A comparison of different focus measures for use in fusion of multi-focus noisy images

V. Aslantas, R. Kurban
Erciyes University, Faculty of Engineering, Department of Computer Engineering, 38039 Melikgazi, Kayseri, Turkey
Email: aslantas@erciyes.edu.tr

ABSTRACT

Impulsive noise (IN) produced by image sensors and/or communication channels corrupts images in many practical applications. This noise may cause miscalculation of sharpness values which, in turn, introduce considerable errors in fused images. In this paper, conventional focus measures and frequency selective weighted median filter (FSWM) based focus measure are evaluated for fusion of multi-focus images in the presence of IN. Experimental results are presented for several sets of images and the results show that FSWM based focus measure can provide better performance than other focus measures.

Keywords: Image reconstruction, Multi-focus image fusion, Focus measures, Impulsive noise

1. Introduction

Recently, intelligent inspection systems have become an important topic in industrial informatics. Image fusion has been widely applied in industrial manufacturing [1-5] areas such as electronic circuitry and component inspection, product surface measurement and inspection and intelligent robots on assembly lines, manufacture process monitoring, etc.

In practice, all cameras used in imaging systems are not pin-hole devices but consist of convex lenses. Therefore, they have a limited depth of field, i.e., when a lens focuses at a particular distance, objects that are in front of or beyond this distance are out of focus and appear blurring [6]. The degree of this blurring is affected by the camera position, the focal length, the *f*-number of the lens and the distance between the lens and the sensor plane [7]. A popular way to extend the depth of field of an optical imaging system is to combine the images of the same scene taken from the identical point of view under different focal settings. This process is named as multi-focus image fusion.

Image fusion produces a sharply focused single image of a scene from partially focused and defocused images of the same scene acquired using either multiple sensors or a sensor whose geometric and sensory parameters can be changed. The sharper parts of the input images are copied into the fused image with the aim of obtaining everywhere in focus image of the scene. For human and machine perception, well-focused images are preferred to out-of-focus images since more information can be acquired from sharply focused images than the defocused images [8,9,10].

Recently, several methods based on multi-scale transforms have been proposed such as Laplacian Pyramid [11] and the discrete wavelet transforms (DWT) [12]. However, multi-resolution approaches are generally shift-variant and sensitive to noise [13]. To overcome this problem, a discrete wavelet frame transform based method has been proposed [9]. However, the implementation of these methods is complicated and the algorithms are time consuming. In addition to that, a spatial domain multi-focus image fusion (SDMIF) method has also been presented [14]. The fundamental idea behind this method is to construct a fused image that encloses the sharper parts of the source images. This process necessitates an analytic focus measure which can be used to evaluate the sharpness of focus in a part of the image. A sharpness criterion should:

a. respond to high-frequency variations in image intensity,
b. be independent from the image content,
c. be computationally efficient for the real-time implementations.

Without considering the noise generated by the CCD of video cameras, the performance of several focus measures was studied with respect to image fusion [14,15]. However, in many practical applications, images are distorted by impulsive noise produced by image sensors and/or communication channels. This noise may cause miscalculation of sharpness values which, in turn, degrade the performance of image fusion. In this paper, conventional focus measures and FSWM based

focus measure are evaluated for multi-focus image fusion in the presence of IN.

2. Focus measures

The general idea behind focus measures is to compute the high frequency content of an image [16,17]. Let $f(i, j)$ be the gray level intensity of the pixel (i, j).

2.1. Variance (VAR)

For an M × N size of an image block, variance can be employed as a focus measure and is computed as:

$$VAR = \frac{1}{M \times N} \sum_i \sum_j (f(i,j) - \overline{f})^2 \quad (1)$$

where $\overline{f}$ is the average grey level over the image region:

$$\overline{f} = \frac{1}{M \times N} \sum_i \sum_j f(i,j)$$

2.2. Energy of laplacian (EOL)

Laplacian of an image can also be utilized as a focus measure that can analyze high frequencies associated with image edges and is calculated as:

$$EOL = \sum_i \sum_j (\nabla^2 f(i,j))^2 \quad (2)$$

The Laplacian of the image intensity can be estimated using any of the Laplacian masks.

