Digital Signal Processing 60 (2017) 200–210

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

All-in-focus imaging using average filter-based relative focus measure

Oh-Jin Kwon, Seungcheol Choi *, Dukhyun Jang, Hee-Suk Pang

Department of Electrical Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-Gu, Seoul 05006, Republic of Korea

ARTICLE INFO

Article history:
Available online 20 September 2016

Keywords:
Multi-focus
All-in-focus
Automatic focus bracketing
Image fusion
Digital camera

ABSTRACT

Digital images are normally taken by focusing on an object, resulting in defocused background regions. A popular approach to produce an all-in-focus image without defocused regions is to capture several input images at varying focus settings, and then fuse them into an image using offline image processing software. This paper describes an all-in-focus imaging method that can operate on digital cameras. The proposed method consists of an automatic focus-bracketing algorithm that determines at which focuses to capture images and an image-fusion algorithm that computes a high-quality all-in-focus image. While most previous methods use the focus measure calculated independently for each input image, the proposed method calculates the relative focus measure between a pair of input images. We note that a well-focused region in an image shows better contrast, sharpness, and details than the corresponding region that is defocused in another image. Based on the observation that the average filtered version of a well-focused region in an image shows a higher correlation to the corresponding defocused region in another image than the original well-focused version, a new focus measure is proposed. Experimental results of various sample image sequences show the superiority of the proposed measure in terms of both objective and subjective evaluation and the proposed method allows the user to capture all-in-focus images directly on their digital camera without using offline image processing software.

1. Introduction

Many photographers are often frustrated when they try to capture a picture of an object in the foreground. When the foreground object is in focus, the background is out of focus, and vice versa. This is due to the limited depth-of-field of cameras. A common solution to this problem is to capture images at various focuses and fuse them together into an all-in-focus (AIF) image, where all visible objects are in focus [1–23].

A variety of multi-focus image fusion methods have been developed. So far, these methods are mainly classified into two types, one of which is used in transform domain and the other in spatial domain [1–3]. Transform domain methods transform the original source images into transform coefficients, apply a unique fusion rule, and reconstruct the fused image by performing the inverse transform. Most commonly used transforms are multi-scale based transforms. They include gradient pyramid transform [4], Laplacian pyramid transform [5], discrete wavelet transform [6], dual tree complex wavelet transform [7], contourlet transform [8,9], log-Garbor transform [10], shearlet transform [11], curvelet transform [12], and others. Even though these transform-based methods have shown satisfactory fusion results for high quality images, there is a known issue of considering limited spatial information. Therefore, such methods often lose edge and texture information, which leads to distortions in fused images [13–22]. Furthermore, the high computing cost of these methods can be inhibitive.

Spatial domain methods can be divided into two categories: region-based methods [13–18] and pixel-based methods [1,19–22]. In region-based methods, fusion results from two-pass processing. These methods typically employ their own focus measure to segment the source images into three regions: focused regions, defocused regions, and transition regions between focused and defocused regions. Subsequently, different fusion rules are applied for each region. Pixel-based methods, however, use one-pass processing. Therefore, these methods are usually faster than region-based. In pixel-based methods, the focus measure is designed to estimate the focus of each pixel in an image. A focus weighting map is generated for each source image to determine the pixel of the fused image by calculating the weighted average of pixels from source images or by selecting the pixel demonstrating optimal focus among source images.

Various focus measures have been proposed in the literature on the subject, which include standard deviation or variance [13,23], average gradient [13], energy of Laplacian [23], spatial frequency [14,19], energy of gradient [23], Tenenbaum gradient [24], sum-

* Corresponding author.
E-mail addresses: ojkwon@sejong.ac.kr (O.-J. Kwon), choisc@sju.ac.kr (S. Choi), jangdh777@sju.ac.kr (D. Jang), hspang@sejong.ac.kr (H.-S. Pang).

http://dx.doi.org/10.1016/j.dsp.2016.09.005

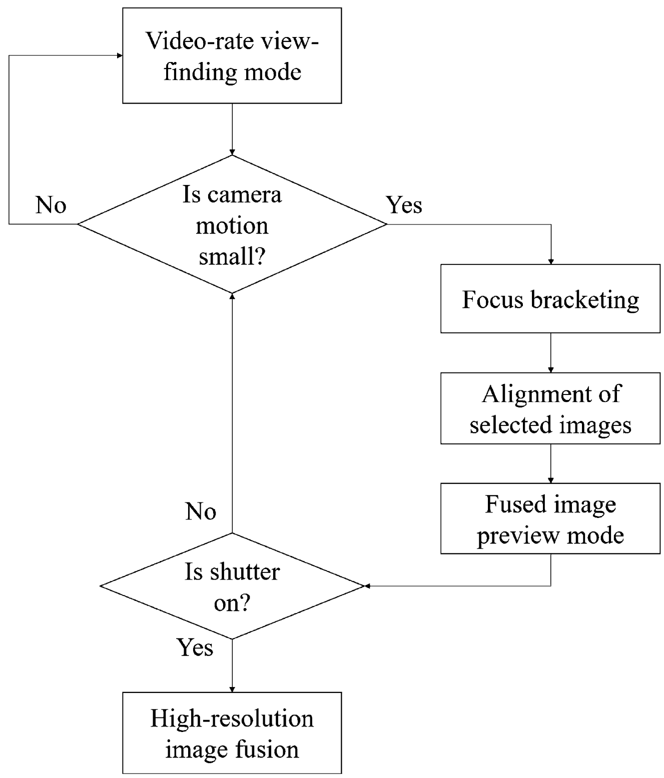

Fig. 1. Exemplified implementation of the proposed AIF imaging method.

