Extending Depth-of-field by Image Fusion using Multi-objective Genetic Algorithm

Veysel Aslantas, Rifat Kurban
Department of Computer Engineering, Faculty of Engineering, Erciyes University, Kayseri, Turkey
{aslantas, rkurban}@erciyes.edu.tr

Abstract- **Optical imaging systems used in industrial vision applications suffer from the limited depth of field problem. Image fusion is a practical way to obtain an everywhere-in-focus image of a scene. In this paper, a spatial domain image fusion method using multi-objective genetic algorithm is proposed to obtain an optimal fused image. Experimental results conducted with industrial images show that the proposed method outperforms Laplacian Pyramid and DWT methods.**

I. INTRODUCTION

In recent years, intelligent inspection systems have become an important topic in industrial informatics. One of the problems related with the imaging systems used in industrial applications is the limited depth of field where objects at different depths cannot be viewed sharply in a single image. [1,2].

Image fusion refers to an image processing task that produces a sharply focused single image of a scene by combining 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 fused image obtained then turns out to be more desirable for human/machine perception [3,4].

In industrial manufacturing, image fusion has been widely used in the areas such as electronic circuitry and component inspection, product surface measurement and inspection, complex machine/device diagnostics and intelligent robots on assembly lines, manufacture process monitoring, etc. [5-10].

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 [11]. 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 [12].

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.

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.

Recently, several methods based on multi-scale transforms have been proposed such as Laplacian Pyramid [13] and the discrete wavelet transforms (DWT) [14]. However, multi-resolution approaches, are generally shift-variant and sensitive to noise [15,16]. In addition to that, a spatial domain multi-focus image fusion (SDMIF) method has also been presented [17]. The fundamental idea behind this method is to construct a fused image by choosing the sharper image blocks within the source images.

SDMIF can avoid the problem of moving objects and misregistration of source images. SDMIF techniques also necessitate the determination of an appropriate block size which significantly affects the fusion results [16,18]. The block size directly depends on the content and defocused regions of the source images. A fixed block size may not be applicable for every application. Therefore, selection of block size inevitably requires heuristic choices. Numerous combinations of block sizes are possible, and it is difficult to find optimal solutions by trial and error. Hence, optimization of the block size is necessary in order to obtain a fused image that comprises the sharper parts of the source images.

The fundamental objective of fusion methods is to obtain a high quality fused image of a scene. Various metrics have been developed for assessing the quality of the fused image. Each has its own advantages and disadvantages. Furthermore, different metrics may be compatible or incompatible with one another. In [19], Genetic Algorithm (GA) is employed to obtain an optimal block size by employing a single metric. However, it may not be appropriate to qualify a fused image based on a single metric. Therefore, quality of a fused image should be evaluated by designing an intelligent hybrid system where two or more metrics are combined in a manner that overcomes the limitations of individual metrics. The common approach to the task is to transform the multi-objective problem into a single objective problem where the objectives are weighted to form a composite objective function. However, relation between the quality metrics is often nonlinear, and the knowledge of appropriate weights for different metrics should be known beforehand. These necessitate the use of multi-objective optimization methods to search the optimal parameters in order to achieve the fused image optimally. The solutions obtained in this way are more adaptive and competitive because they do not depend on the predefined weights. Therefore, adaptive multi-objective particle swarm optimization (AMOPSO) algorithm has been

978-1-4244-3760-3/09/$25.00 © 2009 IEEE

used to optimize the fusion process in DWT domain [20]. However, AMOPSO method requires the reference image in some cases and its performance will deteriorate quickly if there is slight movement or mis-registration of the source images. Another disadvantage of this method is that it requires quite a number of weights to be optimized which, in turn, increases the complexity of the algorithm.

In this paper, a novel multi-objective genetic algorithms (MOGA) approach has been proposed to determine the optimal block size for SDMIF method.

II. Multi-focus Image Fusion using Multi-objective Genetic Algorithm

In this section, first, two well-known objective quality metrics used to measure the performance of image fusion methods are presented. Then, general architecture of a multi-objective genetic algorithm (MOGA) is given. The last subsection explains the proposed spatial domain multi-focus image fusion method.

A. Objective Evaluation Metrics

Mutual Information (MI):

The aim of fusion techniques is to integrate complementary information from multiple sources. Hence, mutual information (MI) can be used to measure the amount of information transferred from the individual source images to the fused image and is defined as:

$$I_{XY} = \sum_{i,j} P_{XY}(i,j) \log \frac{P_{XY}(i,j)}{P_X(i) P_Y(j)} \tag{1}$$

where $P_X(i)$ and $P_Y(j)$ are the probability density functions in individual images and $P_{XY}(i,j)$ is the joint probability density function. Probability distributions can be computed by normalizing the corresponding joint and marginal histograms of the images.

Let A and B are source images and F is a fused image. The MI based objective metric can be formulated as [21]:

$$MI_F^{AB} = I_{AF} + I_{BF} \tag{2}$$

Large MI values indicate that more information has been transferred from source images into the fused image.

Objective Edge Based Measure (QE):

This metric computes the quality of the fused image by evaluating the amount of edge information that is transferred from source images to the fused image [22]. Sobel operator is applied to obtain edge magnitude $g(i,j)$ and orientation $\alpha(i,j)$ for each pixel of the image. The relative magnitude and orientation of input image A with respect to fused image F are defined as:

$$G^{AF}(i,j) = \begin{cases} \dfrac{g_F(i,j)}{g_A(i,j)}, & \text{if } g_F(i,j) > g_A(i,j) \\ \dfrac{g_A(i,j)}{g_F(i,j)}, & \text{otherwise} \end{cases} \tag{3}$$

$$\Phi^{AF}(i,j) = 1 - \frac{\left|\alpha_A(i,j) - \alpha_F(i,j)\right|}{\pi/2} \tag{4}$$

where $G^{AF}(i,j)$ and $\Phi^{AF}(i,j)$ are the relative magnitude and orientation, respectively. These edge parameters are then used to form the edge preservation values Q^{AF} from input image A to fused result F by the product of a sigmoid mapping function.

$$Q^{AF}(i,j) = \frac{\Gamma_g \Gamma_a}{(1+\exp^{k_g(G^{AF}(i,j)-\sigma_g)})(1+\exp^{k_\alpha(\Phi^{AF}(i,j)-\sigma_\alpha)})} \tag{5}$$

The constants which determine the shapes of the sigmoid mapping are chosen $k_g = -15, \sigma_g = 0.5, \Gamma_g = 1.0006$ and $k_\alpha = -25, \sigma_\alpha = 0.8, \Gamma_\alpha = 1.0123$ [22]. The overall objective quality measure $QE^{AB/F}$ is expressed by weighting the normalized edge information values of both input images as:

$$QE^{AB/F} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} Q^{AF}(i,j)\omega^A(i,j) + Q^{BF}(i,j)\omega^B(i,j)}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\omega^A(i,j) + \omega^B(i,j)\right)} \tag{6}$$

where $\omega^A(i,j)$ and $\omega^B(i,j)$ are a function of edge magnitude. QE ranges from 0 to 1 where 0 denotes complete loss of edge information and 1 denotes the perfect fusion without loss.

B. Multi-objective Genetic Algorithms

Optimization problems in engineering seldom regard only one objective. Usually, several performance functions must be considered at the same time. In general, these functions may not be continuous and are not differentiable. A popular way to solve these kinds of problems is to use population based heuristic optimization methods such as multi-objective genetic algorithms (MOGAs) [23]. The vector of objectives to a multi-objective problem can be given as:

$$f(x) = [f_1(x), f_2(x), ..., f_m(x)] \tag{7}$$

Until the systems capabilities are determined and tradeoffs between the objectives fully understood, the relative importance of these objectives is not generally known. While the number of the objective increases, tradeoffs become complex. For multi-objective problems, there is no unique solution. Instead, non inferior solutions (also called Pareto optimality) can be used to characterize the objectives.

