A Novel Automatic Block-based Multi-focus Image Fusion via Genetic Algorithm

Yong Yang[1] , Wenjuan Zheng[1] and Shuying Huang[2]
[1]School of Information Technology, Jiangxi University of Finance and Economics,
Nanchang 330032, China
[e-mail: greatyangy@126.com]
[2]School of Software and Communication Engineering, Jiangxi University of Finance and Economics,
Nanchang 330032, China
[e-mail: shuyinghuang2010@126.com]
*Corresponding author: Yong Yang

Received January 14, 2013; revised March 21, 2013; revised May 8, 2013; accepted June 4, 2013; published July 30, 2013

Abstract

The key issue of block-based multi-focus image fusion is to determine the size of the sub-block because different sizes of the sub-block will lead to different fusion effects. To solve this problem, this paper presents a novel genetic algorithm (GA) based multi-focus image fusion method, in which the block size can be automatically found. In our method, the Sum-modified-Laplacian (SML) is selected as an evaluation criterion to measure the clarity of the image sub-block, and the edge information retention is employed to calculate the fitness of each individual. Then, through the selection, crossover and mutation procedures of the GA, we can obtain the optimal solution for the sub-block, which is finally used to fuse the images. Experimental results show that the proposed method outperforms the traditional methods, including the average, gradient pyramid, discrete wavelet transform (DWT), shift invariant DWT (SIDWT) and two existing GA-based methods in terms of both the visual subjective evaluation and the objective evaluation.

Keywords: Multi-focus image fusion, genetic algorithm, SML, edge information retention

This work was supported by the National Natural Science Foundation of China (No. 60963012, No. 61262034), by the Key Project of Chinese Ministry of Education (No. 211087), by the Natural Science Foundation of Jiangxi Province (No. 2010GZS0052, No. 20114BAB211020, No. 20132BAB201025), by the Young Scientist Foundation of Jiangxi Province (No. 20122BCB23017), and by the Science and Technology Research Project of the Education Department of Jiangxi Province (No. GJJ13302).

http://dx.doi.org/10.3837/tiis.2013.07.009

1. Introduction

Image fusion refers to an image processing technique that combines two or more source images that have been registered into a single image according to some fusion rules. The fused image that is obtained has more information than any single source image, which is more adapted to the human visual perception and is more suitable for the subsequent processing of the computer [1], [2]. Image fusion is usually divided into three levels: pixel-level image fusion, feature-level image fusion and decision-making level image fusion [3]. Currently, image fusion as a very important technology has been widely used in many fields, such as military, public security, computer vision, medicine, and remote sensing [4], [5], [6], [7].

Because of the limited depth of field of optical lenses in CCD devices, it is often impossible to obtain an image that contains all relevant objects in focus, which means if one object in the scene is in focus, another one will be out of focus [8]. Therefore, to obtain all information of the scene, images taken from the same scene that focus on different objects must be fused, which is called multi-focus image fusion [9]. After fusing, the resulting image not only contains redundant information and complementary information for each of the individual source images but also contains information that can not be embodied in any of the single source images.

The simplest multi-focus image fusion method simply takes the average value of the source image pixels as the fused image pixel value. This method is simple and is suitable for real-time processing, but does not consider the correlation among the pixels solely by superposition. As a result, the effect of the images fusion is not satisfactory and has poor contrast [10]. The image pyramid technology that was proposed by Burt et al has been widely applied in image fusion. Various methods based on the pyramid transform have been proposed, such as the Laplacian pyramid [11], the gradient pyramid [12], and the ratio of low-pass pyramid [13], etc. Pyramid decomposition-based image fusion can achieve a good effect, but the tower decomposition of the image is redundant decomposition. Because the information of the different decomposition layers is correlative, which reduce the stability of the pyramid based algorithm. Wavelet technology has emerged to solve the instability of the pyramid decomposition algorithm because wavelet decomposition is a non-redundant decomposition. Wavelet is directional and can be decomposed into different scales and different directions. Moreover, wavelets have good locality with respect to time-frequency. Therefore, multi-focus image fusion algorithms based on wavelet transform have been widely used [14], [15], [16]. The basic idea of this approach is to perform wavelet decomposition on each source image to generate a wavelet pyramid of each source image; then all of the decompositions are integrated according to some fusion rules, to form the fused wavelet pyramid. Finally, we reconstruct the fused image by performing an inverse wavelet transform on the fused wavelet pyramid. Although the methods that are based on wavelet transforms and pyramid transforms can often achieve satisfactory results, their multi-resolution decompositions is shift-variant because of an underlying down-sampling process. In other words, when there is a slight camera/object movement or there is mis-registration of the source images, their performance will quickly deteriorate [17], [18]. To overcome this problem, a method based on a shift-invariant discrete wavelet transform (SIDWT) was proposed in the literature [19]. However, the implementation of the algorithm is more complicated and more time consuming. Additionally, the algorithm requires more memory.

