Infrared Physics & Technology 78 (2016) 133–146

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Regular article

Infrared and multi-type images fusion algorithm based on contrast pyramid transform

Hua Xu, Yan Wang, Yujing Wu, Yunsheng Qian *

Institute of Electronic Engineering and Optoelectronic Technology, Nanjing University of Science and Technology, Xiao Ling Wei 200, Nanjing, Jiangsu 210094, China

HIGHLIGHTS

- Laplacian sharpening is introduced to enhance the image contrast.
- Otsu method combined with improved weighted fusion is employed to fuse the low-frequency images.
- Top-bottom hats transform is utilized to highlight the bright and dark details of objects and preserve the salient features.
- The second fusion process further makes the thermal targets notable and edges legible.
- Four groups of multi-type images are tested and both qualitative and quantitative evaluations are given.

ARTICLE INFO

Article history:
Received 29 April 2016
Revised 19 July 2016
Accepted 19 July 2016
Available online 20 July 2016

Keywords:
Multi-type images fusion
Contrast pyramid transform (CPT)
Otsu method
Morphology

ABSTRACT

A fusion algorithm for infrared and multi-type images based on contrast pyramid transform (CPT) combined with Otsu method and morphology is proposed in this paper. Firstly, two sharpened images are combined to the first fused image based on information entropy weighted scheme. Afterwards, two enhanced images and the first fused one are decomposed into a series of images with different dimensions and spatial frequencies. To the low-frequency layer, the Otsu method is applied to calculate the optimal segmentation threshold of the first fused image, which is subsequently used to determine the pixel values in top layer fused image. With respect to the high-frequency layers, the top-bottom hats morphological transform is employed to each layer before maximum selection criterion. Finally, the series of decomposed images are reconstructed and then superposed with the enhanced image processed by morphological gradient operation as a second fusion to get the final fusion image. Infrared and visible images fusion, infrared and low-light-level (LLL) images fusion, infrared intensity and infrared polarization images fusion, and multi-focus images fusion are discussed in this paper. Both experimental results and objective metrics demonstrate the effectiveness and superiority of the proposed algorithm over the conventional ones used to compare.

1. Introduction

Multi-sensor and multi-spectral image fusion technologies have been widely focused and adopted in many fields such as military surveillance [1], medical diagnosis [2], remote sensing imaging [3], and target detection and object recognition [4] for its significant value and practical meaning. Spectrum is classified into infrared, visible and ultraviolet spectroscopy in accordance with different wavelength regions, thus have distinct imaging principles and characteristics [5]. In practical applications, the infrared and visible or LLL images fusion has become a hot spot in image fusion research. Infrared image can recognize the thermal or concealed targets in the target scene without difficulties and weather condition limitations, but is insensitive to the variations of brightness, which usually leads to a poor contrast and a deficiency of details [6]. Compared with infrared image, visible image is capable of providing a higher contrast and richer detail information, but usually suffers from a terrible quality at night or low irradiance. LLL image has an advantage in performing the textures and contours information when the illumination is dim while has a difficulty in identifying hidden goals. Infrared polarization image is able to reflect the inherent characteristics of objects, and therefore is liable to find out the objects with different materials and offers more messages than the light intensity one does. In order to acquire an improved image with greater accuracy, higher

* Corresponding author.
E-mail address: yshqian@mail.njust.edu.cn (Y. Qian).

http://dx.doi.org/10.1016/j.infrared.2016.07.016
1350-4495/

reliability and superior visual effect than could be obtained by any single source image, the fusion technology is applied to import the complementary and comprehensive information from all source images into a composite one, which can remarkably enhance people's comprehension and analysis of the target scene.