2.3. Tenengrad (TNG)

Edge characteristics are affected by the quality of focus. Since defocusing tends to decrease the gradient of an edge, the focus measure involves measuring the magnitude of the gradients in the region of interest. The criterion estimates the gradient at each image point, and simply sums their magnitudes.

$$TNG = \sum_i \sum_j [\nabla f(i,j)]^2 \text{ for } \nabla f(i,j) > T \quad (3)$$

where T is a threshold value and $\nabla f(i, j)$ is the gradient magnitude given as:

$$\nabla f(i,j) = \sqrt{f_x^2 + f_y^2} \quad (4)$$

where f_x and f_y are gradients along the horizontal and vertical direction, respectively. There are many discrete operators which can estimate f_x and f_y. Tenengrad uses Sobel operators to approximate gradients in the horizontal and vertical directions.

2.4. Frequency selective weighted median filter (FSWM):

The high frequency content of an image can be measured by a gradient estimator because it is inherently a high-pass filter. In the course of detailed analysis between gradient estimator and the high-pass filter, the basic structure for FSWM based focus measure is derived in [18]. The FSWM based focus measure not only responds to high frequency components of the images, but also eliminates the effects of impulsive noise. Therefore, it can measure the sharpness of an image more precisely.

The characteristics of the high-pass filter can be improved by a nonlinear weighted median (WM) filter. The WM filter can be represented with $<W; F>$, where $W = [w_1, w_2, \ldots, w_m]$ and $F = [f_1, f_2, \ldots, f_m]$ are the weight vector and the discrete time continuous valued input vector of a WM filter, respectively. The elements of F are arranged in increasing order as $f_1 < f_2 < \ldots < f_m$. The output of the WM filter is computed by repeating each sample f_i to the number of the corresponding weight w_i followed by sorting the resulting array. Then, the median value from the expanded vector is chosen.

For instance, <[1, 2, 3, 2, 1]; [-3, -2, 0, 1, 4]> refers to $median\{f(i\text{-}3), 2 \lozenge f(i\text{-}2), 3 \lozenge f(i), 2 \lozenge f(i\text{+}1), f(i\text{+}4)\}$, where $\lozenge$ is the duplicating operator, i.e., $w \lozenge f$ represents that f is repeated w times.

WM filters can be linearly combined to form an FSWM filter that can be defined as:

$$H = \sum_i a_i < W_i, F_i > \quad (5)$$

An FSWM filter can be obtained by finding appropriate $a_i \in R$, W_i and F_i. In [18], the frequency characteristics of FSWM filters were investigated to find a desirable filter for autofocusing and the following band pass filter is applied for measuring image sharpness:

$$\begin{aligned} H = {} & med\{f(n-1), f(n), f(n+1)\} \\ & - \frac{1}{2} med\{f(n-3), f(n-2), f(n-1)\} \\ & - \frac{1}{2} med\{f(n+1), f(n+2), f(n+3)\} \end{aligned} \quad (6)$$

In this paper, the filter given in eq. (6) can also be employed to measure image sharpness for the purpose of image fusion. Let H_h and H_v, respectively, be FSWM filtering results that are obtained by applying the filter to an image along the horizontal and vertical directions using (6). The sharpness measure can be defined as:

$$FSWM = \sum_i \sum_j \left[H_h^2 + H_v^2\right] \quad (7)$$

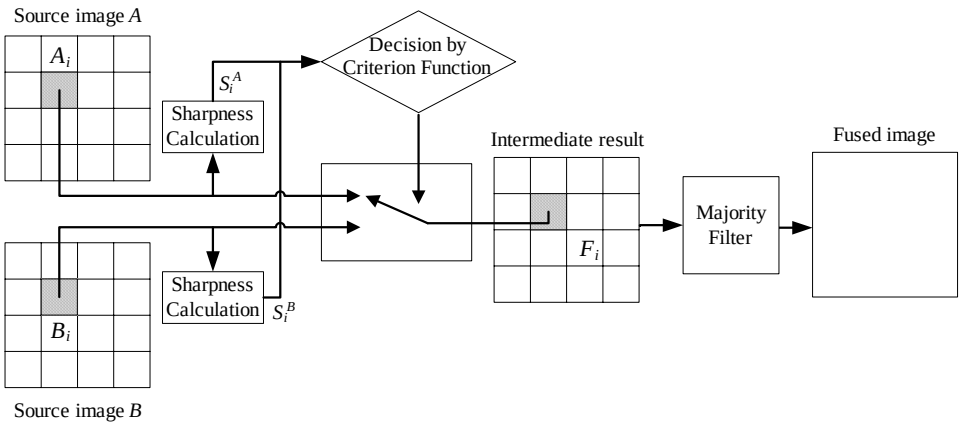

Fig. 1. Schematic diagram of multi-focus image fusion.

