Optik 127 (2016) 1291–1296

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Multi-focus image fusion based on cartoon-texture image decomposition

Yongxin Zhang[a,b], Li Chen[a,*], Zhihua Zhao[a], Jian Jia[c]

[a] School of Information Science and Technology, Northwest University, Xi'an 710127, China
[b] School of Information Technology, Luoyang Normal University, Luoyang 471022, China
[c] Department of Mathematics, Northwest University, Xi'an 710127, China

ARTICLE INFO

Article history:
Received 2 January 2015
Accepted 12 October 2015

Keywords:
Image fusion
Split bregman algorithm
Total variation
Multi-component fusion

ABSTRACT

In order to represent the source images effectively and completely, a multi-component fusion approach is proposed for multi-focus image fusion. The registered source images are decomposed into cartoon and texture components by using cartoon-texture image decomposition. The significant features are selected from the cartoon and texture components, respectively to form a composite feature space. The local features that represent the salient information of the source images are integrated to construct the fused image. Experimental results demonstrate the proposed approach works better in extracting the focused regions and improving the fusion quality compared to the other existing fusion methods in both spatial and transform domain.

1. Introduction

Multi-focus image fusion can be defined as a process of combing substantial information from multiple images of the same scene to create a single composite image that will be more suitable for human visual perception or further computer processing [1]. It has been proven to be an effective way to extend the depth of the field. In general, the fusion methods can be categorized into two groups: spatial domain fusion and transform domain fusion [2]. In this paper, we concentrate on the spatial domain methods.

The spatial domain methods are easy to implement and have low computational complexity. The spatial domain fusion methods can be divided into two categories: pixel based methods and region based methods. The simplest pixel based fusion method is to take the average of the source images pixel by pixel. However, the simplicity may reduce the contrast of the fused image. To improve the quality of fused image, some region based methods have been proposed to combine partitioned blocks or segmented regions based on their sharpness [3]. The sharpness is measured by using local spatial features [4], such as energy of image gradient (EOG) and spatial frequency (SF). Then, the focused blocks or regions are selected from source images by simply copying them into the fused image. However, if the size of blocks is too small, the block selection is sensitive to noise and is subject to incorrect selection of blocks from the corresponding source images. Or else, if the size of blocks is too large, the in-focus and out-of-focus pixels are partitioned in the same block, which are selected to build the final fused image. Accordingly, the blocking artifacts are produced and may compromise the quality of the final fused image. To eliminate the blocking artifacts, researchers have proposed some improved schemes. Li et al. [5,6] have selected the focused blocks by using learning based methods, such as artificial neural networks (ANN) and support vector machine (SVM). Due to the difficulty in obtaining empirical data in most multi-focus image fusion cases, the learning based methods are not widely used. Fedorov et al. [7] have selected the best focus by titling source images with overlapping neighbourhoods and improved the visual quality of the fused image. But this method is afflicted by temporal and geometric distortions between images. Aslantas et al. [8] have selected the optimal block-size by using differential evolution algorithm and enhanced the self-adaptation of the fusion method. But this method requires longer computational time. Wu et al. [9] have selected the focused patches from the source images by using a belief propagation algorithm. But the algorithm is complicated and time-consuming. Goshtasby et al. [10] have detected the focused blocks by computing the weight sum of the blocks. The iterative procedure is time-consuming. De et al. [11] have determined the optimal block-size by using quad tree structure and effectively solved the problem of determining of block-size. These schemes all achieve better performance than the traditional methods and significantly inhibit the blocking artifacts.

* Corresponding author. Tel.: +86 18637930027.
E-mail addresses: tabo126@126.com (Y. Zhang), chenli@nwu.edu.cn (L. Chen), jiajian@nwu.edu.cn (J. Jia).

http://dx.doi.org/10.1016/j.ijleo.2015.10.098

In order to effectively and completely represent the source images, a novel fusion method based on cartoon-texture image decomposition is proposed. Cartoon-texture image decomposition is an important way of image processing, which has been widely used in image analysis and vision applications, such as enhancement, inpainting, segmentation, texture and shape analysis [12]. Cartoon-texture image decomposition separates a given image into cartoon and texture components. The cartoon component holds the geometric structures, isophotes and smooth-piece of the source images, while the texture component contains textures, oscillating patterns, fine details and noise [13]. The cartoon and texture components represent the most meaningful information of the source images, which is important for image fusion. Cartoon-texture image decomposition has been proven to be an effective way to extract the structure information and texture information from image [14]. The objective of this paper is to investigate the potential application of cartoon-texture image decomposition in the multi-focus image fusion. The main contribution of this paper is that a multi-component fusion framework is established. The framework is based on the discriminative features that computed from the cartoon and texture components of the source images. The focused regions are detected by computing the salient feature of the neighbourhood region of each pixel in the cartoon and texture components. The proposed method works well in inhibiting the blocking artifacts and representing the source images.

