Computers and Electrical Engineering 000 (2017) 1–17

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Multi-focus image fusion based on optimal defocus estimation☆

Veysel Aslantas, Ahmet Nusret Toprak*

Erciyes University, Faculty of Engineering, Department of Computer Engineering 38039 Kayseri, Turkey

ARTICLE INFO

Article history:
Received 2 May 2016
Revised 2 February 2017
Accepted 2 February 2017
Available online xxx

Keywords:
Multi-focus image fusion
Extended depth of field
Defocus estimation

ABSTRACT

One of the main drawbacks of the imaging systems is limited depth of field which prevents them from obtaining an all-in-focus image of the environment. This paper presents an efficient, pixel-based multi-focus image fusion method which generates an all-in-focus image by combining the images that are acquired from the same point of view with different focus settings. The proposed method first estimates the point spread function of each source image by utilizing the Levenberg–Marquardt algorithm. Then, artificially blurred versions of the source images are computed by convolving them with the estimated point spread functions. Fusion map is computed by making use of both the source and the artificially blurred images. At last, the fusion map is improved by morphological operators. Experimental results show that the proposed method is computationally competitive with the state-of-the-art methods and outperforms them in terms of both visual and quantitative metric evaluations.

1. Introduction

Multi-focus image fusion is a significant challenge in computer vision and image analysis areas, which aims to create an all-in-focus image by combining the images of the same scene with varying focus settings [1]. The obtained fused image provides more comprehensive information about the scene and it is more suitable for the human/machine perception than each of the source images [2]. Furthermore, analyzing a single fused image is easier than analyzing a collection of images. Therefore, a single fused image is preferred to a series of images with different focal points. Multi-focus image fusion methods are applied to a broad range of applications such as microscopic imaging, industrial vision systems, macro photography, and so on [3,4].

In the aforementioned applications, it is desirable to have the entire image in focus. However, in practice, cameras that are used in imaging systems are not pinhole devices but composed of convex lenses. Such a camera that composed of convex lenses has a limited depth of field (DoF). It can precisely focus on only one plane at a time. Therefore, images of the objects that located at any other plane, are blurred by an amount depending on their distances to the plane of focus. However, if the amount of blur is sufficiently small, it is nearly indistinguishable from a sharp object. Thus, an area of acceptable sharpness occurs between two planes on either side of the plane of focus. This area is named as the depth of field (DoF). To overcome this limitation, two or more images of the same scene are captured from the same point of view with different focus settings. Then these images are combined to create a single composite image that provides the desired depth of field.

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Area Editor Dr. E. Cabal-Yepez.

* Corresponding author.
E-mail address: antoprak@erciyes.edu.tr (A.N. Toprak).

http://dx.doi.org/10.1016/j.compeleceng.2017.02.003

An effective multi-focus image fusion method must meet the following requirements: Firstly, it should preserve all relevant and salient information of the source images. Secondly, fusion process should not produce artifacts. Finally, it should be robust to imperfections such as noise and misregistration [1].

Recently, a large number of multi-focus image fusion methods have been presented. These methods could be classified into two classes: transform-domain and spatial-domain methods. In the transform-domain methods, transform coefficients are obtained by first applying a multi-scale transform to the source images. Then, these coefficients are fused according to particular fusion strategies. Finally, the fused image is obtained by performing an inverse transform over the fused coefficient. Many kind of transforms have been used for image fusion such as Laplacian Pyramid [5] and Discrete Wavelet Transform (DWT) [6]. Another wavelet-based approach have been proposed in [7] that uses variance and spatial frequency of the wavelet coefficients to produce the fused coefficients. Recently developed multi-scale transforms such as Nonsubsampled Shearlet Transform [8], and Discrete Cosine Transform (DCT) [9] have been also employed to multi-focus image fusion. The main drawback of the transform-domain methods is that applying transform prior to the fusion process modifies the original pixel values. Since the fusion is carried out on the transform coefficients, original pixel values cannot be preserved in the transform-domain methods. This may result in loss of information, and color and brightness distortions on the fused image [10].

By contrast with the former class methods, spatial-domain methods fuse source images directly by using their spatial features. Spatial-domain techniques could be basically classified into two categories, which are region-based and pixel-based techniques. In the former category, source images are initially segmented into regions or fixed-sized blocks. Then, the corresponding regions are compared by using sharpness criteria to determine which regions are the sharp ones. At last, sharpest regions are selected to form the fused image. If any of the regions contain the projection of the objects that located at different distances from the plane of focus, these regions will be imaged partially blurred. In this case, the fused image will also contain blurred parts. Therefore, in region-based methods, prior segmentation algorithm plays a crucial role. Some of the well-known region-based methods can be listed as follows: block-based method [11], and block-based method using quadtree approach [12]. Similarly, the method based on ensemble individual features [13] first divides the images into blocks. Then it calculates different feature information that is extracted from both spatial and transform domains of each block to construct the fused image by selecting blocks in a winner-takes-all manner. In addition to these, several optimization-based methods are also proposed in literature to enhance the performance of the region-based methods [14].

The main principle of the pixel-based methods is to select pixels with the maximum sharpness value e.g. PCNN based [15] and dense scale invariant feature transform (DSIFT) [16] based methods. In these methods, a sharpness criterion is computed for each pixel by taking into account a particular neighborhood around that pixel. Then, fused image is obtained by transferring sharpest pixels. On the other hand, some spatial-domain methods create the fused image by using the weighted average of the corresponding pixel values [1,17]. However, averaging pixel values often causes halo artifacts, reduced contrast and reduced sharpness.

In recent times, a pixel-based method that provides satisfactory fusion performance compared to the classical ones, is introduced [18]. In this method, first, the point spread functions (PSF) of each source image is estimated analytically. Then the sharp pixels are detected by using the estimated PSFs. In order to estimate the PSFs, the source images are segmented into fixed sized blocks and then the sharpness values of these blocks are computed by using a predefined sharpness criterion. At last, the PSFs of the source images are obtained by taking average of the PSFs that are calculated over the completely sharp and blurred block pairs.

To compute the PSFs, entirely blurred and sharp regions are to be needed. However, in the multi-focus image fusion applications, the source images are not entirely blurred or sharp but consist of both of these regions. Detecting entirely blurred and sharp region pairs in which PSFs can be computed is a challenging task. Besides, each blurred objects on the images may have a different degree of blur since the distance between the objects and the plane of focus can be different. In this situation, each object could have a different PSF. Considering this; using an optimal PSF instead of an average PSF for each image would be more reasonable.

Based on the aforementioned analysis and motivations, we then propose an efficient pixel-based multi-focus image fusion method using Levenberg-Marquardt algorithm. Levenberg-Marquardt algorithm is utilized to determine the PSFs of each source image. Then, the sharp pixels of the source images are detected by using estimated PSFs to create the fusion map. Some post processes are carried out on the fusion map to enhance the fusion performance. Finally, the fused image is obtained by transferring the detected sharp pixels.

The main contributions of the paper are summarized below:

(1) An optimal defocus estimation method is proposed to estimate the defocus blur of the multiple focused input images.
(2) An improved and efficient algorithm is presented to fuse the multiple focused images which uses the estimated defocus blur of input images.
(3) Several experiments are conducted on real multi-focus image sets. Experiments demonstrate the performance of the proposed technique and provide comparison with the previous methods.

The remainder of the paper is organized as follows. In Section 2 the proposed method is explained in detail. In Section 3 experimental results and discussions are given. With Section 4, the paper is concluded.