Multi-focus Image Fusion for Extended Depth of Field

Wisarut Chantara and Yo-Sung Ho
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology (GIST)
{wisarut,hoyo}@gist.ac.kr

ABSTRACT

In this paper, we address the extended depth of field method, which appropriately applicable to a set of stack images. The proposed fusion algorithm improved a conventional discrete wavelet transform using spatial frequency and sum-modified-Laplacian. Both methods are applied to the approximation coefficient and the detail coefficients (horizontal, vertical and diagonal) for calculating in-focused regions. After that, the sharp areas of each stack images are combined with the fusion image process. Finally, the inverse wavelet transform is utilized to obtain a final result image. The performance of the proposed method is conducted and compared with conventional fusion methods. Experiment results can demonstrate that the proposed method outperforms other reference methods.

CCS CONCEPTS

• **Computing methodologies** → **Image processing**;

KEYWORDS

Depth of field, discrete wavelet transform, spatial frequency, sum-modified-Laplacian

ACM Reference Format:
Wisarut Chantara and Yo-Sung Ho. 2018. Multi-focus Image Fusion for Extended Depth of Field. In *The 10th International Conference on Internet Multimedia Computing and Service(ICIMCS'18), August 17-19,2018, Nanjing, China*. ACM, Nanjing, China, 4 pages. https://doi.org/10.1145/3240876.3240894

1 INTRODUCTION

The depth of field (DOF), also called focus range or effective focus range, is the distance between the nearest and farthest objects in a scene that appears acceptably sharp in an image. Extended DOF creates an image with an extended focal range from a series of images focused at a different depth, particularly useful for computer vision, digital photography, and macro photography. So that multi-focus image fusion is the process which combines two or more different focal image with the same scene into a single image, maintaining the significant information from each of the stack images. Therefore, all important objects in the image are in-focused. Generally, the digital camera can set only one distance focus at a time. Then, to acquire a set of stack images with the same scene by captured in a single shot, a light field camera such as Lytro camera [8] and Raytrix camera [10] provide a set of the stack images with different focal display in the same view.

In recent year, many depth of field extension techniques have been developed, e.g., Li et al. [6], introduced a method based on the selection of image blocks from source images to construct the fused image using spatial frequency. While Desale et al. [2], proposed a study and analysis of image fusion technique by way of PCA, DCT, and DWT. Moreover, Wang et al. [11], presented multisource image fusion using spatial frequency and simplified pulse coupled neural network. In this paper, we proposed an improved discrete wavelet transform using spatial frequency and sum-modified-Laplacian to detect focal regions and reconstruct in-focused regions to an all-in-focused image.

2 PROPOSED METHOD

The following techniques are involved in the proposed methods implementation. Light field camera [8] is utilized as a capture device. Light field picture splitter [9] is applied to split the raw file into different focus level images. This process provided a set of different focal images which display the same position as shown in Fig. 1.

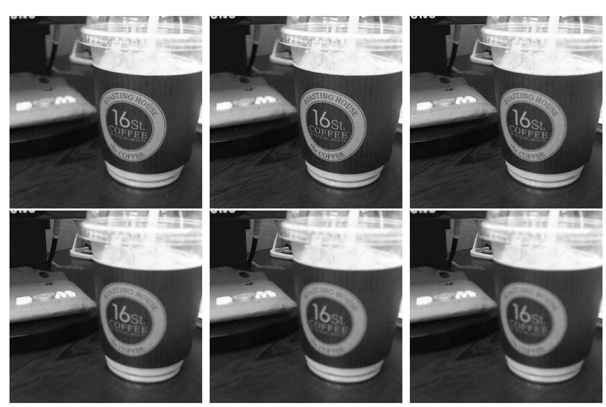

Figure 1: A set of different focal images.

ICIMCS'18, August 17-19,2018, Nanjing, China

ACM ISBN 978-1-4503-6520-8/18/08. . . $15.00
https://doi.org/10.1145/3240876.3240894

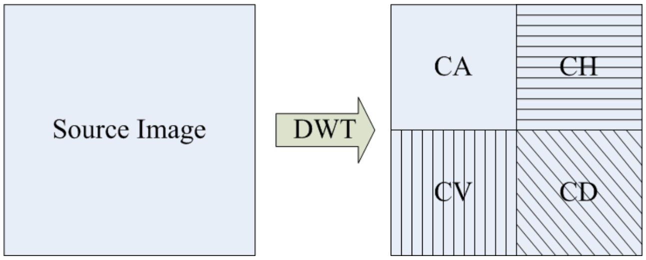

Figure 2: DWT decomposition.

