# Multi-focus Image Fusion based on Improved DWT for Light Field Camera

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

**Abstract**

In this paper, we address the multi-focus image fusion method, which appropriately applicable to a set of stack images. 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 outperform other reference methods.

## 1. Introduction

Multi-focus image combination is the process by which two or more different focal image with the same scene are fused into a single image, maintaining the significant information from each of the stack images. Therefore, all 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 [1] and Ratrix camera [2] provide a set of the stack images with different focal display in the same view.

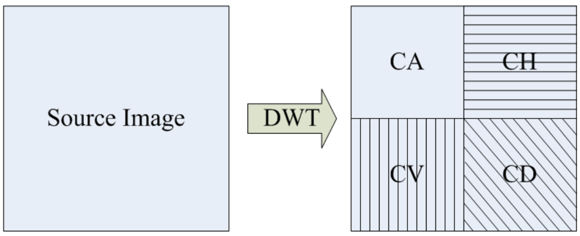
In recent year, many fusion techniques have been developed, e.g., Li et. al. [3], introduced a method based on the selection of image blocks from source images to construct the fused image using spatial frequency. While Desale et. al. [4], proposed a study and analysis of image fusion technique by way of PCA, DCT and DWT. Moreover, Wang et. al. [5], 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-focused image.

## 2. Proposed method

The following techniques are involved in the proposed method's implementation. First, we use the lytro light field camera [1] as a capturing device to get the stack images. The sub-images are the same viewpoint with different focus level [6].

### 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 contain four decomposition parts. Those are CA, CH, CV and CD, where CA is the approximation coefficient, which is sensitively with human eyes [7]. 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.




(Figure 1) DWT Decomposition

### 2.2 Spatial frequency (SF)

Spatial frequency, which originated from the human visual system, indicates the overall active level in an image. The use of SF has led to an effective for image fusion [8].

SF is defined as:

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

Where RF and CF are the row frequency:

$$\mathrm{RF} = \sqrt{\frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=2}^{N}[f(i,j) - f(i,j-1)]^2} \qquad (2)$$

and column frequency:

$$\mathrm{CF} = \sqrt{\frac{1}{M \times N}\sum_{j=1}^{N}\sum_{i=2}^{M}[f(i,j) - f(i-1,j)]^2} \qquad (3)$$

For an *M*×*N* image with gray value $f(i,j)$ at *(i, j)*

2.3 Sum-modified-laplacian (SML)

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 [9]. SML is defined as:

$$\mathrm{F(x,y)} = \sum_{i=x-N}^{i=x+N}\sum_{j=y-N}^{j=y+N} \nabla_{ML}^2 I(i,j)$$
$$for \ \ \nabla_{ML}^2 I(i,j) \geq T \qquad (4)$$
$$where \qquad \nabla_{ML}^2 I(i,j) = \left|\frac{\partial^2 I}{\partial x^2}\right| + \left|\frac{\partial^2 I}{\partial y^2}\right|$$

Where *T* is a threshold value of SML.

2.4 Image fusion process

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

$$F_{i,j} = D_{i,j}^{m}$$
$$\text{where} \qquad (5)$$
$$\mathrm{m} = \underset{t}{\mathrm{argmax}}\{C_{i,j}^{t}\ \}, t = 1,2,\dots,N$$

Where $F_{i,j}$ is a final coefficient in all-focused image at *(i, j)*. While $D_{i,j}^{m}$ is a maximum coefficient information at *(i, j)*. $C_{i,j}^{t}$ is a coefficient information at *t* stack image. N is a set of stack images.

2.5 Summary of the proposed method

A summary of the proposed method is provided as follows:

*Step1:* Apply a discrete wavelet transform on each stack image

*Step2:* For the approximation coefficients, apply the spatial frequency using Eq. (1) to (3). And for the detail coefficients (horizontal, vertical and diagonal), apply the sum-modified-laplacian using Eq. (4).

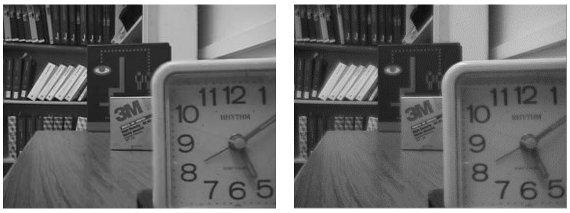
*Step3:* Combine each coefficient of the approximation and detail sub-bands using Eq. (5)

*Step4:* Apply an inverse discrete wavelet transform to obtain an all-focused image.