The main goal of multi-objective optimization problem is to minimize the vector of objectives $f(x)$. The set of all values defines the feasible region Ω and any point $x \in \Omega$ is a feasible solution. A point $x^* \in \Omega$ is Pareto Optimal if for every $x \in \Omega$ and $I = \{1,2,\ldots k\}$ either, $\forall_{i \in I}[f_i(x) = f_i(x^*)]$ or, there is at least one $i \in I$ such that $f_i(x) > f_i(x^*)$. This definition states that x^* is a Pareto optimal if there exists no feasible vector x which would decrease some criterion without causing a simultaneous increment in at least on other criterion. Hence, the aim of the multi-objective optimization algorithm is to

obtain the Pareto optimal solutions to the problem [24]. The basic flowchart of MOGA is given in Fig. 1.

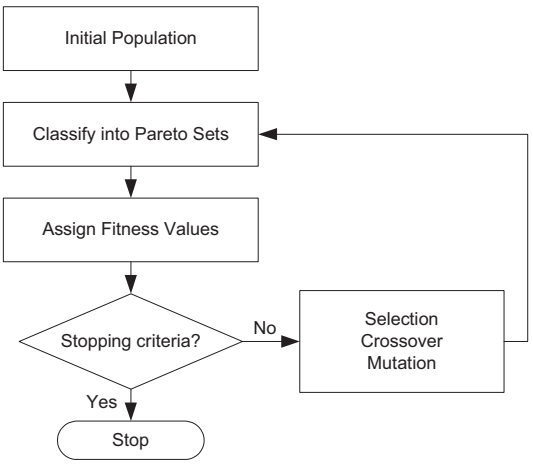


Fig. 1. Basic flowchart of MOGA.

C. Spatial Domain Multi-focus Image Fusion

It is assumed that the images of the same scene (A and B) are acquired from the same point of view with different focal settings. Therefore, they consist of focused and defocused regions. The focused regions of the images need to be determined properly to obtain a single focused image. The algorithm first divides both source input images into non-overlapping blocks. Then, the sharpness of the corresponding blocks is compared. Finally, the sharpest block is transferred into the fused image. The block size plays an important role in the fused image. Therefore, an optimal block size increases the quality of the fused image.

In order to obtain an optimal fused image, the proposed method in this paper employs the MOGA algorithm to search for optimal parameters. The parameters that are subject to the optimization process are the dimension (width and the height) of image blocks. The steps of the proposed method can be described as follows:

Step 1. Define the control parameters of MOGA and termination criteria.

Step 2. Generate initial population of possible solutions. The individuals of the population consist of two parameters corresponding to width (m) and height (n) of the optimal block size for the desired fused image.

Step 3. Evaluate the population and calculate the fitness values of each individual, x, consisting of two integer values is represented as $x_i=\{m_i, n_i\}$:

a. Divide the input images A and B into non-overlapping blocks with the size of $m \times n$. Refer to the ith image blocks of A and B images by A_i and B_i, respectively.

b. Calculate the spatial frequency (SF) of A_i and B_i for determining the sharpness values of the blocks and denote the results of A_i and B_i by SF_i^A, and SF_i^B, respectively. SF can be defined as [25]:

$$SF = \sqrt{C^2 + R^2} \tag{8}$$

where C and R represents the row and column gradients, respectively:

$$C = \left[\frac{1}{M \times N} \sum_i \sum_j \left[f(i,j) - f(i-1,j) \right]^2 \right]^{1/2}$$
$$R = \left[\frac{1}{M \times N} \sum_i \sum_j \left[f(i,j) - f(i,j-1) \right]^2 \right]^{1/2} \tag{9}$$

c. Compare the sharpness values of two corresponding blocks A_i and B_i to determine the sharper image block, and construct ith block F_i of the fused image as:

$$F_i = \begin{cases} A_i & SF_i^A > SF_i^B \\ B_i & SF_i^B > SF_i^A \\ \dfrac{A_i + B_i}{2} & otherwise \end{cases} \tag{10}$$

d. After evaluating Eq. (10) for each corresponding block of the source images, calculate the fitness value of the fused image by employing objective evaluation metrics MI and QE. The larger the MI and QE, the better the solution (or individual) is. Therefore, in order to calculate cost values of the each solution, the following expression is adopted as the objective function to be minimized:

$$f_1(M,N) = 1/MI_F^{AB}$$
$$f_2(M,N) = 1/QE^{AB/F} \tag{11}$$

Step 4. Apply the operators of MOGA (selection, crossover and mutation) to find new solutions.

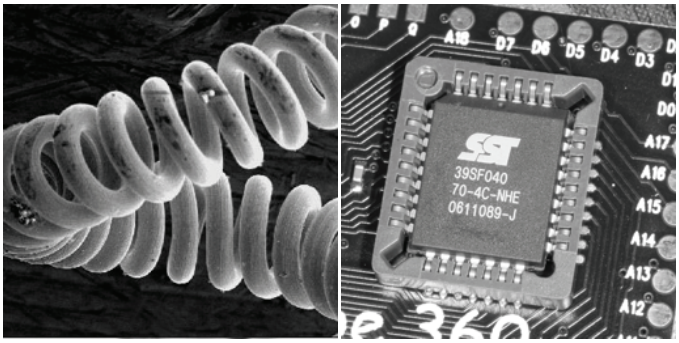
Step 5. Repeat the steps 3 and 4 until a predefined condition is satisfied.

It must be noted that, MOGA algorithm used in the implementation uses real valued numbers. However, the dimensions of blocks (m and n) are integer values. Thus, the individual solutions produced by MOGA are truncated in the optimization process.

III. Experimental Results

Experiments are performed on a set of images to compare the proposed method with well-known image fusion methods: Laplacian Pyramid and DWT. First, fusion of artificially produced multi-focus images is realized (*Filament* and *PCB*). Then, the methods are tested on a real multi-focus image (*Pepsi*).

The images in the dataset are 256×256 in sizes and 8-bit gray level. *Filament* is obtained from a scanning electron microscope (SEM) with $200\times$ zoom and shown in Fig. 2a. The artificial blurred versions of *Filament* obtained by a 3×3 Gaussian filter of radius 1 are shown in Figs. 3a and 3b, respectively. *PCB* (Fig. 2b) is an optical image and artificial multi-focus images of *PCB* are obtained by using a 3×3 Gaussian filter of radius 2 (Figs. 3c and 3d). *Pepsi* images are optical multi-focus image pair and shown in Figs. 3e and 3f.



(a) (b)

Fig 2. Reference images for producing artificial multi-focus images.

In the experiments, four levels of decomposition are employed for both Laplacian Pyramid and DWT methods. The filter chosen for DWT is Daubechies 'db4'. In the proposed method, the control parameters of MOGA are chosen as: population size is 40, crossover fraction is 0.8, Pareto fraction is 0.35 and maximum generation number is 500. Mutation scheme is chosen as adaptive.

Peak signal-to-noise ratio (PSNR) is also used to evaluate the quality of the artificial 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:

$$PSNR = 10\log_{10}\left(\frac{255^2}{\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[R(i,j)-F(i,j)\right]^2}\right) \qquad (12)$$

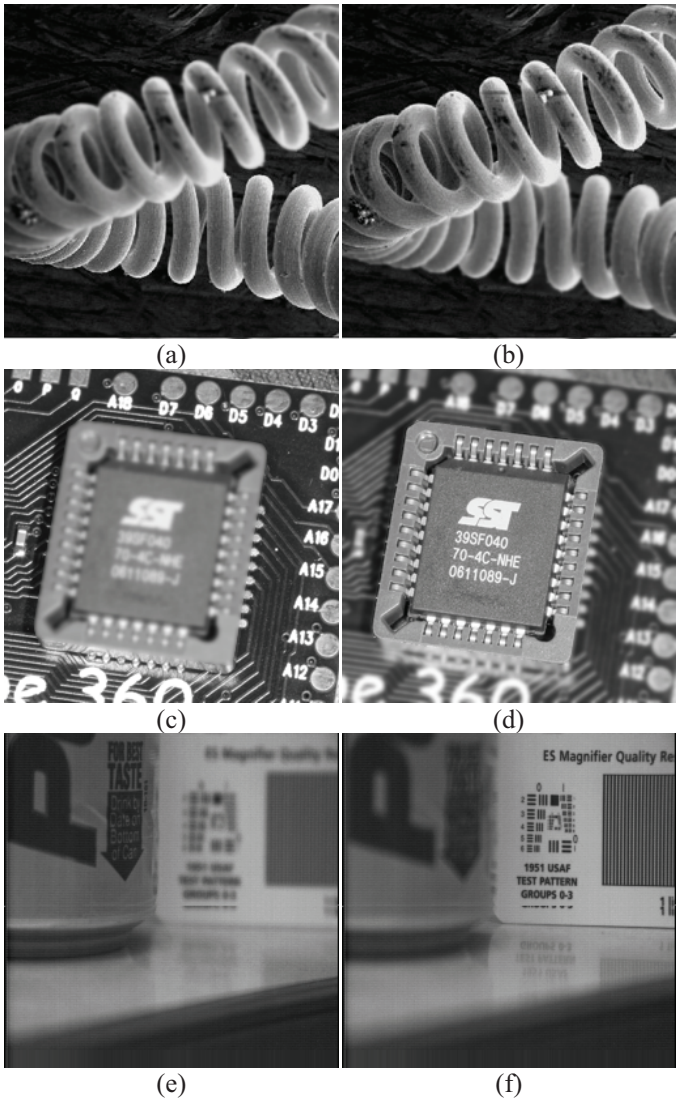
where R is the everywhere-in-focus ideal reference image, F is the fused image. PSNR can not be used to evaluate the fusion of *Pepsi* image since the everywhere-in-focus image under analysis is not available

The results of methods for the fusion of *Filament*, *PCB* and *Pepsi* images are given in the tables I, II and III, respectively. In Table I and II, PSNR results are also given together with the MI and QE results. It must be noted that, only three selected Pareto optimal results are given for proposed MOGA method in the following tables. The optimal block dimensions obtained by using MOGA are also shown in the parenthesis in the tables.

As can be seen from Table I and II, the proposed method outperforms Laplacian Pyramid and DWT in terms of MI, QE and PSNR. Moreover, the PSNR results are also in agreement with the results of MI and QE metrics.

It can be also observed that from Table III, proposed MOGA method is better than Laplacian Pyramid and DWT methods with respect to the objective metrics MI and QE.

Another important issue is that the optimal block sizes of the proposed method for different experimental images are different from each other. Therefore, it is not possible to obtain a universal block size for SDMIF method, as our experiments prove this fact.



(a) (b) (c) (d) (e) (f)

Fig 3. Multi-focus image dataset. (a) *Filament* far focused, (b) *Filament* near focused, (c) *PCB* focus on the board, (d) *PCB* focus on the chip, (e) *Pepsi* focus on the can, (f) *Pepsi* focus on the card.

TABLE I
COMPARISON OF FUSION METHODS ON *FILAMENT* IMAGE.