3. Multi-focus image fusion

The spatial domain pixel-level multi-focus image fusion algorithm is illustrated in Fig. 1. This work only deals with the fusion of two images. However it can easily be adopted to handle more than two images.

First the algorithm partitions the source images into equi-sized square tiles. As a result, there are no gaps or overlaps between the neighboring sub-image blocks. Then, the sharpness values of each corresponding blocks of the source images are computed. Lastly, by comparing these values, the blocks with the high sharpness values are copied into the corresponding positions in the fused image.

The computation steps of the algorithm can be expressed as the followings:

1. Decompose the input images A and B into equi-sized square blocks. The corresponding image blocks of A and B are referred by A_i and B_i, respectively, where i represents the position of the block.
2. Calculate the sharpness values of A_i and B_i by applying the focus measure and refer the results of A_i and B_i by S_i^A, and S_i^B, respectively.
3. Determine the sharper image block by comparing the sharpness values of the corresponding blocks, and copy the sharper block into the intermediate fused image F_i as:

$$F_i = \begin{cases} A_i, & S_i^A > S_i^B \\ B_i, & otherwise \end{cases} \qquad (8)$$

4. Repeat Steps 2 and 3 for all corresponding blocks of the input images.
5. Compute the final fused image by applying a majority filter on the intermediate fused image. This filter eliminates the isolated blocks. That is, if a particular block in the intermediate fused image is copied from image A but the majority of its surrounding blocks are from image B, this block will be replaced with the corresponding block of B and vice versa. In the implementation, the size of the majority filter employed was 3×3 window.

4. Experimental results

This section presents several experiments conducted on artificially and naturally blurred images for comparing the image fusion performance of focus measures.

The first experiment was carried out on two 256×256 reference (everywhere-in-focus) images. The images (*Kids* and *Sulfur*) are shown in Fig. 2a and Fig. 3a, respectively. The blurred versions of the *Kids* image were then obtained by convolving the regions of the children with a Gaussian of radius 1,4 and distorted by adding different amount of random IN. As an example, two of the resulting noisy (1% IN) and blurred images are shown in Figs. 2c and 2d. In a similar manner, Figs. 3c and 3d were obtained by blurring Fig. 3a with a Gaussian of radius 1 and degraded by 1% random IN.

Peak signal-to-noise ratio (PSNR) is used to evaluate the quality of the artificially produced multi-focus images. PSNR is a metric for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation and is given by:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [R(i,j) - F(i,j)]^2} \right) \qquad (9)$$

where R is the everywhere-in-focus reference image, F is the fused image. The larger the PSNR value, the better the fused image.

Experiments were conducted for evaluating the performance of the focus measures on

image fusion with respect to various block sizes (8×8, 16×16 and 32×32) and different levels of IN (0.5%, 1%, 1.5%, 2% and 2.5%). Simulations were performed 30 times due to the random nature of artificial IN. For *Kids* images (Fig. 2c and 2d), the results of PSNR values obtained by averaging the results of independent runs are given in Table 1. The standard deviation (StdDev) for the PSNR values were also computed showing the variations of the PSNR values with respect to the average PSNR values. Table 1 gives the computed StdDev values of the results.

The same experiments were also carried out for the *Sulfur* images (Fig. 3c and 3d) and the results obtained are given in Table 2. As can be seen from Table 1 and 2, it is clear that FSWM performs better than other methods under various amount of noise and block sizes. All the experiments show that the calculated average PSNR values of FSWM are less than those of other functions. In addition, the standard deviations computed for all the experiments show that the difference between various runs is lower for FSWM than that of the others. These indicate that the FSWM is more reliable, robust and stable than the other focus measures. Examples of fused images obtained by employing FSWM with the block size of 8×8 are given in Fig. 2b and 3b.

