Signal Processing 128 (2016) 474–493

Contents lists available at ScienceDirect

Signal Processing

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

Performance improvement scheme of multifocus image fusion derived by difference images

Huafeng Li, Xinkun Liu, Zhengtao Yu*, Yafei Zhang

Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Yunnan, Kunming 650500, PR China

ARTICLE INFO

Article history:
Received 7 December 2015
Received in revised form
26 March 2016
Accepted 16 May 2016
Available online 24 May 2016

Keywords:
Multifocus image fusion
Difference image
Updating mechanism

ABSTRACT

In multiscale transform (MST)-based multifocus image fusion, the fusion rules of different subbands are a significant factor that affects the fusion performance. However, dependence only on new fusion rule will see no significant performance gain for a MST-based method. To address this problem, this paper proposes two novel multifocus image fusion techniques based on multi-scale and multi-direction neighbor distance (MMND), in which the improvements of the fusion performance are respectively achieved by two new developed updating schemes. These two schemes are constructed according to the fact that the difference between a low quality fused result and the source image in the focused region is sharper than those generated by a high quality fused result. Based on this fact, the pixels of the source images are classified into three types in the updating mechanism, namely, pixels of focused significant regions, pixels of smooth regions, pixels of transition area between the focused and defocused regions. According to the categories of source images pixels, we can update the fused result produced by the MMND method in spatial and the MMND domain. Extensive experimental results validate that the proposed two fusion schemes can achieve better results than some state-of-the art algorithms.

1. Introduction

In recent years, multifocus image fusion has received considerable researchers' attention due to the effectiveness in overcoming the limited focus depth of the optical lens in imaging camera. Owing to this technology, one can easily obtain an image that contains all relevant objects in focus, and a significant number of methods about this technology have been proposed [1–12] in recent years. These methods can be roughly classified into two categories: transform domain-based and spatial domain-based techniques [11]. The spatial domain-based methods construct the fused result directly by selecting the pixels, blocks or regions from the focused regions of the source images. Therefore, the quantitative assessment results of these methods are usually larger than those obtained from the MST-based fusion methods. In general, this kind of method includes region-based [13], block-based [14] and focused region detection-based methods [15–17].

For the region-based methods, the basic idea is to apply a segmentation algorithm to the source images, and then a sharpness measure is employed to identify the focused regions. After that the fused image is established by restructuring these focused regions. However, the performance of this approach is highly dependent on the segmentation algorithm, which is very complex and time-consuming. Moreover, discontinuities or erroneous results may arise along the boundary of the regions. In block-based methods, the source images are first divided into blocks, and then a clarity measure is employed to decide which block is from the focused regions. This kind of approach usually has low computational-complexity. But, it is very sensitive to the size of the blocks, and if the block size is not suitable, the blocking effect, which significantly affects the visual perception of the fused results, may appear in the fused result. Although some methods [18,19] have been developed to address this problem, the blockness cannot be deleted completely due to the intrinsic ambiguity arising from the textureless and edgeless regions.

Relative to the above methods, the focused region detection-based methods can effectively avoid the disadvantages discussed above. So they have received more and more attention in recent years. For example, Zhang and Lu et al. [4] proposed to use Graph-based visual saliency algorithm to locate the focused region in the source image; based on depth extraction with inhomogeneous diffusion equation, the source images were divided into clear regions, fuzzy regions and transition regions [20]; In [21], the focused region is detected by comparing the difference between the original log spectrum of the source image and its smoothed version; Moreover, in [22], Zhou and Li et al. used a novel multi-scale weighted gradient approach to detect the definite focused region; Li and Kang et al. [23] applied image matting to obtain the focused

* Corresponding author.
E-mail address: ztyu@hotmail.com (Z. Yu).

http://dx.doi.org/10.1016/j.sigpro.2016.05.015
0165-1684/

region; Liu et al. [24] employed the dense scale invariant feature transform to detect the focused regions. However, the fusion performances of these methods depend on the detection methods which usually yield incorrect solutions due to the complexity of image content. In addition, between focused and defocused regions, there is no clear and definite boundary for many images. As a result, when the pixels located at the transition area are mismanaged, the boundary seams may present along the boundaries of the focused regions in fused result.

Compared with the spatial domain-based method, the MST-based method suffers from little such deficiencies. Therefore, this kind of method is still very popular nowadays, and many effective methods have been developed [12,25–29]. In this type of method, the MST is an important factor to the qualities of the fused results. The popular MSTs include Laplacian pyramid [30], the discrete wavelet transform (DWT) [31], stationary wavelet transform (SWT) [32], lifting stationary wavelet transform (LSWT) [16], dual tree complex wavelet [25,33,34], contourlet transform (CT) [35] and nonsubsample contourlet transform (NSCT) [36]. Among them, the NSCT can generally achieve excellent performance in multifocus image fusion because it possesses the multi-scale, localization, multi-direction, anisotropy and shift-invariance. But then Zhao et al. found that the multi-scale neighbor distance-based method can achieve a better result than the NSCT-based method [7].