2.1 Discrete wavelet transform (DWT)

Discrete wavelet transform uses a cascade of special low-pass and high-pass filter and a sub-sampling operation. The output from first-order of DWT contains four decomposition parts as shown in Fig. 2. Those are CA, CH, CV and CD, where CA is the approximation coefficient, which is sensitive to human eyes [5]. While CH, CV, and CD are the detail coefficient (horizontal, vertical and diagonal) which have more detail information more than CA. Since DWT of image signals produces a nonredundant image representation, it can provide better spatial and spectral localization of image information.

2.2 Spatial frequency (SF)

Spatial frequency, which originated from the human visual system, indicates the overall active level in an image. It is difficult to completely comprehend the human visual system with current physiologic means. While the use of SF has led to an effective contrast criterion for image fusion [7].
SF is defined as:

$$SF = \sqrt{(CF)^2 + (RF)^2} \tag{1}$$

whereRF and CF are the row frequency and column frequency respectively:

$$RF = \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=2}^{N} [I(x,y) - I(x,y-1)]^2} \tag{2}$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{j=1}^{N} \sum_{i=2}^{M} [I(x,y) - I(x-1,y)]^2} \tag{3}$$

for M x N image with a grayscale $I(x, y)$ at (x, y) coordinate.

2.3 Sum-modified-Laplacian

Sum-modified-Laplacian is developed to compute a local measure of the quality of image focus. SML can provide better performance in focus measurement criterion [4]. SML is defined as:

$$SML(x,y) = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} \nabla^2_{ML} I(i,j) \tag{4}$$

$$for \quad \nabla^2_{ML} I(i,j) \geq T_{SML}$$

$$where \quad \nabla^2_{ML} I(i,j) = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right|$$

2.4 Enhancement SML (eSML)

In the homogeneous region, the focus measure can be affected by the pixel noise [1]. In order to decrease this effect, the SML values are computed in a small window to determine the eSML focus measure of the center pixel (x, y) as:

$$eSML = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} SML(i,j) \tag{5}$$

2.5 Image fusion

The wavelet decomposition coefficient structures (approximation and detail coefficients) are combined as:

$$F_{i,j} = D^m_{i,j} \tag{6}$$

$$where \quad m = \arg\max_t C^t_{i,j}, t = 1, 2, ..., N$$

where $F_{i,j}$ is a final coefficient in an all-focused image at (i, j). While $D^m_{i,j}$ is a maximum coefficient information at (i, j). $C^t_{i,j}$ is a coefficient information at t stack image. N is a set of input stack images.

2.6 The proposed method procedure

A summary of the proposed method is provided as follow:

(1) Apply the DWT on each image of the input stack images,
(2) Apply the SF by using Eq. (1) to (3) for the approximation coefficients (CA),
(3) Apply the SML by using Eq. (4) for the detail coefficients (CH, CV, and CD),
(4) Apply the eSML by using Eq. (5) for the noise reduction improvement in the detail coefficients,
(5) Combine each coefficient of the approximation and detail coefficients by using Eq. (6), and
(6) Apply an inverse discrete wavelet transform to obtain a multi-focus image with extended depth of field as a final image.

3 EXPERIMENT RESULTS

The experiments are conducted to compare the performance of the proposed method with those of popular widely conventional methods (such as pixel averaging method, DWT-averaging method [2], DWT-maximum method [2], and the spatial frequency method). We conducted experiments on 512 × 512 pixels sample images. The test images are a set of cup and stuff images, which contain multiple objects at different focus level images capturing by light filed camera [8]. In all our experiments, we set T_{SML} to be 10. The experiment results are presented in Fig. 3 and 4. From the fusion results, we can easily observe that the results of the pixel averaging and DWT methods have a lower contrast

Table 1: Objective evaluation of the extended depth of field image (non-reference fusion metrics)