The second experiment was performed on the images which were acquired by a real lens. The images were composed of multiple objects at different distances from the camera. Therefore, the images had naturally defocused parts because points on the surface of the world at a particular distance from the lens were focused whereas points at other distances were defocused (or blurred) by varying degrees depending on their distances. In the experiments, 1% artificial impulsive noise was added to the out of focus images. Focus measures were performed on the images shown in Fig. 4a (near focused) and Fig. 4b (far focused) by employing 8×8 block size. The resulting fused images are shown in Figs. 4c through f. The fused images given in Figs. 4c through e have some problematic parts. For example, some areas on the person head and the first monitor are copied from the wrong source images. On the other hand, the fused image by FSWM, as shown in Fig. 4f, is clearer than the results of the other techniques.

5. Conclusion

Almost all focus measures depend directly on the amount of high frequency information in the image. It should be noted that much of the image noise is also related to high frequencies. This dilemma causes miscalculation of the image sharpness and consequently of the fused image. The effects of noise on the measurement of image sharpness should be reduced as much as possible. Efforts to mitigate image distortions before the image fusion level can modify pixel values of the original images and may result in unjustified complication of image fusion algorithms since the decision for fusion has to be done on processed images rather than original images. Therefore, image fusing algorithms should not use filtering operations for noise removal before fusing stage.

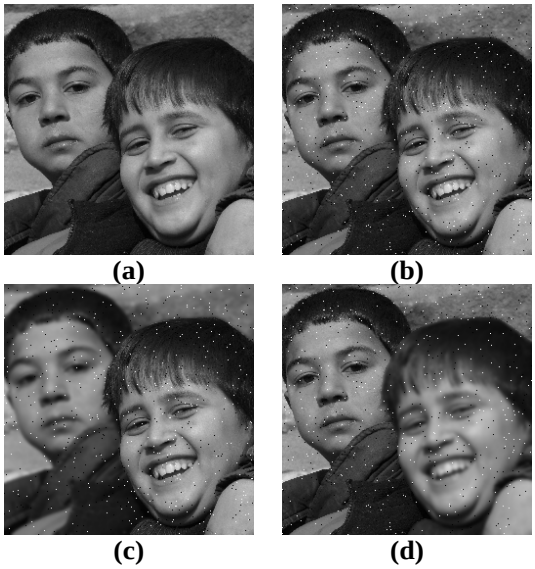

Fig. 2. *Kids* image: (a) reference image, (b) fused image using FSWM, (c) noisy source image (near focused), (d) noisy source image (far focused).

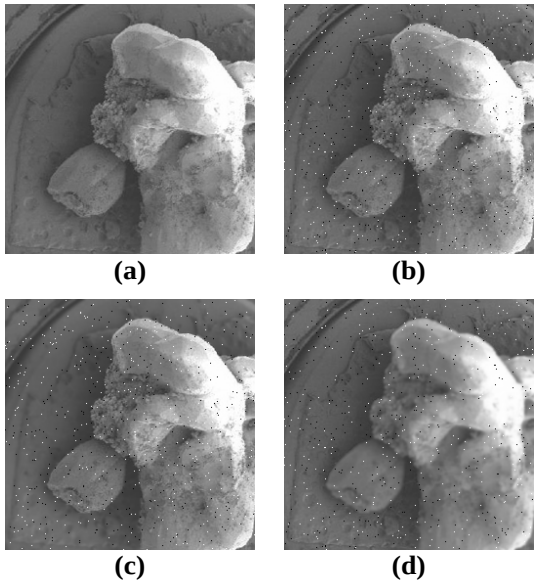

Fig. 3. *Sulfur* image: (a) reference image, (b) fused image using FSWM, (c) noisy source image (near focused), (d) noisy source image (far focused).

Table 1. PSNR results for focus measures after 30 replicates for *Kids* image.