modified Laplacian (SML) [25], frequency selective weighted median filter [26], and other measures. Comparisons of traditional focus measures for the multi-focus image fusion are provided in [13,20,23,27]. More recently proposed measures for multi-focus image fusion include multi-scale weighted gradient [15], morphological filtering [17], bilateral gradient-based sharpness [20], homogeneity similarity [21], neighbor distance [28], and other measures.

In this paper, we propose a pixel based AIF imaging method. Professional photographers may create AIF images using popular image editing software in post-processing. However, this offline approach is cumbersome for the ordinary photographer. We propose a method for previewing and capturing AIF scenes directly on digital cameras without using offline image processing software. The proposed method attempts to include the following features:

1) An automatic bracketing algorithm to determine the minimum number of focus settings required to accurately capture AIF scenes.
2) A real-time preview for the user to examine the AIF image and freeze a shot.
3) Direct creation and review of the high-resolution AIF image on the digital camera.
4) A new multi-focus image-fusion algorithm employed for fast processing.

The paper is organized as follows. Section 2 describes the proposed AIF imaging method. Section 3 presents the experimental results. Finally, Section 4 concludes the paper.

2. Proposed AIF imaging method

Fig. 1 exemplifies an implementation of the proposed method. One attractive feature of modern digital cameras is their video-rate view-finding (VRVF) mode. The exemplified implementation automatically starts focus bracketing when camera motion in the VRVF mode is small for a few seconds and shows a fused image in preview resolution size in real time. The camera switches between the VRVF mode and fused image preview (FIP) mode based on detected camera motion. For fast processing, a simple frame difference detecting method detects camera motion by thresholding the sum of the absolute differences between centered blocks of half the horizontal and vertical frame sizes of consecutive frames. A high-resolution fused image is captured at the user's demand in the FIP mode.

In the following subsections, we first propose a new focus measure. We explain the focus-bracketing algorithm that implements the proposed focus measure. Finally, we describe the multi-focus image-fusion algorithm for previewing a reduced sized image in real-time and capturing a high-resolution AIF image.

2.1. Proposed focus measure

The following attributes have been proposed for perceptual image quality assessment: overall brightness, contrast, sharpness, details, naturalness, and colorfulness [29–31]. Researchers have investigated the effect of focus change on these attributes and have found that focus change has little effect on overall brightness, naturalness, and colorfulness [13–22]. Therefore, most focus measures proposed have been designed mainly by utilizing contrast, sharpness, and details. As these attributes may be reflected on the luminance component of an image, the proposed focus measure is calculated based on the luminance value.

Focus measures can be broadly organized into two main categories: measures calculated in the transform domain and the spatial domain [1–3]. Even though transform domain measures result in satisfactory fusion to produce high quality images, their computing cost is too high for real-time processing in state-of-the-art digital cameras. Spatial domain measures in multi-focus image fusion were evaluated by Tian et al. [20] and Huang and Jing [23]. The authors reported that the SML-based measure produced the best results among the real-time implementable measures.

Most spatial domain focus measures estimate the focus of a pixel based on its neighboring pixels. For example, the SML is calculated as follows:

$$f(m,n) = \sum_{(i,j)\in B_{m,n}} \{|2x(i,j) - x(i-1,j) - x(i+1,j)| + |2x(i,j) - x(i,j-1) - x(i,j+1)|\}, \tag{1}$$

where $f(m,n)$ and $x(m,n)$ denote the focus measure and the luminance intensity of a pixel at (m,n), respectively, and $B_{m,n}$ is the block centered at (m,n).

In this paper, we propose a new focus measure that results in better quality fused images than the SML does with comparable computing cost. To the authors' knowledge, all focus measures proposed to date calculate the absolute focus for each source image independently, as exemplified in Eq. (1); the proposed measure calculates the relative focus of an image compared to another. This new approach improves fusion performance.

For two images taken at different focuses, a well-focused region in one image shows better contrast, sharpness, and detail than the corresponding defocused region in the other image. Therefore, average-filtering a well-focused region in one image increases the correlation to the corresponding defocused region in the other image. On the other hand, average-filtering a defocused region in one image decreases its correlation to the corresponding well-focused region in the other image. We verify this observation using an experiment below.

Consider images 1 and 2 denoted by $X^{(1)} = \{x^{(1)}(m,n)\}$ and $X^{(2)} = \{x^{(2)}(m,n)\}$, with average-filtered versions $\overline{X}^{(1)} = \{\overline{x}^{(1)}(m,n)\}$ and $\overline{X}^{(2)} = \{\overline{x}^{(2)}(m,n)\}$, respectively. We define the correlation map between $X^{(1)}$ and $X^{(2)}$ as follows: