Optics Communications 338 (2015) 138–144

Contents lists available at ScienceDirect

Optics Communications

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

High quality multi-focus image fusion using self-similarity and depth information

Di Guo [a], Jingwen Yan [b], Xiaobo Qu [c,*]

[a] School of Computer and Information Engineering, Fujian Provincial University Key Laboratory of Internet of Things Application Technology, Xiamen University of Technology, Xiamen 361024, China
[b] Guangdong Provincial Key Laboratory of Digital Signal and Image Processing Techniques, Department of Electronics Engineering, Shantou University, Shantou 515063, China
[c] Department of Electronic Science, Fujian Provincial Key Laboratory of Plasma and Magnetic Resonance Research, Xiamen University, Xiamen 361005, China

ARTICLE INFO

Article history:
Received 27 August 2014
Received in revised form
5 October 2014
Accepted 15 October 2014
Available online 22 October 2014

Keywords:
Image fusion
Self-similarity
Adaptive
Depth

ABSTRACT

Due to the limited depth of field in a camera, some imaging objects will be blurred if they are located far from the focus plane and the other objects on the plane will be clear. Multi-focus image fusion synthesizes a sharp image from multiple partially focused images. However, traditional fused images usually suffer from blurring effects and pixel distortions. In this paper, we explore two unique characteristics of multi-focus images: (1) The self-similarity of a single image and the shared similarity among multiple source images; (2) The distances from object to focal plane. The former characteristic is used to identify image structure-driven regions while the latter refine the image clarity by automatically estimating depth information of blurred images. Experimental results demonstrate that the proposed method outperforms the state-of-the-art fusion methods on image quality and objective fusion criteria.

1. Introduction

Focusing is important for acquiring a clear image in photography. However, images will be blurred if an object lies far from the focal plane. Therefore, it will be hard to acquire a clear image for a camera with limited depth length. Multi-focus image fusion technology can provide a clearer and more reliable image by combining the information from multiple images focusing on different objects of the same or similar scene [1–3]. It has been applied in non-diffracting imaging system [4], mobile microscope processing software [5] and fluorescence intraoperative surgery [6] in biophotonics.

Existing fusion methods can be roughly divided into two types, which perform the fusion in the spatial domain [7–11] or multi-scale decomposition (MSD) domains [1,12], using sparse transforms, e.g. wavelets [1,13,14], bandelets [15], contourlets [16,17], shearlets [18], surfacelets [19], trained dictionaries [3,20,21]. Although many MSD methods produce nice images, they lead to pixel distortion (Fig. 1(c) and (e)) due to nonlinear operations in the MSD domain. Making use of the coefficients statics in MSD domain helps to suppress this limitation but needs to train parameters in statics models comprehensively [22]. On the contrary, much less distortions are introduced if linear fusion rules in the spatial domain, e.g. the maximum rule, directly choose the pixel values in a well-focused region (Fig. 1(d) and (f)). But the shape of a region will seriously affect fused images. For example, the commonly used isotropic square regions in the spatial domain will easily lead to blocky artifacts in blurred areas around the edge (Fig. 1(d) and (f)). How to choose image pixels in an adaptive region with high clarity remains open.

In this paper, a new multi-focus image fusion method in the spatial domain is proposed. A data-driven scheme, based on the shared similarity of source images, is proposed to generate adaptive regions. Moreover, the distance from an object to the focal plane is automatically estimated. This prior information can hopefully improve the fusion performance since the distance will seriously affect clarity of images. Our contributions are summarized as follows:

1) Adaptive regions based on shared similarity of source images are generated to measure the clarity, thus will potentially avoid the blurring on one object and make more reliable decisions for fusion.
2) A weighting fusion rule based on adaptive regions voting is proposed to further suppress the blurring artifacts.

* Corresponding author.
E-mail address: quxiaobo@xmu.edu.cn (X. Qu).

http://dx.doi.org/10.1016/j.optcom.2014.10.031
0030-4018/

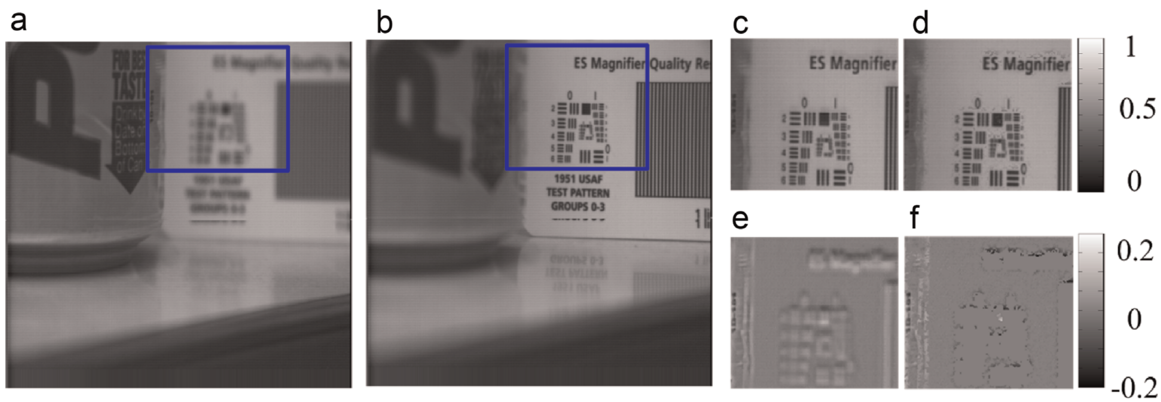

Fig. 1. Artifacts produced by the wavelet and the traditional spatial method. (a) and (b) are source images, (c) and (d) are cropped parts of fused images using wavelet and typical spatial method with isotropic regions, (e) and (f) are fusion errors compared with the sharp image regions in (b). Note: Some-of-Modified-Laplacian (SML) is applied in the spatial method.

3) The distances between multiple objects and the focal plane are incorporated in adaptive regions to measure the clarity.

2. Method

The outline of the proposed method is illustrated in Fig. 2. First, adaptive regions are generated using image similarity information. Next, the depth information is estimated, and combined with the image gradient to measure the clarity in adaptive regions. Finally, a pixel is fused based on the weightings from source images.

2.1. Shared similarity regions

The adaptive region is generated using the shared similarity of source images. Images are first divided into multiple overlapped square patches, and similar patches are searched [23,24]. The self-similarity is very effective for medical image reconstruction [24].

Here, a shared self-similarity of source images is defined as follows to generate an adaptive region of pixels for fusion. Given a reference patch $P_r \in \mathbb{R}^{m\times m}$ and a region $R(r) \in \mathbb{R}^{n\times n}$ centered at pixel r, the similarity of any candidate patch $P_q \in \mathbb{R}^{m\times m}$ to the P_r is defined as

$$\eta_q = \|P_q - P_r\|_F, \tag{1}$$

where $\|A\|_F = \sqrt{\sum_{i=1}^{I}\sum_{j=1}^{J}|a_{ij}|^2}$ denotes the Frobenius norm of matrix A. By sorting the η_q by the descending order for all the patches in this region, the most k similar patches to P_r are found and the collection of this patches are expressed as $L_R(r) = \left\{P_{q_1}, P_{q_2}, \cdots, P_{q_k}\right\}$. Similar patches shared by both L_{R^A} and L_{R^B} are

$$L_R^S(r) = L_{R^A}(r) \cap L_{R^B}(r) \tag{2}$$

where $L_R^S(r)$ is one of the adaptive regions for fusion, $R^A(r)$ and $R^B(r)$ denote the same region of source images f^A and f^B, respectively. The locations and the number of similar patches of each adaptive region vary with the shared similarity of source images. For example, the adaptive region is composed of 4 similar patches as shown in Fig. 3.

How to make use of these adaptive regions in image fusion? Traditionally, a pixel of the fused image is chosen from a source image with higher clarity, which is called the maximum rule [1,7], as follows:

$$f^F(r) = \begin{cases} f^A(r) & \text{if } C^A(r) > C^B(r) \\ f^B(r) & \text{otherwise} \end{cases} \tag{3}$$

where C denotes the clarity metric and f^F is the fused image. In order to overcome the blocky artifacts in fusion, instead of selecting the pixel with a larger value, $C(r)$ is usually computed as a sum of pixel-wise charity measure, e.g. gradients like Some-of-Modified-Laplacian (SML) in a fixed size of region centered at

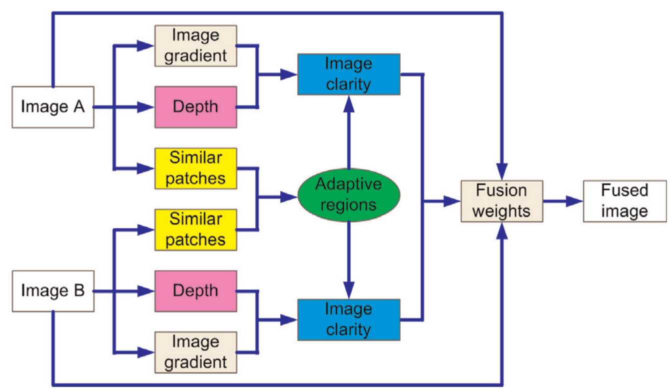

Fig. 2. Flowchart of the proposed method.

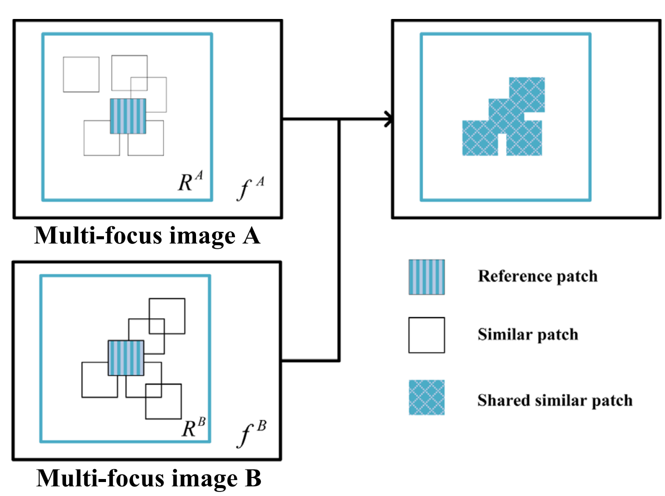

Fig. 3. An adaptive region of shared similar patches.