Image	Criteria	Pixel averaging	DWT[2] Averaging	DWT[2] Maximum	SF	Proposed method
`Cup`	FMI	0.9227	0.9186	0.9122	0.9262	**0.9264**
	$Q^{AB/F}$	0.6292	0.6159	0.5207	0.6533	**0.6552**
`Stuff`	FMI	0.9221	0.9178	0.9059	0.9228	**0.9230**
	$Q^{AB/F}$	0.6129	0.5837	0.5055	0.6498	**0.6500**

than those of the spatial frequency method and the proposed method. However, it is hard to tell the difference between the results of the spatial frequency method and the proposed method by subjective evaluation. Hence, the paper applies some non-reference fusion metrics such as Feature Mutual Information (FMI) [3] and Petrovics metric ($Q^{AB/F}$) [12] are then introduced and employed. These evaluation criteria metrics are calculated without respect to the reference images. FMI measures the amount of information that the fused image contains the source images, while $Q^{AB/F}$ measures the relative amount of edge information that is transferred from the source into the fused image. The higher the FMI or $Q^{AB/F}$ value, the better the fused image performance. The comparison results are summarized in Table 1.

The above two evaluation criteria are then applied to evaluate the four fusion methods in Fig. 3 and 4, the detailed quantitative results are given in Table 1. From Table 1, we can observe that the values of all quality indices of the proposed method are larger than those of pixel averaging, DWT-averaging, DWT-Maximum, and the conventional spatial frequency methods, which means the proposed algorithm can effectively combine sharp parts of the original image to the fused image, and yield superior quality than conventional methods.

4 CONCLUSIONS

In this paper, we proposed a multi-focus image fusion for extended depth of field imaging. The reconstructed image provides all-focused scene. We apply the enhanced discrete wavelet transform algorithm to measure focus regions and fuse the final image. The main contribution of this work is that we reform the conventional DWT algorithm with the spatial frequency and the sum-modified-Laplacian algorithms. As a result of that, the proposed method has more efficiently than other conventional methods.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIP) (No. 2011-0030079).

REFERENCES

[1] Xiangzhi Bai, Yu Zhang, Fugen Zhou, and Bindang Xue. 2015. Quadtree-based multi-focus image fusion using a weighted focus-measure. *Inf. Fusion* 22 (March 2015), 105–118.
[2] Rajenda Pandit Desale and Sarita V. Verma. 2013. Study and analysis of PCA, DCT & DWT based Image Fusion Techniques. In *Int. Conf. on Signal Processing, Image Processing and Pattern Recognition*. IEEE, Coimbatore, India, 66–69.
[3] Mohammad Haghighat and Masoud Amirkabiri Razian. 2014. Fast-FMI: non-reference image fusion metric. In *Proceedings of 8th Int. Conf. on Application of Information and Communication Technologies*. IEEE, Astana, Kazakhstan, 1–3.
[4] Wei Huang and Zhongliang Jing. 2007. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters* 28, 4 (March 2007), 493–500.
[5] Chu-Hui Lee and Zheng-Wei Zhou. 2012. Comparison of image fusion based on DCT-STD and DWT-STD. In *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2012)*. Hong Kong, 720–725.
[6] Shutao Li, James T. Kwok, and Yaonan Wang. 2001. Combination of images with diverse focuses using the spatial frequency. *Inf. Fusion* 2, 3 (Sept. 2001), 169–176.
[7] Shutao Li and Bin Yang. 2008. Multifocus image fusion using region segmentation and spatial frequency. *Image and Vision Computing* 26, 7 (July 2008), 971–979.
[8] Lytro. [n. d.]. The Lytro camera. http://www.lytro.com
[9] Nirav Patel. 2013. lfptools. http://github.com/nrpatel/lfptools/
[10] Raytrix. [n. d.]. The Raytrix camera. http://www.raytrix.de
[11] Nianyi Wang, Yide Ma, and Weilan Wang. 2014. DWT-based multisource image fusion using spatial frequency and simplified pulse coupled neural network. *Journal of Multimedia* 9, 1 (Jan. 2014), 159–165.
[12] C.S. Xydeas and Vladimir Petrovic. 2000. Objective image fusion performance measure. *Electronics Letters* 36, 4 (Feb. 2000), 308–309.

(a) (b) (c)

(d) (e)

Figure 3: Comparison of "Cup": (a) Pixel Avg., (b) DWT-Avg., (c) DWT-Max, (d) SF, (e) Proposed method.

(a) (b) (c)

(d) (e)

Figure 4: Comparison of "Stuff": (a) Pixel Avg., (b) DWT-Avg., (c) DWT-Max, (d) SF, (e) Proposed method.