Infrared Physics & Technology 60 (2013) 81–93

Contents lists available at SciVerse ScienceDirect

Infrared Physics & Technology

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

Multiscale top-hat selection transform based infrared and visual image fusion with emphasis on extracting regions of interest

Xiangzhi Bai [a,b,c,*], Xiaowu Chen [b], Fugen Zhou [a,c], Zhaoying Liu [a], Bindang Xue [a]

[a] Image Processing Center, Beijing University of Aeronautics and Astronautics, 100191 Beijing, China
[b] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China
[c] Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, China

HIGHLIGHTS

- A multiscale top-hat selection transform based algorithm is proposed.
- Top-hat selection transform well differentiates and extracts the regions of interest.
- The algorithm appropriately combines the extracted useful image information.•The algorithm produces a clear result contained more useful image information.

ARTICLE INFO

Article history:
Received 29 September 2012
Available online 19 March 2013

Keywords:
Infrared and visual image fusion
Top-hat selection transform
Multiscale morphology
Region extraction
Mathematical morphology

ABSTRACT

To effectively combine regions of interest in original infrared and visual images, an adaptively weighted infrared and visual image fusion algorithm is developed based on the multiscale top-hat selection transform. First, the multiscale top-hat selection transform using multiscale structuring elements with increasing sizes is discussed. Second, the image regions of the original infrared and visual images at each scale are extracted by using the multiscale top-hat selection transform. Third, the final fusion regions are constructed from the extracted multiscale image regions. Finally, the final fusion regions are combined into a base image calculated from the original images to form the final fusion result. The combination of the final fusion regions uses the adaptive weight strategy, and the weights are adaptively obtained based on the importance of the extracted features. In the paper, we compare seven image fusion methods: wavelet pyramid algorithm (WP), shift invariant discrete wavelet transform algorithm (SIDWT), Laplacian pyramid algorithm (LP), morphological pyramid algorithm (MP), multiscale morphology based algorithm (MSM), center-surround top-hat transform based algorithm (CSTHT), and the proposed multiscale top-hat selection transform based algorithm. These seven methods are compared over five different publicly available image sets using three metrics of spatial frequency, mean gradient, and Q. The results show that the proposed algorithm is effective and may be useful for the applications related to the infrared and visual image fusion.

1. Introduction

Fusion of image information from multiple imaging sensors is an important technique in different applications [1–4]. Especially, fusion of images obtained from infrared and visual imaging sensors is important to improve the performance of some military or civil applications, such as image based guidance, security surveillance and targeting [3,4]. Infrared (IR) images contain important image regions which could not be displayed by the visual image. However, a visual image usually contains more image details than an infrared image. So, to obtain an image which contains the important image regions and more image details, the technique of infrared and visual image fusion is an important way to achieve this purpose. Usually, the used original infrared and visual images should be already registered. To obtain a good fusion result image with more useful image information, different algorithms have been proposed. Direct average algorithm is widely used in different applications [5,6]. But, this algorithm smoothes some important regions in the infrared image and image details in the visual image, which produces an un-clear result. The wavelet pyramid (WP) and curvelet transforms decompose the original images into different images representing the multiscale features of the original image, and these multiscale features are used for image fusion [6–13]. But, some important image details may also be smoothed, which will affect the performance of these algorithms. Although the shift invariant discrete wavelet transform algorithm (SIDWT) [10,11], which is the improved wavelet pyramid algorithm, performs better

* Corresponding author at: Image Processing Center, Beijing University of Aeronautics and Astronautics, 100191 Beijing, China. Tel.: +86 10 82338578.
E-mail address: jackybxz@buaa.edu.cn (X. Bai).

1350-4495/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.infrared.2013.03.002

for detail protection, some image details may be still smoothed and the result image may be not clear. Laplacian pyramid algorithm (LP) [6,11,14,15] could also extract the multiscale image features for image fusion. Although the result image is clear because some edge features may be strengthened, some image details may be smoothed and the contrast of the result image may be not good. Independent component analysis (ICA) or principal component analysis (PCA) based algorithms extract the main information of the original images and combine the extracted information together to form the final fusion result [16–18]. However, some detailed information which represents important image regions may not be preserved in the final fusion result. Neural networks are also applied for image fusion, which are mainly used for the multi-focus image fusion [19,20].