One serious limitation of the multi-focus image fusion methods based on pixel or decomposition coefficients is that they are sensitive to noise or mis-registration. Additionally, pixel-based fusion methods consider only a single pixel or its neighboring pixel, without

accounting for the correlation between pixels. As a result, the block-based multi-focus image fusion methods have been proposed [20], [21]. The basic idea is to divide each source image into blocks size of M * N; then the clearest blocks are selected and are measured by some image evaluation standards to construct the fused image. This method is shift-invariant and can overcome the problems caused by a slight movement of camera/object or a mis-registration of the source images. However, the block-based fusion method is faced with some problems. The first problem is how to determine the size of the sub-block. If the size of the sub-block is too small, then the fusion result is prone to cause block effects. If the size of the sub-block is too large, then it is possible that some sub-blocks contain not only the clear portion but also contain the blurred portion. Another problem is which one index is to be chosen as the evaluation criterion to measure the sharpness of the sub-blocks. A multi-focus image fusion method based on visual characteristics has been proposed in the literature [22], in which the visibility was chosen as the evaluation criterion to measure the clarity of the sub-blocks; the clearest block was selected to construct the fused image when the difference in the sharpness between the two sub-blocks is greater than the threshold value (which was artificially set beforehand). Otherwise, it takes the average value of the pixel of the two sub-blocks as the pixel value of the fused image block. However, in this method, the initial size of a sub-block is determined by experience and the threshold is set artificially beforehand, which will have some impact on the fusion result. A multi-focus image fusion method based on an adaptive block is proposed in the literature [23]. In this method, each source image is divided into sub-blocks of 32 * 32 by experience, selecting the clearest block measured by spatial frequency to construct the fused image when the difference in the clarity between the two sub-blocks is greater than the threshold value (artificially set beforehand); otherwise, it takes the two sub-blocks and divides them into smaller blocks. In addition, this method has the same shortcomings as the method in the literature [22].

A multi-focus image fusion method that applies a genetic algorithm(GA) to seek the optimal size of the sub-block has been proposed in the literature [24], in this method the visibility was chosen as the evaluation criterion to measure the clarity of the sub-blocks, and the spatial frequency is taken as a fitness function to calculate the fitness of each individual, selecting the clearest block measured by spatial frequency to construct the fused image, and then generating the optimal sub-block by selection, crossover and mutation. GA is also used to find the optimal size of the sub-block in the literature [25], in this method the spatial frequency was chosen not only as the evaluation criterion to measure the clarity of the sub-blocks but also as the fitness function to calculate the fitness of each individual. The rule of the fusion is the same as in the literature [22]. The method generates the optimal sub-block by selection, crossover and mutation. However, in this method, the parameters that are artificially set have an influence on the effect of the fusion. To overcome the limitations of the conventional block-based multi-focus image fusion, a novel automatic block-based multi-focus image fusion with GA is proposed in this paper. The first step is to divide the source images that have been registered into blocks of size M * N, selecting the clearest block measured by Sum-modified-Laplacian (SML) to construct the fused image. The second step is to verify the consistency of the fusion result and to take the edge information retention as the fitness function to calculate the fitness of each individual. Finally, the optimal target solution is generated in a global sense by the selection, crossover and mutation procedures. The experiments show that the proposed method outperforms a series of traditional multi-focus fusion algorithms and two existing GA based methods in the literatures [24], [25].

The remainder of the paper is organized as follows. Section 2 briefly describes the related theory of SML, edge information retention and GA. The proposed fusion method is presented

in Section 3. Experimental results and analysis are given in Section 4. Finally, some concluding remarks are drawn in Section 5.

2. Related Concepts

2.1 Sum-modified-Laplacian (SML)

The Laplacian operator is often used as a high-frequency estimator. A study [26] has noted that the second derivatives of the Laplacian in the horizontal and vertical directions can have opposite signs, which can lead to the two values canceling each other out. Therefore, the modified Laplacian is proposed. The expression of the ML is as follows:

$$\begin{aligned}\nabla^2_{ML} f(x,y) &= \left|\frac{\partial^2 f(x,y)}{\partial x^2}\right| + \left|\frac{\partial^2 f(x,y)}{\partial y^2}\right| \\ &= \left|2f(x,y) - f(x-step,y) - f(x+step,y)\right| \\ &+ \left|2f(x,y) - f(x,y-step) - f(x,y+step)\right|\end{aligned} \tag{1}$$

Here, *f(x, y)* is the gray-level intensity of pixel *(x, y)*. In this method, 'step' always equals 1. Hence, the expression for SML becomes the following:

$$SML = \sum_{i=1}^{M}\sum_{j=1}^{N} \nabla^2_{ML} f(i,j) \tag{2}$$

where M , N determine the window size.

In this paper, the reason that we apply SML to measure the clarity of the sub-block is that the reference [18] has proved that SML can provide better performance than other evaluation metrics, such as the spatial frequency (SF), the energy of Laplacian (EOL), and the energy of image gradient (EOG). In addition, we have conducted a group of experiments to verify the effectiveness of the SML. **Fig.1 (a)** shows an image block that has a size of 64*64 and that was extracted from the 'Lena' image. **Fig.1 (b)-(d)** show the degraded versions blurred with a Gaussian radius of 0.5, 1.0 and 1.5, respectively. As shown in **Table 1**, when the image becomes more blurred, the value of the SML becomes smaller. The results presented here can demonstrate that SML can be effectively used to measure the clarity of the image sub-blocks.

(a) (b) (c) (d)

Fig. 1. Original and blurred versions of an image block extracted from the 'Lena' image. (a) Original image, (b) blurred result of radius 0.5, (c) blurred result of radius 1.0, (d) blurred result of radius 1.5.