Apart from the MSTs, another important component of MST-based algorithms is the fusion rule. In almost MST-based methods, the key point of research is to design the fusion rules for different subbands, and many effective methods have been developed in different MST domains. These methods include pixel-based [37], local neighborhood-based [2,37–42] and local region-based selection principles [37,38]. In general, the local region-based and local neighborhood-based methods outperform the pixel-based methods because the relationship between pixels in a local region and the region-character of object are all considered. In addition to above selection principles, Liu et al. [43] proposed a general framework based on sparse representation (SR) in MST domains; Yu et al. [25] proposed a new method by combining the dual-tree complex wavelet transform and support vector machine; Redondo et al. [44] developed a new selection principle based on multi-size windows technique in log-Gabor domain.

Although some reasonable fusion rules have been presented in different MST domains, the improvement of fusion performance is not so obvious. The reason is that some coefficients corresponding to the important features in the source images cannot be correctly identified due to the intrinsic ambiguity of image content. Studies show that the difference between a low quality fusion result and the focused regions of its source images is sharper than those generated by a high-quality fused result. Based on this conclusion, we developed two novel simple schemes in MMND domain to make up for the defects derived from MST-based methods.

In this scheme, we first use the "averaging" scheme to produce the low-quality fused result. Then, we utilize the MMND as the MST to generate the intermediate product which is considered as the initial value of the updated result. In this step, to achieve the multi-scale decomposition, the neighbor distance is iteratively used, and then the nonsubsampled direction filter banks (NDFB) are employed to achieve multi-direction decomposition, meanwhile, to achieve a fusion result whose visual effect is better than the first fused result, the lowpass subbands and the highpass subbands are respectively merged by 'averaging' scheme and a new 'intensity of pixel gray-values change' derived scheme. After that, the transition fused image (TFI), namely, the intermediate product is constructed by taking an inverse MMND on the fused coefficients. In order to improve the fusion performance, two updated versions of the intermediate product are reconstructed according to two new updating mechanisms which are established based on the properties of different differences originated from these fused results (the low-quality fused result and the intermediate fused result) and their source image. The contributions of this paper and the advantages of our proposed method are summarized as follows:

(1) Development of a new focus measure that combines the idea of major operator and the relationship of pixels located in the same local region, which can properly assess the focusing properties of the highpass subbands at different directions.

(2) Design of a novel method for generating the updating decision maps according to the fact that the clearer the fused result, the nearer the difference between focused regions of the source image and the fused result approximate to zero. This design gives full consideration to the pixels located in the smooth regions, which can effectively improve the accuracy of identification of pixel focusing properties.

(3) Two novel effective updating mechanisms are respectively constructed in spatial domain and MMND domain based on the updating decision maps. These mechanisms can significantly improve the fusion performance of the transition fusion method because they integrate the advantages of MST domain-based and spatial domain-based methods in an efficient manner.

The rest of this paper is organized as follows: Section 2 introduces the theory of MMND in brief. In Section 3, it introduces the fusion method to produce the transition image and the updating mechanisms for the transition image. Our experimental results with some discussion and performance analysis are presented in Section 4. Finally, we give our conclusions in Section 5.

2. Multi-scale and multi-direction neighbor distance filter

In a smooth surface, the oriented distance between two given points can measure the degree of bending in the given direction at a particular point. Therefore, the oriented distance can be employed to extract the details of nformation of an image if we restore the smooth image surface from its discrete version. Moreover, the MST based on the oriented distance outperforms the NSCT in fusion of multifocus images [7]. Therefore, for multifocus image fusion, this MST would be a better choice.

2.1. Multi-scale neighbor distance

In this subsection, we will briefly review the theory of the local neighbor distances developed by Zhou et al. in [7] which will be utilized in combination with NDFB to produce the transition fused result. Firstly, we assume that the smooth image surface have been restored from a given digital gray image and denoted by $S = z(x, y)$. The oriented distance (OD) [7] from point (x,y) to $(x + \Delta x, y + \Delta y)$ is given by

$$OD((x, y), (x + \Delta x, y + \Delta y)) = \frac{1}{2}(l_{xx}(x, y)\Delta x\Delta y + 2l_{xy}(x, y)\Delta x\Delta y + l_{yy}(x, y)\Delta x\Delta y). \quad (1)$$

In Eq. (1), $l_{\mathbf{xy}}$, $(\mathbf{x}, \mathbf{y} = x, y)$ is defined as

$$l_{\mathbf{xy}}(x, y) = z_{\mathbf{xy}}(x, y)/\sqrt{1 + z_x^2(x, y) + z_y^2(x, y)} \quad (2)$$

Let Ω be a local neighborhood centered at (x,y), then the neighbor distance (ND) of (x,y) in Ω is defined by

$$ND(x, y) = \iint_{\Omega} OD((x, y), (x_0, y_0))dx_0dy_0. \quad (3)$$

If we denote by $\{(x_k, y_k)\}_{k=1}^{8} = \{(x + i, y + j): i, j = 0, \pm 1\}$ the eight neighbors around (x,y), the neighbor distance of point (x,y)