Mathematical morphology based algorithms are effective ways for image fusion [5,21–25]. These algorithms use morphological operations to extract the useful image features and combine these image features together to form the final fusion image. Among these morphological operations used for feature extraction, top-hat transforms show promising results. Through the multiscale extension of the classical top-hat transform by using the multiscale morphological theory (MSM), image features in the original images may be well extracted and combined into the final fusion image [5]. However, because the performance of opening and closing in the classical top-hat transform may be not good for the extraction of the regions of interest [22], the image details in the fusion result are not clear. To improve the performance of top-hat transform through constructing structuring elements, the center-surround top-hat transform based algorithm (CSTHT) is proposed [23]. Although it is effective for obtaining a fusion result with good contrast [23], some image details are smoothed. This may make the final result un-clear. Also, extracting the focus regions from the multiscale features by top-hat transform performs well for multi-focus image fusion [25]. The algorithm may be not effective for infrared and visual image fusion. Moreover, by utilizing the morphological operations in the pyramid decomposition theory, the constructed morphological pyramid algorithm (MP) could also extract the multiscale pyramid features for image fusion [21]. However, because of the sampling, the morphological pyramid algorithm may produce some artifacts in the result image. This will affect the application of the fusion result.

Actually, the regions of interest in infrared images are usually bright or dim regions comparing with the surrounding regions. Also, the useful image details in visual image are image regions which are different from the surrounding regions. These regions of interest in the infrared and visual images usually represent the useful and important image regions which are different from the surrounding regions. And, these regions of interest are the useful and important image information which should be combined into the final fusion image. So, one important part of infrared and visual image fusion is effectively extracting the important regions of interest which are different from the surrounding regions. The top-hat selection transform [24], which is the modification of the classical top-hat transform, could selectively output image regions following different application purposes. So, top-hat selection transform may be well used for the extraction of the regions of interest in infrared and visual images. Then, based on top-hat selection transform, an effective infrared and visual image fusion algorithm may be constructed.

Through extracting the regions of interest, a multiscale top-hat selection transform based infrared and visual image fusion algorithm is proposed in this paper. First, specifying the top-hat selection transform to extract the regions of interest and constructing multiscale top-hat selection transform using multiscale structuring elements with increasing sizes are discussed. Second, the multiscale image regions of the original infrared and visual images are extracted. Third, the regions of interest represented by the final fusion regions are constructed from the extracted multiscale image regions. Finally, the final fusion image is obtained through importing the extracted regions of interest into a base image using the weight strategy. Because of the effectively extracting the regions of interest using the specified top-hat selection transform, the proposed algorithm performs well for infrared and visual image fusion. Experimental results on infrared and visual images verified the good performance of the proposed algorithm.

2. Top-hat selection transform

2.1. Mathematical morphology

Mathematical morphology has been an important theory for image processing and pattern recognition after being proposed [21]. Morphological operations are mainly based on set theory. And, the two basic operations are dilation and erosion which use two sets. One set is the image and the other set is called structuring element.

Let f and B represent the grayscale image and structuring element, respectively. The dilation and erosion of $f(x, y)$ by $B(u, v)$, represented by $f \oplus B$ and $f \ominus B$, are given as follows,

$$f \oplus B = \max_{u,v} (f(x-u, y-v) + B(u, v)) \tag{1}$$

$$f \ominus B = \min_{u,v} (f(x+u, y+v) - B(u, v)) \tag{2}$$

(x, y) and (u, v) are the pixel coordinates of f and B, respectively.

Through combining the dilation and erosion, the opening and closing of $f(x, y)$ by $B(u, v)$, represented by $f \circ B$ and $f \bullet B$, are as follows,

$$f \circ B = (f \ominus B) \oplus B, \tag{3}$$

$$f \bullet B = (f \oplus B) \ominus B. \tag{4}$$

By comparing the result of opening or closing with the original image, the white top-hat transform and black top-hat transform of image f, represented by WTH and BTH, are as follows,

$$WTH(x,y) = f(x,y) - f \circ B(x,y), \tag{5}$$

$$BTH(x,y) = f \bullet B(x,y) - f(x,y). \tag{6}$$

Opening is usually used to smooth bright image regions whereas closing is usually used to smooth dim image regions. So, WTH is used to extract bright image regions and BTH is used to extract dim image regions.

2.2. Top-hat selection transform

Opening smoothes bright image regions and closing smoothes dim image regions. So, $f(x, y) \geqslant f \circ B(x, y)$, $f(x, y) \leqslant f \bullet B(x, y)$. Take WTH as an example, because $f(x, y) \geqslant f \circ B(x, y)$, WTH could be recognized as a selective output, which could be rewritten as follows [24].

$$WTH(x,y) = \begin{cases} f(x,y) - f \circ B(x,y), & \text{if } f(x,y) - f \circ B(x,y) \geqslant 0 \\ k, & \text{else}, \end{cases} \tag{7}$$

where k is an arbitrary value.

Because top-hat transform does not well utilize the difference information between the regions of interest and the surrounding regions, the performance of top-hat transform is not effective. Through expanding and controlling the difference information between the original image and the result of opening or closing following expression (7), the performance of top-hat transform