Table 1. SML of the image blocks in Fig.1

	(a)	*(b)*	*(c)*	*(d)*
SML	94473	43130	25972	17752

2.2 Edge Information Retention

The edge information retention is used to evaluate the amount of edge information that is transferred from the input images to the fused image [27], [28]. This metric is a measure for objectively assessing the pixel-level fusion performance and is perceptually meaningful. The edge information retention is obtained by the following procedures:

(1) Assume that A and B are source images, and that F is the fused image. The edge information and direction information of the source image and the fused image are extracted by the Sobel operator, respectively. The expression of the edge strength denoted by $g_i(m,n)$ and the direction information denoted by $\partial_i(m,n)$ are defined, respectively, as

$$g_i(m,n) = \sqrt{s_i^x(m,n)^2 + s_i^y(m,n)^2} \tag{3}$$

$$\partial_i(m,n) = \tan^{-1}\left[\frac{s_i^y(m,n)}{s_i^x(m,n)}\right] \tag{4}$$

where, $i = A, B, F$, and s_i^x and s_i^y represent the horizontal edge detection image and the vertical edge detection image, respectively.

(2) Calculate the relative intensity denoted by $G^{AF}(m,n)$, and the relative angle, denoted by $A^{AF}(m,n)$, between the source image A and the fused image F. The expression of the $G^{AF}(m,n)$ and $A^{AF}(m,n)$ are defined, respectively, as

$$G^{AF}(m,n) = \begin{cases} \dfrac{g_F(m,n)}{g_A(m,n)} & g_A(m,n) > g_F(m,n) \\ \dfrac{g_A(m,n)}{g_F(m,n)} & g_A(m,n) \le g_F(m,n) \end{cases} \tag{5}$$

$$A^{AF}(m,n) = \frac{\left|\left|\partial_A(m,n) - \partial_F(m,n)\right| - \pi/2\right|}{\pi/2} \tag{6}$$

(3) Calculate the edge information retention from the source image A to the fusion image F, denoted by $Q^{AF}(m,n)$. The expression of $Q^{AF}(m,n)$ is defined as

$$Q^{AF}(m,n) = Q_g^{AF}(m,n) Q_\partial^{AF}(m,n) \tag{7}$$

where $0 \le Q^{AF}(m,n) \le 1$, $Q_g^{AF}(m,n)$ and $Q_\partial^{AF}(m,n)$, respectively, represent the retention value of the edge strength and the edge angle.

$$Q_g^{AF}(m,n) = \frac{\Gamma_g}{1 + e^{k_g(G^{AF}(m,n) - \sigma_g)}} \tag{8}$$

$$Q_\partial^{AF}(m,n) = \frac{\Gamma_\partial}{1 + e^{k_\partial(A^{AF}(m,n) - \sigma_\partial)}} \tag{9}$$

The parameters of $\Gamma_g, k_g, \sigma_g, \Gamma_\partial, k_\partial$ and σ_∂ determine the shape of the sigmoid functions that are used to form the values of $Q_g^{AF}(m,n)$ and $Q_\partial^{AF}(m,n)$. Similarly, the edge information retention from the source image B to the fusion image F, denoted by $Q^{BF}(m,n)$, can be obtained.

Thus, the average value of the edge information retention from the source image A and B to the fusion image F is defined as

$$Q^{AB/F} = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N}(Q^{AF}(m,n)\times\omega_A(m,n)+Q^{BF}(m,n)\times\omega_B(m,n))}{\sum_{m=1}^{M}\sum_{n=1}^{N}(\omega_A(m,n)+\omega_B(m,n))} \tag{10}$$

where $\omega_X(m,n)$ reflects the importance of $Q^{XF}(m,n)$ and $\omega_X(m,n)=[g_X(m,n)]^L$, L is a constant, and $X=A,B$.

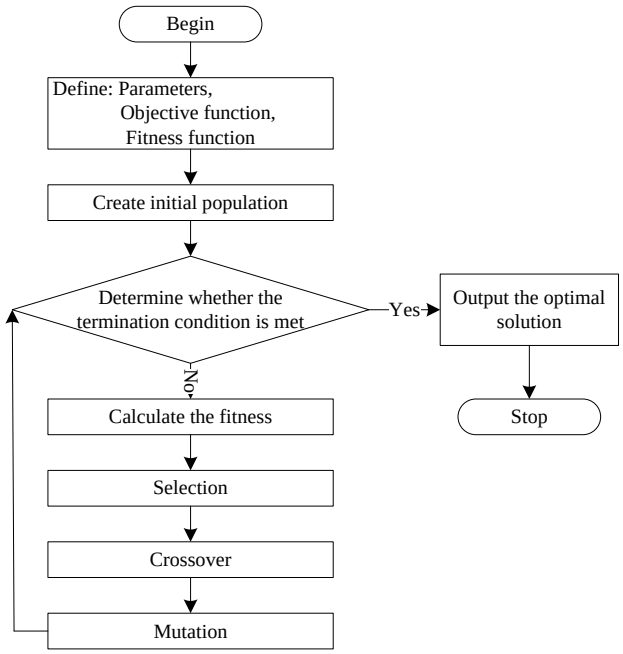


Fig. 2. Flow chart of the genetic algorithm

2.3 Genetic Algorithm

The genetic algorithm (GA) is an optimization strategy that simulates the process of biological evolution in the natural world [29]. In the GA, the solutions of a variety of complex problems are represented with a simple encoding. The nature of the GA is to perform an iterative process on a population. The basic idea is to generate an initial population randomly, to represent the problem solutions, and then to generate the next population by selection, crossover and mutation, which is based on the principle of survival of the fittest, until the termination condition is met. **Fig. 2** shows a flow chart of the genetic algorithm. The GA mainly includes the following three operations:

(1) Selection: Selection based on the fitness of the individual reflects the principle of survival of the fittest. The greater the fitness of the chromosome is, the greater the probability it is selected. The traditional selection strategies include the roulette selection strategy and the tournament selection strategy. Here, we introduce only the roulette selection algorithm. First, we calculate the probability of an individual who is selected in the population based on the individual fitness. Second, we calculate the cumulative probability of each individual. Third, we generate N random numbers in the interval [0, 1] and sort those random numbers from small to large; then we compare the result with the accumulated probabilities of each

individual one by one, where N is the number of individuals in the population. If the random number is in the interval (the cumulative probability of individual $i-1$, the cumulative probability of individual i), then the i th individual will be selected to produce offspring.

(2) Crossover: Crossover reflects the information exchange that takes place in nature. The chromosomes that have already been paired randomly will be crossed to produce new chromosomes based on the crossover probability. The position of the crossover is also random. Crossover methods mainly include single-point crossover, multiple-point crossover and uniform crossover. Here, we introduce only the single-point crossover. The single-point crossover is that chromosome information after the crossover point is exchanged.

(3) Mutation: Mutation reflects the gene mutations that take place in nature. Each individual will mutate by a certain probability to generate a new chromosome. The position of the mutation is also random. Typically, if the position in the selected chromosome has the value of 1, then the position will have the value of 0 after the mutation, and vice versa. Mutation is conducive to maintaining the diversity of the population so that the best solution can be searched for in a space that is as large as possible, avoiding premature convergence into a local solution.

3. The proposed GA-based Fusion Method

As described in Section 1, we know that the important problem of block-based multi-focus image fusion methods is how to determine the size of the sub-block. If the block is too large or too small, the fusion result will be not very good. How can we find a suitable size for the sub-block? The most direct way is to perform a large number of experiments, in which we use all combinations of the size of the sub-block to find a suitable block. However, this method is computationally intensive and time consuming.

Section 2.3 has described that the GA approach is an optimization strategy which can guide a search in a problem space according to the competition mechanism of the survival of the fittest and it can reduce the complexity of the problem dynamically and effectively [30], [31]. The probability of the optimal solution or near-optimal solution of the original problem obtained in a complex and large search space is high. Therefore, the method that searches the optimal size of the sub-block based on a GA has better robustness.

Based on the above analysis, a novel automatic block based multi-focus image fusion with a GA is proposed in this paper. The basic idea of our method is to handle the size of the sub-block as a chromosome and to use the GA to search for a suitable size of the sub-block. The first step is to divide the source images that have been registered into blocks of size M * N, selecting the clearest block measured by SML to construct the fused image. Then, the consistency of the fusion result is verified. The edge information retention is taken to be a fitness function to calculate the fitness of each individual. Finally, we generate the optimal target solution in a global sense by selection, crossover and mutation.

Fig. 3 shows the flow chart of the proposed multi-focus image fusion method. Here, we consider only the case of two source image fusions, although the method can be extended straightforwardly to handle more than two, with the assumption that the source images have always been registered.

Here, the focus is about the generation of the objective function. In the method, the objective function and the fitness function are two different functions because a chromosome expressed in binary is composed of two parts: one part is the width of the sub-block and the other part is the length of the sub-block. When the source images are to be fused, it is necessary to divide the source images that have been registered into blocks according to the

size of the sub-blocks. Therefore, the objective function is to calculate the decimal value of the two parts of each chromosome respectively, corresponding to the width and length of the sub-block. The fitness function is to calculate the fitness of each individual, where fitness is the edge information retention from the source image to the fusion image.

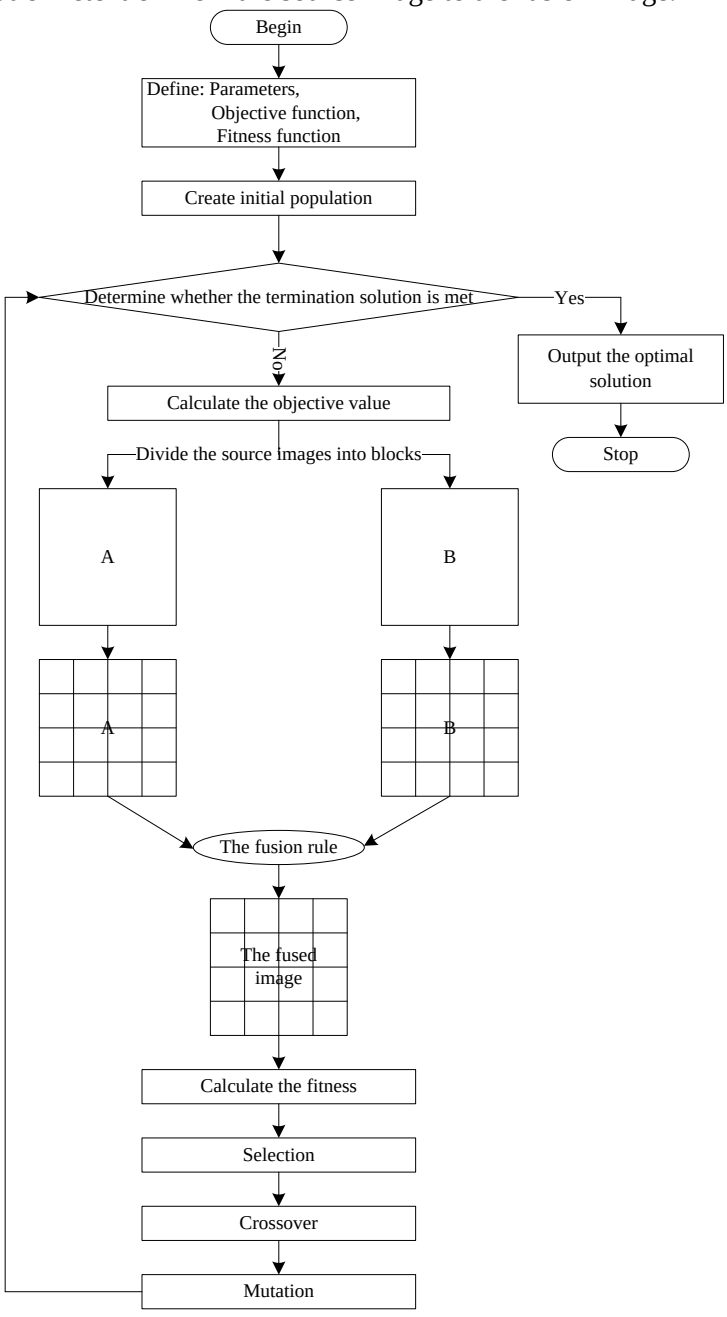


Fig. 3. A flow chart of the proposed method

Assume that the source images A and B have been registered. The steps of the proposed method are as follows:

(1) Define the length of the chromosome as l. Handle the size of the sub-block as a chromosome expressed in binary. Assume that the size of the sub-block is $m \times n$. Define l as

$l = width + length$, where $width = \lfloor \log_2 m \rfloor$, $length = \lfloor \log_2 n \rfloor$. Here, the expression of l can consider all types of images (any size), but the literature [24] can consider only the square images. Assume that C_i is denoted as the i th chromosome. The expression of C_i is

$$C_i = [a_{i,1}, a_{i,2}, ..., a_{i,width}, a_{i,width+1}, ..., a_{i,l}],$$

where the high $width$ bits of the chromosome represent the width of the sub-block, and the low $length$ bits of the chromosome represent the length of the sub-block.

(2) Generate the initial population pop. The size of the pop is $N * l$, where N represents the number of individuals in the population.

(3) Define the objective function as $calobjvalue(pop, width)$, where pop is a population and $width$ is a constant that represents the length of the binary string of the width of the sub-block. The output value is a matrix of $N * 2$ denoted by $pop\_regionSize$. Each row of the matrix represents a combination of the size of the sub-block. The first column of each row represents the width of the sub-block, and the second column of each row represents the length of the sub-block.

(4) Fuse the source images according to the matrix $pop\_regionSize$ and obtain the fused image matrix denoted by $F(1, N)$. Assume that the source images are fused based on the i th combination of the size of the sub-block. Divide the source images that have been registered into blocks of $pop\_regionSize(i,1) * pop\_regionSize(i,2)$. Denote the j th image block of the source images A and B by A_j and B_j, respectively. Denote the SML of the image block A_j and B_j by SML_j^A and SML_j^B, respectively. Therefore, the rules of constructing the j th block of the i th fusion image are as follows:

$$F\{i\}_j = \begin{cases} A_j & SML_j^A > SML_j^B \\ B_j & SML_j^A \le SML_j^B \end{cases} \tag{11}$$

The fused image $F\{i\}$ based on the i th combination of the size of the block can be obtained by fusing all of the sub-blocks of the source images. Then, by verifying the consistency of $F\{i\}$, we can obtain the final fused image $F\{i\}$.

(5) Calculate the fitness of the fused image based on different combinations of the size of the sub-block, where fitness is the edge information retention from the source image to the fusion image.

(6) Generate the new population by selection, crossover and mutation. The crossover probability and mutation probability is denoted by pc and pm, respectively.

(7) Iterative calculation. If the termination condition is not satisfied, then go to step (3). Otherwise, go to the next step. Here, the termination condition is that the number of iterations is reached.

(8) Obtain the optimal solution. The optimal chromosome is that fitness is the greatest. The corresponding objective function value of the optimal chromosome is the suitable size of the sub-block and the final fused image is obtained based on this optimal sub-block size.

4. Experimental Results and Analysis

4.1 Experimental Setup

In this section, the experiment is performed on one set of artificially produced images and two sets of naturally acquired multi-focus images. To evaluate the effectiveness and the robustness of the proposed fusion method, three sets of multi-focus images are also analyzed by using the conventional and classical methods, such as taking the average of the source images pixel by pixel, the gradient pyramid method, the DWT-based method, the SIDWT-based method, and the GA-based methods proposed in the literatures [24] and [25]. The decomposition level of the multi-scale transform is 4 layers. The wavelet basis of the DWT and SIDWT is DBSS(2,2) and Haar, respectively. The fusion rules of lowpass subband coefficients and the highpass subband coefficients are the 'averaging' scheme and the 'absolute maximum choosing' scheme, respectively. To compare the performance of the different image fusion methods mentioned above, the visual subjective evaluation and the objective evaluation are used. All of the algorithms are coded by MATLAB and are run on a 1.7GHz Pentium IV personal computer with a memory of 512 MB.

In the experiment, the parameters of GA are set as follows: $N = 20$, $pc = 0.6$, $pm = 0.05$, which are most widely used for GA. However, the number of iterations in GA is obtained experimentally. We take a set of values for the number of iterations to test its effect on the performance of our method. To obtain this parameter more reasonably, the experiments are repeated 10 times with different combinations of parameter sets on the multi-focus source images. The blue line in **Fig. 4** shows the average standard deviation (STD) of the fitness values errors under different values of the number of iterations, and the red line in Fig.4 shows its corresponding computation time. Based on **Fig. 4**, the STD tends to be stable in the 50 generations. Therefore, in this paper, the parameter for the number of iterations is set to 50.

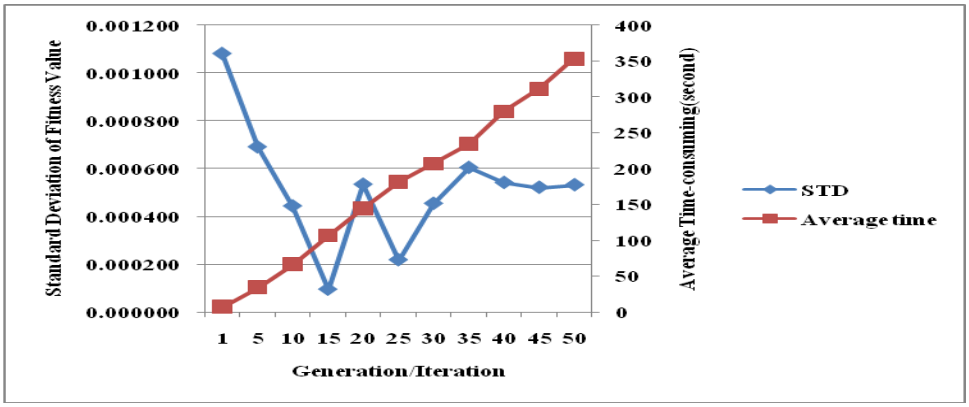


Fig. 4. Convergence and computation time for the experimental results on different iterations.

4.2 Fusion of Artificial Test Images

Fig. 5(a) is the 256-level standard reference image *cameraman*, which has the size of 256*256. We artificially produce a pair of out-of-focus images by blurring the top part to obtain the image in **Fig. 5(b)**, and then, we blur the bottom part to produce the image in **Fig. 5(c)**. The blurring is accomplished by using a Gaussian function. The two images are then fused by the seven different fusion methods mentioned above and the corresponding fused images are shown in **Fig. 5 (d)-(j)**, respectively. The size of the sub-blocks is automatically found by the

methods proposed in the literatures [24], [25], and our proposed method yields 153*218, 2*64 and 4*32, respectively. As can be clearly observed from **Fig. 5(d)-(j)**, the fused results from the pixel averaging and gradient pyramid method have poor contrast compared to the results from the DWT-based method, the SIDWT-based method, the methods proposed in the literatures [24], [25], and the proposed method. However, it is difficult for us to directly perceive the difference among the results of the latter five methods. Therefore, to make better visual comparisons, the difference images between the fused images and the cameraman reference image are given in **Fig. 6 (a)–(e)**. As shown in **Fig. 6 (a)–(e)**, the difference between the fused image obtained by the proposed method and the reference image is the lowest, which indicates that the fused result using the proposed method is better than the results from the other six methods.

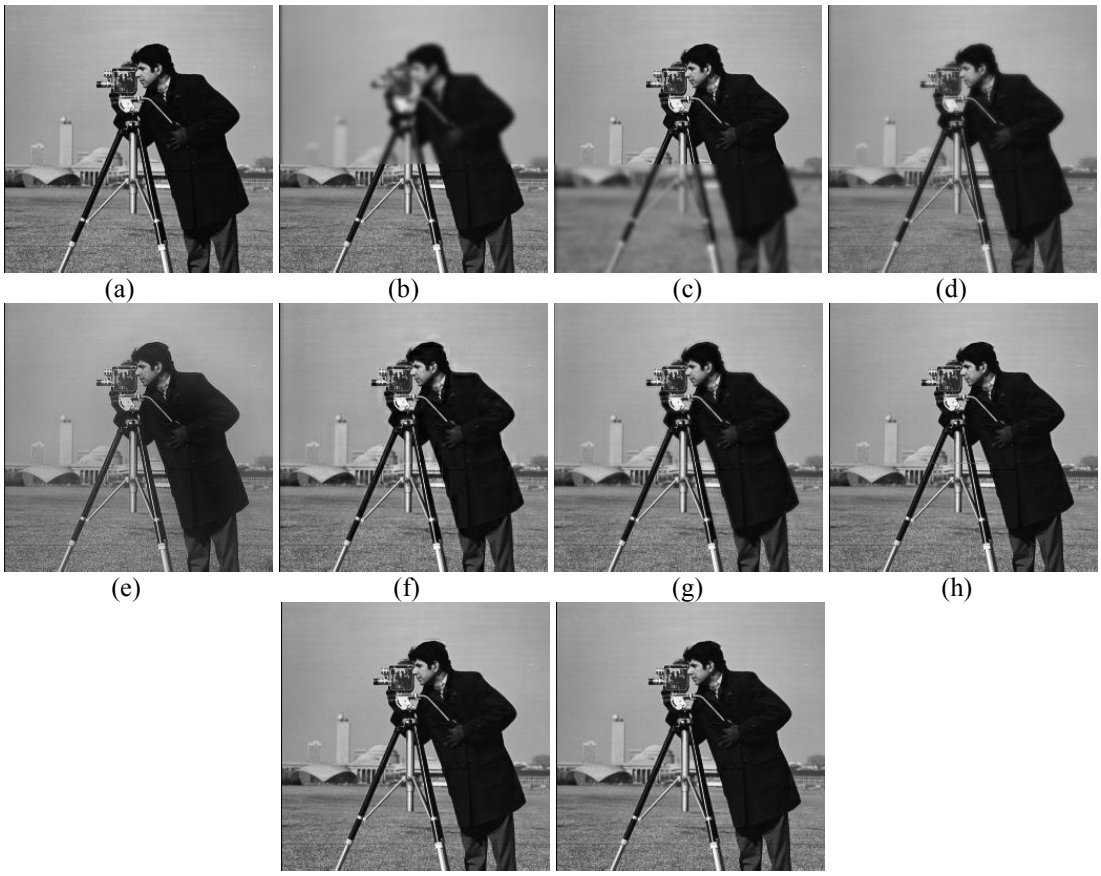


Fig. 5. Original and fused image of cameraman: (a) reference image (all in focus); (b) focus on the bottom; (c) focus on the top; (d) fused image using an average; (e) fused image using the Gradient pyramid; (f) fused image using DWT; (g) fused image using SIDWT; (h) fused image using the method proposed in the literature [24]; (i) fused image using the method proposed in the literature [25]; (j) fused image using the proposed method.

In addition to visual subjective evaluation, five evaluation criteria, including the edge information retention[27]($Q^{AB/F}$), the mutual information[21](MI), the root mean square error[20](RMSE), the distortion index[32](D) and the correlation coefficient[33](CORR) are employed to further compare the performance of the seven methods mentioned above. In these evaluation criteria, the $Q^{AB/F}$ evaluates the sum of the edge information preservation values from the source images to the fused image, the MI reflects the total amount of information that

the fused image contains compared with that in the source images, and the CORR reflects the degree of correlation between the fused image and the standard reference image. Therefore, the larger these values are, the better the performance is. However, because the RMSE measures the difference between the fused image and the standard reference image, and the D reflects the degree of distortion of the fused image compared with the standard reference image, the smaller their values are, the better the fusion effect is. The results of quantitative assessments are given in **Table 2**. From **Table 2**, we can observe that the proposed method is the best of the seven methods because it not only has the largest values of $Q^{AB/F}$, MI, and CORR, but also has the smallest values of RMSE and D. These quantitative results indicate that by the proposed method, the fused image can retain much more image information and has a smaller difference from the source images compared to the other six methods. The results of the subjective and objective evaluation presented here can clearly verify that the overall performance of the proposed method is significantly superior to that of the other methods.

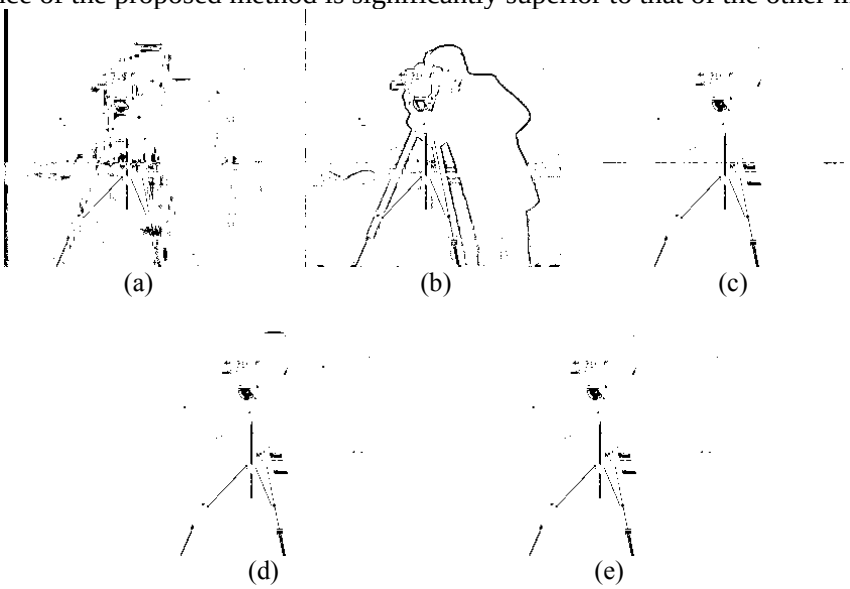


Fig. 6. Differences between the fused images in Fig.5(f)-(j) and the reference image in Fig.5(a). (a) difference between the fused image using DWT(Fig.5(f)) and the reference image(Fig.5(a)); (b) difference between the fused image using SIDWT(Fig.5(g)) and the reference image(Fig.5(a)); (c) difference between the fused image using the method proposed in the literature [24](Fig.5(h)) and the reference image (Fig.5(a)); (d) difference between the fused image using the method proposed in the literature [25](Fig.5(i)) and the reference image (Fig.5(a)); (e) differences between the fused image using the proposed method(Fig.5(j)) and the reference image(Fig.5(a)).