	IN levels	**0.5%**			**1%**			**1.5%**			**2%**		
	Block Size	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**
VAR	Avg.	42.586	40.527	40.203	40.301	39.167	39.761	39.000	38.771	39.118	37.864	38.067	38.784
	StdDev	0.638	0.636	0.969	0.711	0.786	1.089	0.576	0.973	1.076	0.718	0.867	1.208
EOL	Avg.	42.410	37.455	36.267	39.463	35.159	35.692	37.010	34.858	34.833	35.760	33.983	34.698
	StdDev	0.818	1.003	1.318	0.835	0.929	1.588	0.746	0.997	1.072	0.680	0.779	1.369
TNG	Avg.	42.745	40.684	40.196	40.545	39.014	39.698	38.842	38.523	38.786	37.642	37.916	38.549
	StdDev	0.535	0.684	1.002	0.640	0.763	1.113	0.640	0.707	1.050	0.639	1.074	1.291
FSWM	Avg.	*44.470*	*44.344*	*42.613*	*44.337*	*43.944*	*41.965*	*44.374*	*43.341*	*41.271*	*44.057*	*42.936*	*40.413*
	StdDev	*0.135*	*0.106*	*0.657*	*0.257*	*0.543*	*0.829*	*0.291*	*0.577*	*1.017*	*0.474*	*0.750*	*1.182*

Table 2. PSNR results for focus measures after 30 replicates for *Sulfur* image.

	IN levels	**0.5%**			**1%**			**1.5%**			**2%**		
	Block Size	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**
VAR	Avg.	46.886	44.012	45.349	43.462	42.438	44.077	41.759	42.073	43.663	40.881	41.289	42.986
	StdDev	1.169	1.028	1.315	0.893	1.220	1.107	0.953	1.077	1.098	0.788	1.524	1.653
EOL	Avg.	46.474	40.381	40.263	43.177	38.253	39.142	40.732	38.065	39.476	39.474	37.799	38.441
	StdDev	1.128	1.177	1.721	0.988	1.166	1.765	0.940	0.985	1.625	0.603	0.777	1.701
TNG	Avg.	47.368	44.673	45.683	44.380	43.278	44.913	42.241	42.567	43.868	41.214	41.910	43.553
	StdDev	1.086	0.854	1.127	0.772	1.290	1.329	0.796	0.915	1.477	0.866	1.143	1.273
FSWM	Avg.	*51.800*	*51.136*	*49.179*	*51.587*	*50.290*	*48.284*	*51.380*	*49.199*	*47.049*	*50.642*	*48.334*	*45.943*
	StdDev	*0.312*	*0.299*	*0.619*	*0.484*	*0.891*	*1.000*	*0.496*	*1.389*	*1.007*	*0.979*	*1.183*	*1.406*