	MI	QE	PSNR
MOGA (5x20)	10,8913	0,8689	43,5630
MOGA (4x20)	10,8878	0,8689	43,2138
MOGA (6x20)	10,8855	0,8690	42,9772
DWT	8,9686	0,8519	36,9953
Lap. Pyramid	9,7518	0,8597	40,3930

Fused images of *Filament*, *PCB* and *Pepsi* which were obtained by using different fusion methods are given in Fig. 4. The results of the proposed MOGA method are given in the first column of Fig. 4. Similarly, the results of Laplacian Pyramid and DWT are given in the second and third columns of Fig. 4, respectively. As can be seen from the fused images, the fused images of the proposed method are better than those of Laplacian Pyramid and DWT with respect to the visual evaluations.

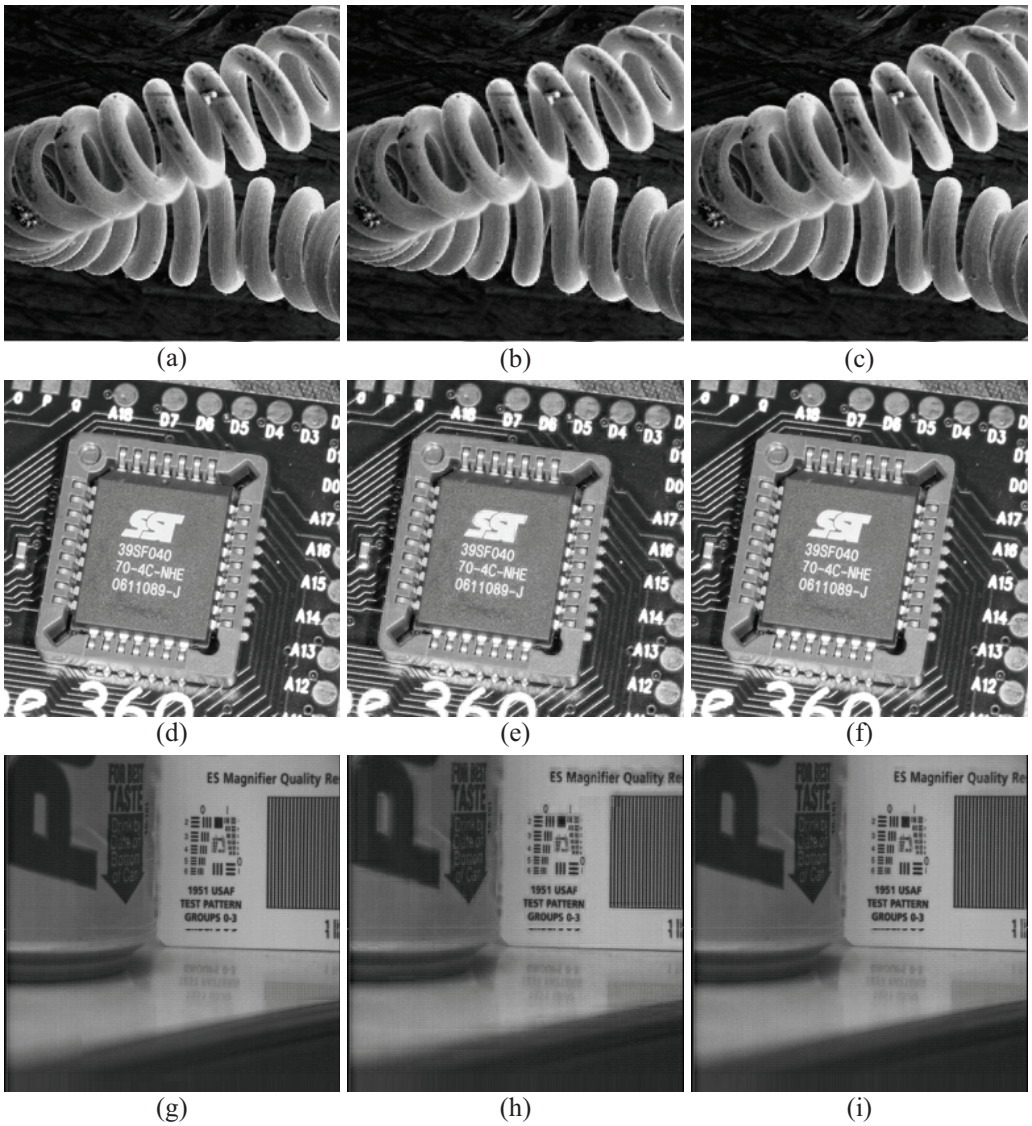


Fig. 4. Fused images. First column represents the results of MOGA, second column represents the results of DWT, and third column represents the results of Laplacian Pyramid.