The rest of the paper is organized as follows. In Section 2, the basic idea of cartoon-texture image decomposition will be briefly described, followed by the new method based on cartoon-texture image decomposition for image fusion in Section 3. In Section 4, extensive simulations are performed to evaluate the performance of the proposed method. In addition, several experimental results are presented and discussed. Finally, concluding remarks are drawn in Section 5.

2. Cartoon-texture image decomposition

Nowadays, in many problems of image analysis [15], an observed image f represents a real scene. The image f may contain texture or noise. In order to extract the most meaningful information from f, most models [16–25] try to find another image u, “close” to f, such that u is a cartoon or simplification of f. These models assume that the following relation between f and u:

$$f = u + v \tag{1}$$

where v is noise or texture.

In 1989, Mumford et al. [16] have established a model to decompose the black and white static image by using bounded variation function. In 1992, Rudin et al. [17] have simplified the Mumford–Shah model and proposed total variation minimization energy functional model of Rudin–Osher–Fatemi (ROF) as:

$$E_{\mathrm{ROF}}(u) = \int\int_R (||\nabla u||\mathrm{d}x\mathrm{d}y) + \lambda \int\int_R (u - u_0)^2 \mathrm{d}x\mathrm{d}y \tag{2}$$

The ROF model is very efficient for de-noising images while keeping sharp edges. However, the ROF model will remove the texture when λ is small enough [18]. In 2002, Vese et al. [19] have developed a partial differential equation (PDE) based iterative numerical algorithm to approximate Meyer's weaker norm $||\cdot||_G$ by using L^p. However, this model is time consuming. To improve the computation efficiency, many models and methods have been proposed. They proposed Osher–Sole–Vese (OSV) [20] model based on total variation (TV) and norm H^{-1}. Aujol et al. [21] have introduced dual norm to image decomposition. Chana et al. [22] have proposed $CEP - H^{-1}$ model based on OSV. However, these methods are still complicated. In 2008, Goldstein et al. [23] have proposed Split Bregman algorithm by combining the split method [24] with Bregman iteration [25]. This algorithm is easy to implement and has low computational complexity. This paper performs cartoon-texture image decomposition based on ROF model by using Split Bregman algorithm.

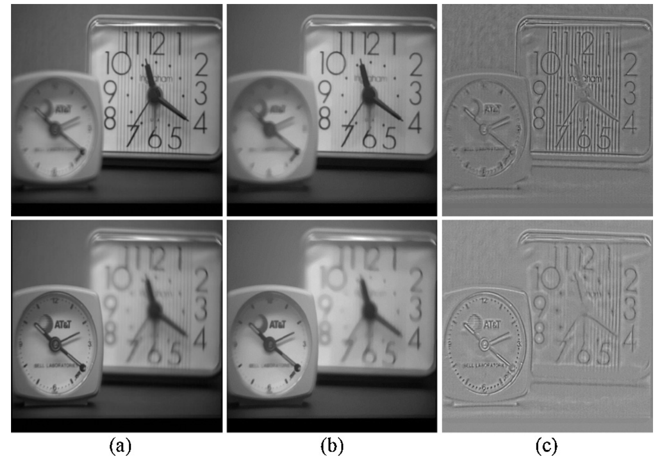

Fig. 1. Cartoon-texture decomposition results of multi-focus image ‘Clock’ using image decomposition: (a) source images *I*. (b) cartoon component *U* and (c) texture component *V*.

Fig. 1(b and c) shows the cartoon-texture image decomposition results of source images ‘Clock’. It is obvious that the salient features of the cartoon and texture components are corresponding to the local feature of objects in focus. Thus, the cartoon and texture components can be useful to build a robust fusion scheme to accurately discriminate the focused regions from defocused regions. In this paper, the salient features of the cartoon and texture components are used to detect the focused regions.

3. Fusion method based on cartoon-texture image decomposition

3.1. Fusion algorithm

In this Section, a novel method based on image decomposition is proposed. The source images must be initially decomposed into cartoon and texture components, respectively. Then, both components are integrated according to certain fusion rules, respectively. The proposed fusion framework is depicted in Fig. 2 and the detailed design is described as follows. For simplicity, this paper assumes that there are only two source images, namely I_A and I_B, here. The rationale behind the proposed scheme applies to the fusion of more than two multi-focus images. The source images are assumed to pre-registered and the image registration is not included in the framework. The fusion algorithm consists of the following 3 steps:

Step 1: Perform cartoon-texture image decomposition on the source images I_A, I_B to obtain cartoon and texture components, respectively. For the source image I_A, let U_A, V_A denote the cartoon and texture components, respectively. For the source image I_B, U_B, V_B have the roles similar to U_A and V_A. Moreover, the color source

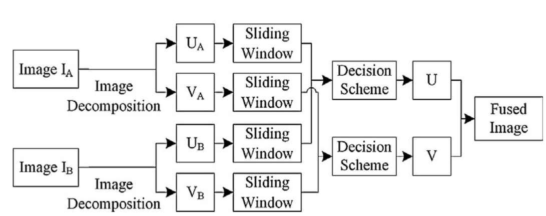

Fig. 2. Block diagram of proposed multi-focus images fusion framework.