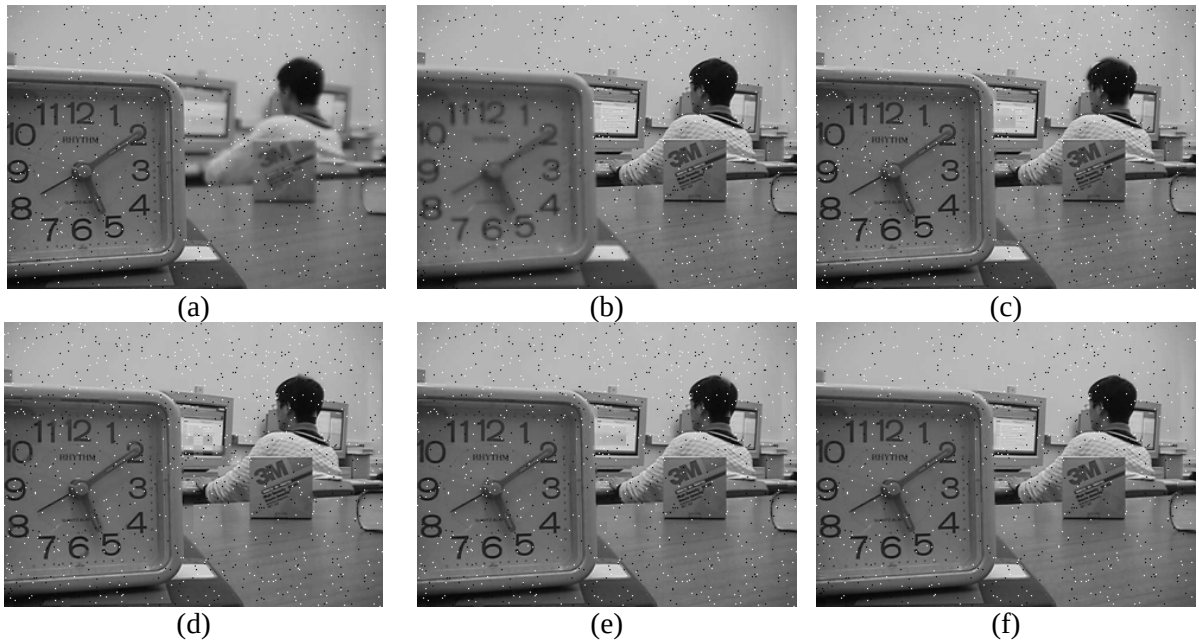

Fig. 4. *Lab* image (320×240) and fusion results under 1% IN: (a) near focused, (b) far focused, (c)-(f) fused images using VAR, EOL, TNG, and FSWM, respectively.

The simulations have been performed with different block sizes and amounts of IN. The experimental results obtained show that FSWM focus measure is less sensitive to noise, exhibits the sharpness of an image more precisely and outperform the other focus measures and fusion methods on the fused noisy images in terms of quantitative and subjective evaluation.

The main objective of this paper is to explore the suitability and effectiveness of the focus measures for the fusion of multi-focus noisy images. The performances of the functions are compared by employing the various block sizes and amount of noise. By employing an optimal block size obtained by

an intelligent optimization technique may improve the performance of the fusion results. Development of an optimization tool to aid in this task is beyond the scope of this work, and could be an interesting future work.

References

[1] G. Piella, “A general framework for multiresolution image fusion:from pixels to regions”, Information Fusion 4, pp. 259–280, 2003.

[2] H.S. Sarraf, and J.S. Goddard, “Vision System for On-Loom Fabric Inspection”, IEEE Trans. on Industry Applications 35, 1999.

[3] R.R. Murphy, “Sensor and information fusion improved vision-based vehicle guidance”, IEEE Intelligent Systems 13, pp. 49–56, 1999.

[4] K.N. Lou, and L.G. Lin, “An intelligent sensor fusion system for tool monitoring on a machining centre”, International Journal of Advanced Manufacturing Technology 13, pp. 556–565, 1997.

[5] J.M. Reed, and S. Hutchinson, “Image fusion and subpixel parameter estimation for automated optical inspection of electronic components”, IEEE Transactions on Industrial Electronics 43, pp. 346–354, 1996.

[6] V. Aslantas and D. T. Pham, “Depth from automatic defocusing” Opt. Express 15, pp. 1011-1023, 2007.

[7] D.T. Pham, and V. Aslantas, “Depth from defocusing using a neural network”, Pattern Recogn. 32, pp. 715-727, 1999.

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

[9] S. Li, J.T. Kwok, I.W. Tsang, and Y. Wang, “Fusing images different focuses using support vector machines”, IEEE T. Neural Network. 15, pp. 1555-1561, 2004.

[10] W.B. Seales, and S. Dutta, “Everywhere-in-focus image fusion using controllable cameras”, Proc. SPIE 2905, pp. 227-234, 1996.

[11] P.T. Burt, and E.H. Andelson, “The Laplacian pyramid as a compact image code”, IEEE T. Commun. 31, pp. 532-540, 1983.

[12] D.A. Yocky, “Image merging and data fusion by means of the discrete two-dimensional wavelet transform”, J. Opt. Soc. Am. A 12, pp. 1834-1841, 1995.

[13] G.H. Qu, D.L. Zhang, and P.F. Yan, “Medical image fusion by wavelet transform modulus maxima”, Opt. Express 9, 184-190, 2001.

[14] S. Li, J.T. Kwok, and Y. Wang, “Combination of images with diverse focuses using the spatial frequency”, Inform. Fusion 2, 169-176, 2001.

[15] W. Huang, and Z. Jing, “Evaluation of focus measures in multi-focus image fusion,” Pattern Recognition Letters, vol. 28, (2007) 493-500.

[16] E.P. Krotkov, “Focusing,” Int. Journal of Computer Vision, vol.1, no.3, (1987) 223-237.

[17] F.C.A. Groen, I.T. Young, and G. Ligthart, “A Comparison of Different Focus Functions for use in Autofocus Algorithms,” Cytometry, vol.6, (1985) 81-91.

[18] K. Choi, J. Lee, and S. Ko, “New autofocusing technique using the frequency selective weighted median filter for video cameras,” IEEE Trans. on Consumer Electronics, vol. 45, (1999) 820-827.