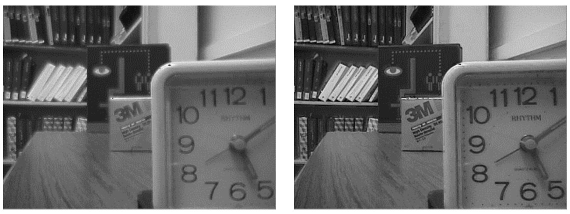
## 3. Experiment results

Experiments are conducted to compare the performance of the proposed method with the conventional method. The first experiment is to conduct image combination using two sets of images with different focus levels. The second experiment is to fuse the stack of light field images. The experiment results are presented in Fig.2 and 3, where one can see that the proposed algorithm can effectively combine sharp parts of the original image to the fused image, and yield superior quality than conventional methods.

(a) Source images



(b) Ref. images (c) Avg. DWT



(d) Max DWT (e) Proposed method

(Figure 2) The experiment images “Clock”

In this paper, the evaluation criteria of Mutual Information (MI) and Root Mean Square Error (RMSE) are used to perform the fusion results. They are expressed as:

$$\mathrm{MI(R;I)} = \sum_{r,i} p(r,i) log \frac{p(r,i)}{p(r)p(i)} \qquad (6)$$

Where $p(r,i)$ is the join probability distribution function of the reference image *(R)* and interested image *(I).*

While $p(r)$ and $p(i)$ are the marginal probability distribution function of $R$ and $I$ respectively. Fusion performance would be better with MI increasing.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_x \sum_y [I(x,y) - R(x,y)]^2}{M \times N}} \qquad (7)$$

Where $I$ and $R$ are the interested image and reference image respectively. $M \times N$ is the image size. The lower RMSE means that the fusion image is close to the actual reference image.

(a) Source images

(b) Ref. images (c) Avg. DWT

(d) Max DWT (e) Proposed method

(Figure 3) The experiment images “Stuffs”

From Table1, we can observe that the performance of the proposed algorithm outperforms other conventional algorithm.

## 4. Conclusion

In this paper, we proposed a method for combination multi-focus images. 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 contrition 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.

(Table 1) Performance Comparison

| **Images** | **Quality measures** | **DWT [4]** | | **Proposed method** |
|---|---|---|---|---|
| | | **Avg.** | **Max** | |
| Fig. 2 (Clock) | MI | 4.116 | 3.445 | 8.794 |
| | RMSE | 12.745 | 17.121 | 11.826 |
| Fig. 3 (Stuffs) | MI | 3.763 | 3.425 | 4.140 |
| | RMSE | 15.354 | 18.452 | 14.545 |

**Acknowledgement**

This research was supported by the ‘Cross-Ministry Giga KOREA Project’ of the Ministry of Science, ICT and Future Planning, Republic of Korea (ROK). [GK15C0100, Development of Interactive and Realistic Massive Giga-Content Technology]

**References**

[1] Lytro, “The lytro camera,” https://www.lytro.com/.

[2] Ratrix, “The ratrix camera,” http://www.raytrix.de/.

[3] S. Li, J. T. Kwok, and Y.Wang, “Combination of images with diverse focuses using the spatial frequency,” Inf. Fusion, Vol.2 (3), pp.395–403, 2001.

[4] R. P. Desale and S. V. Verma, “Study and analysis of PCA, DCT & DWT based Image Fusion Techniques,” Int. Conf. on Signal Processing, Image Processing and Pattern Recognition, pp.66–69, 2013.

[5] N. Wang, Y. Ma, and W. Wang, “DWT-based multisource image fusion using spatial frequency and simplified pulse coupled neural network,” Journal of Multimedia, Vol.9 (1), pp.159–165, 2014.

[6] N. Patel, “lfptools,” http://github.com/nrpatel/fkptools/.

[7] C.H. Lee and Z.W. Zhou, "Comparison of image fusion based on DCT-STD and DWT-STD," Int. Multi-Conf. of Engineers and Computer Scientists, Hong Kong, pp. 720-725, 2012

[8] S. Li and B. Yang, “Multifocus image fusion using region segmentation and spatial frequency,” Image and Vision Computing, Vol.26 (7), pp.971–979, 2008.

[9] W. Huang and Z. Jing, "Evaluation of focus measure in multi-focus image fusion," Pattern Recognition Letters, Vol.28 (4), pp. 493–500, 2007.