Optik 126 (2015) 5127–5133

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Fusion of infrared intensity and polarization images using embedded multi-scale transform

Su-zhen Lin*, Dong-juan Wang, Xiao-hong Zhu, Shang-min Zhang

Key Laboratory of Instrumentation Science and Dynamic Measurement, North University of China, Taiyuan 030051, China

ARTICLE INFO

Article history:
Received 24 October 2014
Accepted 9 September 2015

Keywords:
Image fusion
Infrared polarization image
Infrared intensity image
Embedded multi-scale transform
Support value transform

ABSTRACT

Multi-scale transforms are considered to be the effective methods in image fusion. However, infrared intensity and polarization images fusion based on these methods frequently arise the issues such as limited contrast ratio improvement and marginal area distortion. The embedded multi-scale decomposition has been proposed to integrate the advantages of different methods in this paper, where the low-frequency image continues to be decomposed by another multi-scale transform. The detailed procedure is shown as the following. Firstly, the low-frequency components and support value image sequences of infrared intensity and polarization images are obtained respectively with support value transform (SVT). Secondly, multi-scale bright and dim information are extracted from the last layer of the low-frequency images using the multi-scale top-hat decomposition, respectively. Thirdly, the resulted images are respectively fused and enhanced, and then fused with the two low-frequency images to get the fused low-frequency image. After that, the image is reversely transformed with the support value image sequences fused by extracting maximum gray to get the final image. Compared with the simple SVT and the multi-scale top-hat decomposition, the method suggested successfully improves the contrast ratio, the distortion of marginal area and the local coarseness. The validity of the method proposed is proven.

1. Introduction

Unlike the traditional infrared intensity imaging technology, infrared polarization imaging uses infrared polarization angle, degree and other physical quantities to form images [1], which means different results will be produced if these two types of technology are applied to explore the same scene. The research conducted by Rogne et al. [2] to study the polarization feature of paint samples shows that the infrared polarization image can effectively suppress the disordered background in the infrared intensity image. Bendor et al. [3] tests polarization features of different targets in the infrared band, which finds that polarization images can be used to separate artificial targets from the natural background. Related polarization experiments done in mines [4], automobiles and tents [5] also prove the effectiveness of infrared polarization imaging in eliminating background noise and improving target detection ratio. Researchers in National University of Defense Technology and other universities of China conduct similar experiments in metal and camouflage screen [6–8]. The results prove the above-mentioned advantages of the infrared polarization imaging and its complementarities with the infrared intensity imaging [9]. Thus, image fusion becomes the key element in developing the bimodal imaging.

In recent decades, researches conducted on image fusion mainly focus on the different wave bands, such as the visible light and infrared wave band [10], medium and long [11], short and medium [12], dual-color medium wave infrared [13] and multi-bands. Depending upon the phases of unification, fusion scheme is classified into three levels namely pixel, feather and decision level [14]. The pixel level fusion combines directly the pixel values of the fusing images and renders a composite image [15]. It is attracted by many researchers and various techniques have been developed. Basically, there are two common methods of pixel level fusion technology: spatial domain methods and transform domain methods. Multi-scale transforms are considered to be the effective methods in the image fusion, such as pyramid transform [16], the discrete wavelet transform [17], the dual-tree complex wavelet transform [18], ridgelet transform [19], the curvelet transform [20], the contourlet transform [21], the nonsubsampled contourlet transform [22], and the support value transform [23]. Despite of their distinctive features, all the multi-scale transform methods share one common rule, that is, they firstly divide the original images into the high and low components, respectively, fuse these two kinds of images, and then inversely transform the fused results to form

* Corresponding author. Tel.: +86 13753167088; fax: +86 10351 3923623.
E-mail address: 1966921@sina.com (S.-z. Lin).

http://dx.doi.org/10.1016/j.ijleo.2015.09.154
0030-4026/

the final image. All the purpose is to use the same rule to fuse the sub-images of the same or similar property. It can integrate the difference information between original images better, keeping the original edge smooth and complete in these methods. With the development of the infrared polarization imaging, these methods are applied in infrared intensity and polarization images [24,25]. Similar studies are also conducted by the Chinese researchers. Chen et al. use wavelet packet to fuse the infrared polarization images [26]. Researchers in our team fuse the bimodal infrared images through support value transform and fuzzy combination rules [27,28]. Xu et al. hold deep research on the long wave infrared polarization image fusion and false polarization information [29]. Zhao et al. study an adaptive weighting multi-band polarization images fusion method [30]. In addition, the technology of using colorful fusion to process polarization images [31], of using D–S evidence theory to fuse polarization images [32] and the proposition of near-infrared polarization and thermal imagery data level fusion model [33], all of which have produced good results. But it fails to improve the contrast ratio and clarify the marginal area of the infrared imaging, which have, up till now, been the deep-seated problems in this field.