TABLE II
COMPARISON OF FUSION METHODS ON *PCB* IMAGE.

	MI	QE	PSNR
MOGA (15x7)	8,4864	0,7585	40,8794
MOGA (16x10)	8,5060	0,7584	39,9186
MOGA (17x10)	8,5146	0,7577	39,3465
DWT	5,4450	0,7253	33,7109
Lap. Pyramid	6,7389	0,7447	38,5888

To emphasize the errors of the fused images, absolute difference images between fused images and the ideal reference images are illustrated in Fig. 5. For *Pepsi* image, difference images between the fused image and the source images are given since the ideal reference image does not exist. It can be observed from Fig. 5 that fusion error of the proposed method is less than Laplacian Pyramid and DWT.

TABLE III
COMPARISON OF FUSION METHODS ON *PEPSI* IMAGE.

	MI	QE
MOGA (62x59)	8,9572	0,7637
MOGA (59x59)	8,9724	0,7627
MOGA (59x61)	8,9866	0,7620
DWT	6,2180	0,6957
Lap. Pyramid	7,0357	0,7275

IV. CONCLUSIONS

In this paper, a novel approach is presented for the fusion of multi-focus images. The proposed method decomposes the source defocused images into blocks, and then selects the blocks with higher sharpness values obtained by employing spatial frequency (SF) sharpness criterion. The selected blocks are finally combined to form the fused image. In the experiments, both artificial and natural multi-focus images