The weighted average (WA), principal component analysis (PCA) [7], intensity-hue-saturation (IHS) transform [8] fusion method, etc., which are all based on spatial domain, have the merits of briefness, convenience and efficiency. The drawbacks are obvious, however, some valuable information is weakened, and details of objects are seriously lost. Therefore, the result image is usually not desirable enough for us. Multi-scale transform (MST) fusion method, which is based on transform domain, mainly includes Laplacian pyramid transform (LPT) [9,10], CPT [5,11], discrete wavelet transform (DWT) [12,13], contourlet transform (CT) [14,15], etc., is receiving increased attention in image fusion field. Generally, MST-based fusion method is composed of three procedures: firstly, a certain MST is employed to decompose each source image into a sequence of images with high frequency at different scales and directions and one low frequency image. Secondly, the series of images are separately fused in their own spatial frequency via certain fusion rules. Finally, inverse MST (IMST) is performed on the fused series of images to obtain the final fusion image. Therefore, there are two factors, namely MST method and fusion rules, that affect the fusion performance. The CPT fusion method is proposed by Toet based on human vision system (HVS) on the theoretical basis that people's vision is sensitive to local contrast. Because each decomposition level from the contrast pyramid reflects the contrast information in corresponding spatial frequency, the result image can achieve a preferable effect closer to human visual characteristics. To achieve a satisfying result, fusion rule must be designed scrupulously. Scholars have proposed many fusion rules to obtain integrated images, and some of them are proved effective and practical in certain situations. However, most of them are suitable for a certain kind of images, for example, the infrared and visible images [5,11,16], the infrared and LLL images [6,17], remote sensing images [3,18], multi-focus images [13,19] and so forth.

In the meantime, morphology can extract contour features and highlight texture details of targets, thus is widely used in image fusion applications in recent years [20–22]. Through morphological processing, the thermal targets in infrared source image can be strengthened, which will contribute to contrast advancement in final fusion image.

In light of these conditions, we put forward a new fusion strategy based on CPT combined with Otsu algorithm and morphology, which can preserve the detailed information effectively and extract more meaningful information from source images. The scheme is suitable for all the types mentioned above, especially the infrared and multi-type images, and we mainly discuss the fusion applications between infrared and visible images, infrared and LLL images, infrared intensity and infrared polarization images, and multi-focus images in this paper.

2. The proposed algorithm and theory

The overall flow diagram of the proposed algorithm is shown in Fig. 1. Firstly, two input source images should be converted to grayscale ones and then implement median filtering and Laplacian sharpening operations. Secondly, two enhanced images are combined weighted according to the information entropy to obtain the first fused image. Thirdly, two enhanced images and the first fused image are decomposed to a sequence of images with multi-scale and multi-resolution based on CPT method. Afterwards, an optimal threshold T is obtained to segment the top decomposition layer of the first fused image based on Otsu method. Then the top decomposition layer is fused according to the comparison result between the pixel value of the first fused decomposition layer and threshold T. And the other decomposition layers are fused by selecting the absolute maximum of pixels after morphological top-bottom hats operations. Then a fused image is recovered by using the inverse CPT (ICPT). Finally, the final fusion image is obtained by a second fusion with the morphological gradient outcome of the enhanced infrared image. The theories of each step will be introduced in detail in the remaining section.

2.1. Laplacian sharpening

As can be seen in the flow diagram, the Laplacian sharpening is firstly employed to two source images after grayscale conversion and median filtering to get two enhanced images. Consequently, the information of source images is strengthened and the contrast is improved enormously.

Laplace operator is a second differential operation with properties of linearity, isotropy and shift invariance. The Laplace transform of a binary function of image $f(x,y)$ is defined as

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \tag{1}$$

The difference form is

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y) \tag{2}$$

And it can also be written as a template form

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{3}$$

In order to get the sharpened image, the change presented above needs to be superimposed to the original image pixel, i.e., to make substraction operation between the original image and the image after Laplacian filtering. Then the template can be represented as

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{4}$$

2.2. The first fusion

The two enhanced images are combined weighted according to information entropy to obtain the first fused image, which will give an instruction in the following fusion process.

Information entropy (IE) [23,24] is a vital factor to represent the information richness of an image in image processing applications. Shannon proposed the use of entropy as a means of quantifying information content in his paper *A Mathematical Theory of Communication*, and the theory was based on probabilism of information contents. Suppose the probability of certain gray level i in an image is p_i, and the maximum gray level is L, then the information entropy can be calculated by

$$IE = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{5}$$

Generally, the larger the entropy value, the more information the image contains.

If the weighted coefficients of infrared and another enhanced image are α and β respectively, then they can be determined as