Moreover, we human beings are more sensitive to the contrast ratio in detecting the targets [34]. Refs. [35,36] use multi-scale top-hat decomposition [37] in the mathematic morphology to extract the bright and dim features of visible light and infrared images to fuse with the original images so as to improve the contrast ratio and extend the scope of gray value. Despite of these advantages, the method fusing the bimodal images has the following limits,

(1) Improvement of the contrast ratio is not great.
(2) Distortion of the marginal area is generally serious.
(3) The method raised in Ref. [36] can help to reduce distortion but lead to decrease the contrast ratio.

A new method is presented in this paper to solve these issues. The support values are first transformed to form the low-frequency components (LFC) and support value image sequences (SVIS) of the two modal images, respectively. Multi-scale top-hat decomposition is then applied to extract the bright (bright information, BI) and dim information (dim information, DI) from the last layer of low-frequency images. Following that, the bright and dim information fused images are enhanced to fuse with the two low-frequency images. Last, the low-frequency fused image is conducted inverse transform with the support value fused image sequences to produce the final image. Please note that only multi-scale pixel level image fusion will be addressed in the course of this work. In addition, all input images are assumed to be registered.

2. Fusion method

2.1. Fusion scheme

Fig. 1 shows the method, in which, embedded multi-scale decomposition (EMD) refers to the further decomposition of the low-frequency image by another multi-scale decomposition.

The detailed procedure is shown as follows:

(1) Embedded multi-scale decomposition is employed to extract the low-frequency, high-frequency, bright and dim information from the bimodal images, respectively.
(2) The bright and dim information of the bimodal images are fused and enhanced, respectively.
(3) The low-frequency images, enhanced bright and dim information fused images are integrated.
(4) The high-frequency images are fused layer by layer.

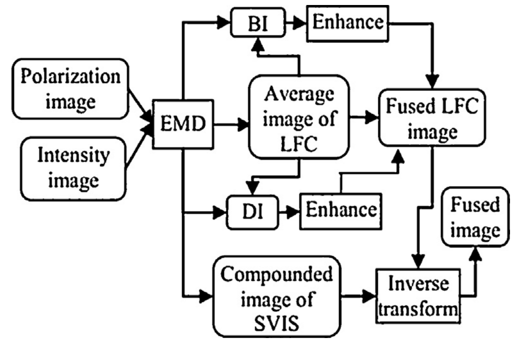

Fig. 1. Schematic diagram of infrared polarization and intensity images fusion.

(5) The fused results in step three and four are reversely transformed to get the final image.

2.2. Basic algorithm

2.2.1. Embedded multi-scale decomposition

Considering the sensitivity to the local information of human eyes, multi-scale transform can decompose the original image into multiple scales, and use the different rules to fuse low-frequency and high-frequency information, respectively, which can achieve a good result. While SVT, a multi-scale transform method developed from the least squares support vector machine, LS-SVM, can run fast, produce no ring effect and achieve shift-invariance property [23]. Furthermore, the study conducted by our team also proved the effectiveness of this method in simplifying the algorithm and achieving better fusion result [13,38]. Besides, since the inverse transform of the support value needs only the last layer of the low-frequency image and the support value image sequences, it can greatly facilitate the embedded multi-scale decomposition. Top-hat decomposition is proven as a effective method to extract the bright and dim information and therefore, together with SVT mentioned above, becomes the choose of this paper.

(1) Support value decomposition.

The following is the decomposition method of image SVT,

$$\begin{cases} S_j = SV_j * P_j \\ P_{j+1} = P_j - S_j, \quad j = 1, 2, \ldots, r \\ P_1 = P \end{cases} \tag{1}$$

where r refers to the decomposition levels. SV_j is the series of support value filters, commonly with Gaussian radial basis function kernel as its initial filter, which are obtained by filling zeros in the basic support value filter. S_j represents the support value image sequences. P_j implies low-frequency components images.

(2) Multi-scale top-hat decomposition.

It is achieved by using the opening-and-closing operation of the mathematical morphology,

$$P_{r+1} \circ D_N(x, y) = ((P_{r+1} \Theta D_N) \oplus D_N)(x, y) \tag{2}$$

$$P_{r+1} \bullet D_N(x, y) = ((P_{r+1} \oplus D_N) \Theta D_N)(x, y) \tag{3}$$

where $\oplus$ refers to dilatation. Θ is erosion. $\circ$ implies opening operation. $\bullet$ represents closing operation. (x, y) stands for position of the pixel and P_{r+1} denotes the last layer of low-frequency image got in Eq. (1). $D_N = \underbrace{(D_b \oplus D_b \oplus \cdots \oplus D_b)}_{N-1}$, construction method of the multi-scale operator. In which, D_N is the construction operator of