de-noising, time registration, spatial registration and re-sampling, and then the fused image is combined with the inherent complementary information extracted from the multiple images with different multi-modal censors. Image fusion have been widely used in various fields, such as remote sensing [1], medicine [2], military [3].

Pixel level image fusion algorithms can be divided into two types: spatial domain algorithms [4] and transform domain algorithms [5]. Pixel level image fusion algorithms mainly involve three steps: extracting image features, merging the image features and reconstructing the fused image. With the development of the theory of multi-scale transformation (MST), various multi-resolution fusion algorithms were proposed, including the ratio of low-pass pyramid (RP) [6], discrete wavelet transform (DWT) [7], dual-tree complex wavelet transform (DTCWT) [8], stationary wavelet transform (SWT) [9], curvelet transform (CVT) [10] and non-subsampled contourlet transform (NSCT) [11]. These fusion algorithms based

* Corresponding author at: The State Key Laboratory of Astronautic Dynamics, Xian 710038, China.
E-mail address: xihua_gzs@sina.com (Z. Gao).

http://dx.doi.org/10.1016/j.ijleo.2016.09.126

Please cite this article in press as: Z. Gao, C. Zhang, Texture clear multi-modal image fusion with joint sparsity model, Optik - Int. J. Light Electron Opt. (2016), http://dx.doi.org/10.1016/j.ijleo.2016.09.126

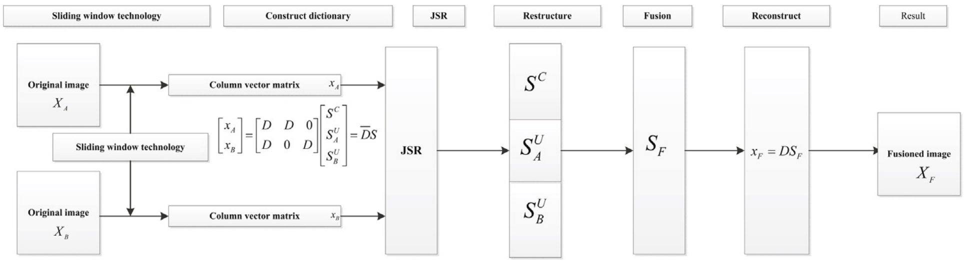

Fig. 1. The framework of joint sparse representation model.

on multi-scale transform have achieved pretty good fusion effect [12]. However, two main drawbacks still exist in them [13]: the first one is the loss of contrast. Since multi-modal images have different physical characteristics, the same area of different source images may have different brightness characteristics. The second one is the difficulty in selecting the decomposition level and decomposition type.

With the mature development of the sparse theory [14,15], various image fusion methods based on sparse representation (SR) have been proposed. Image fusion based on sparse representation was proposed by Yang and Li [16] in 2010. In [16], DCT is used to construct a dictionary. The maximal-one norm of all source sparse coefficients vectors was selected as the fused rule. Although the fusion effect of Ref. [16] is better than multi-scale transform, there are also have some defects: the first one is the fine details in source images such as textures and edges tend to be smoothed from sliding window technology. The second one is computational efficiency is suffered by sliding window technology. The third one is the spatial inconsistency in fused image would be brought by the using the maximal-one norm simply. And the fourth one is the intrinsic links among source images are not taken into account.

Image is a special kind of signal, and there would have certain correlations among the images with same scenes. In order to fuse all the information of source images adequately, multi-modal image fusion based on joint sparsity representation (JSR) is firstly proposed by Yin and Li [17]. In their work, a new fusion rule and the designed joint sparsity model are applied to multi-modal image fusion. Although the fusion effect of Ref. [17] is better than multi-scale transform, there are still some limitations in the fusion rules. Such as, Ref. [18] suggested that subtracted local means are not stable for texture features. However, Yin et al. used them as the weights of low frequency region. Moreover, in [17], fusion strategy only focused on the fusion weight of low frequency, while ignoring the image fusion weight of high frequency. However, the high frequency component of the image plays a more important role in the texture preserving and texture clarity of the fused image.

Most algorithms such as MST, SR and JSR have received good effect in some images. However, their results on clear texture and consistent edge are non-ideal when treating images such as medical images, infrared and visible images. To solve such problems, we presented a new novel multi-modal image fusion scheme based on joint sparse representation inspired by Yin and Li [17]. Our main contributions are the block average fusion rules and the innovation joint sparse coefficients fusion rules. The main idea of our method is that the innovation joint sparse coefficients can represent the high frequency of source images, which contains a lot of information such as edge and texture. After using our designed new fusion rules, the fused image can not only preserve the complementary information of the original image, but also has a better texture clarity, so as to get a better visual observation effect. Furthermore, the fused image allows to highlight the key features of the image more easily.

The rest of this paper is organized as follows. Section 2 introduces a novel method of fusion algorithm based on dictionary learning and joint sparse representation. Section 3 describes the experiments conducted to examine the effectiveness of the proposed method. Finally, Section 4 concludes this work.

2. Texture clear image fusion scheme

2.1. Joint sparsity model

Ref. [19] has proposed three different joint sparse models: JSM-1, JSM-2 and JSM-3. JSM-1 is the most suitable for image fusion [17]. In JSM-1 model, all signals in the signal group share a shared sparse component while each exclusive signal contains a sparse innovation component. The general framework of image fusion based on JSM-1 is shown in Fig. 1. Note that all signals are composed of shared sparse components and exclusive sparse component by the theory of the distributed source encoding. Suppose that $X = \{x_1, x_2, \ldots, x_J\} \in \Re^{JN}$ is a set of signals, and $x_i \in R^N$, $i = 1, 2, \ldots, J$, then:

$$x_i = Ds^C + Ds_i^U, \quad i = 1, 2, \ldots, J \tag{1}$$

Please cite this article in press as: Z. Gao, C. Zhang, Texture clear multi-modal image fusion with joint sparsity model, Optik - Int. J. Light Electron Opt. (2016), http://dx.doi.org/10.1016/j.ijleo.2016.09.126