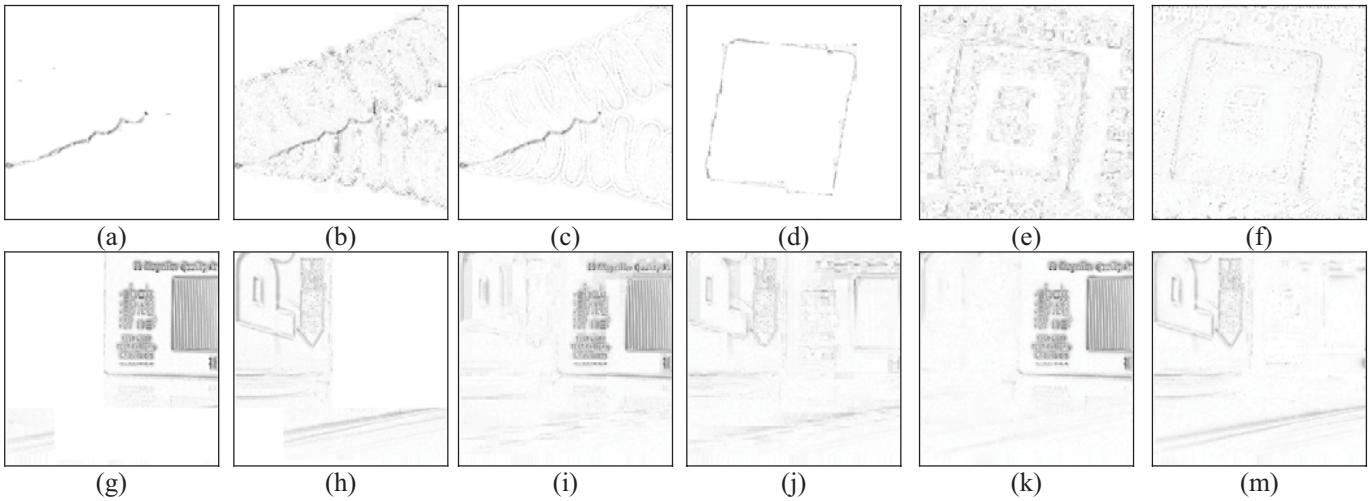


Fig. 5. Difference images between fused images and the reference/source images: (a) Fig. 4a and Fig. 2a, (b) Fig. 4b and Fig. 2a, (c) Fig. 4c and Fig. 2a, (d) Fig. 4d and Fig. 2b, (e) Fig. 4e and Fig. 2b, (f) Fig. 4f and Fig. 2b, (g) Fig. 4g and Fig. 3e, (h) Fig. 4h and Fig. 3f, (i) Fig. 4i and Fig. 3e, (j) Fig. 4g and Fig. 3f, (k) Fig. 4h and Fig. 3e, (m) Fig. 4i and Fig. 3f.

are evaluated. Experimental results show that, the proposed multi-objective genetic algorithm (MOGA) based method is better than Laplacian Pyramid and DWT based methods in terms of both objective fusion metrics such as mutual information (MI) and non-reference based objective edge metric (QE) and qualitative evaluations.

The advantages of the method can be summarized as follows:

- It does not require the reference image. Hence, it can be used in practical applications.
- It has a simple structure. The method does not require a transform; whole operation is carried out in the spatial domain (pixel level).
- MOGA optimizes only two variables (width and height of the block size in SDMIF method), so it is computationally feasible than other intelligent optimization algorithms based methods [20].

In an industrial context, the implementation of the proposed algorithm is complicated and time consuming. However, if the algorithm is implemented on a digital signal processor (DSP) or a field programmable gate array (FPGA) chips, the computational performance of the algorithm can be improved.

As a future work, different objective image fusion metrics and sharpness criterion functions can be studied to increase the performance of the proposed method.

REFERENCES

[1] D.T. Pham, and R.J. Alcock, "Smart Inspection Systems: Techniques and Applications of Intelligent Vision", Academic Press, 2003.

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

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

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

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

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

[7] Z. Zhang, "A stereovision system for a planetary rover: calibration, correlation, registration, and fusion", Machine Vision and Applications 10, pp. 27–34, 1997.

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

[9] 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.

[10] 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.

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

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

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

[14] 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.

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

[16] V.Aslantas, and R.Kurban, "A comparison of criterion functions for fusion of multi-focus noisy images", Optics Communications, accepted for publications.

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

[18] W. Huang, and Z. Jing, "Evaluation of focus measures in multi-focus image fusion", Pattern Recogn. Lett. 28, pp. 493-500, 2007.

[19] X.M. Zhang, J.Q. Han, and P.F. Liu, "Restoration and fusion optimization scheme of multifocus image using genetic search strategies", Optica Applicata 35,pp. 927-940, 2005.

[20] Y. Niu, and L. Shen, "A Novel Approach to Image Fusion Based on Multi-Objective Optimization", in Proceedings of the 6th World Congress on Intelligent Cont. and Automation, pp.9911-9915, 2006.

[21] G.H. Qu, D.L. Zhang, and P.F. Yan, "Information measure for performance of image fusion", Electronic Letters 38, pp.313-315, 2001.

[22] C.S. Xydeas, and V. Petrovic, "Objective image fusion performance measure", Electronic Letters 36, pp. 308-309, 2000.

[23] K. Deb, "Multi-Objective Optimization Using Evolutionary Algorithms", Wiley, 2009.

[24] E. Alba, B. Dorronsoro, F. Luna, A.J. Nebro, P. Bouvry, and L. Hogie, "A cellular multi-objective genetic algorithm for optimal broadcasting strategy in metropolitan MANETs", Computer Communications 30, pp. 685-697, 2007.

[25] A.M. Eskicioglu, and P.S. Fisher, "Image quality measures and their performance", IEEE Trans. on Commun. 43, pp. 2959-2965, 1995.