Table 2. Performance comparison of different fusion algorithms in Fig.5

Fusion algorithms	$Q^{AB/F}$	*MI*	*CORR*	*RMSE*	*D*
Average method	0.64149	6.7572	0.98853	12.812	9.7742
Gradient pyramid	0.68412	5.3354	0.99016	13.923	10.192
DWT	0.71006	6.8049	0.99836	7.9218	7.1221
SIDWT	0.70928	7.0703	0.99505	9.9778	7.6937
Literature [24]	0.74362	9.1562	0.99944	7.5706	7.0283
Literature [25]	0.74388	9.1413	0.99971	7.4219	6.9958
Proposed method	0.74414	9.1634	0.99974	7.4123	6.9909

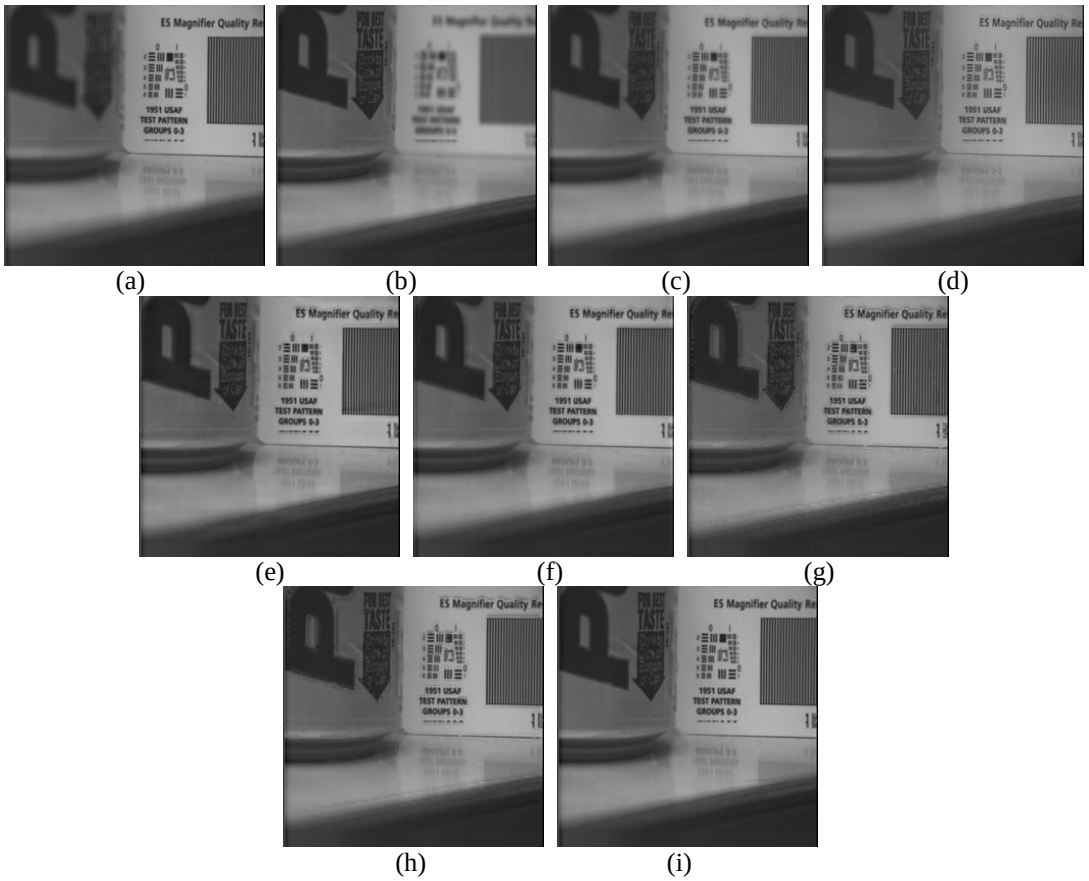


Fig. 7. Original and fused image of 'pepsi': (a) focus on the right; (b) focus on the left; (c) fused image using the average; (d) fused image using the Gradient pyramid; (e) fused image using DWT; (f) fused image using SIDWT; (g) fused image using the method proposed in the literature [24]; (h) fused image using the method proposed in the literature [25]; (i) fused image using the proposed method.

4.3 Fusion of Real Multi-focus Images

In this section, experiments have been performed on images that were acquired naturally. Tests were first realized on the *pepsi* multi-focus images, as shown in **Fig. 7(a)** and **(b)**. **Fig. 7(a)** is focused on the right, while Fig.7(b) is focused on the left. The size of the sub-blocks is automatically found by the methods proposed in the literatures [24], [25], and our proposed method has sub-block sizes of 2*2, 2*4 and 32*64, respectively. The fused images obtained by using the above seven methods are shown in **Fig. 7(c)-(i)**. From the fusion results, we can observe that the fusion effects acquired by using the pixel averaging method and gradient pyramid method are not satisfactory and have poor contrast, but it is difficult to discriminate the difference between the fused images of the DWT-based method, the SIDWT-based method, the methods proposed in the literatures [24], [25], and the proposed method. Therefore, to better show the difference among **Fig. 7(e)**, **Fig. 7(f)**, **Fig. 7(g)**, **Fig. 7(h)** and **Fig. 7(i)**, we calculate the difference images between the fused images and the source images, as shown in **Fig. 8(a)-(j)**. We know that, for the focused regions, the difference between the source image and the fused image should be zero after fusing. For example, in **Fig. 7(a)**, the testing card (the right) is focused, and in **Fig. 8(e)**, the difference between **Fig. 7(i)** and **Fig. 7(a)** in the testing card region is slight, which indicates that the whole focused area in **Fig. 7(a)** is contained in the fused image successfully, while the differences in **Fig. 8(a)-(d)** in the

testing card region are greater compared to in **Fig. 8(e)**, which demonstrate that the fused results using the DWT, the SIDWT, the methods proposed in the literatures [24], [25] are worse than the results of our proposed method. From the difference images in **Fig. 8(f)-(j)**, we can also find that, by the proposed method the pepsi can (the left focused region of **Fig. 7(b)**) is well focused after fusing because the left region in the difference image **Fig. 8(j)** is the smallest difference compared to **Fig. 8(f)-(i)**.

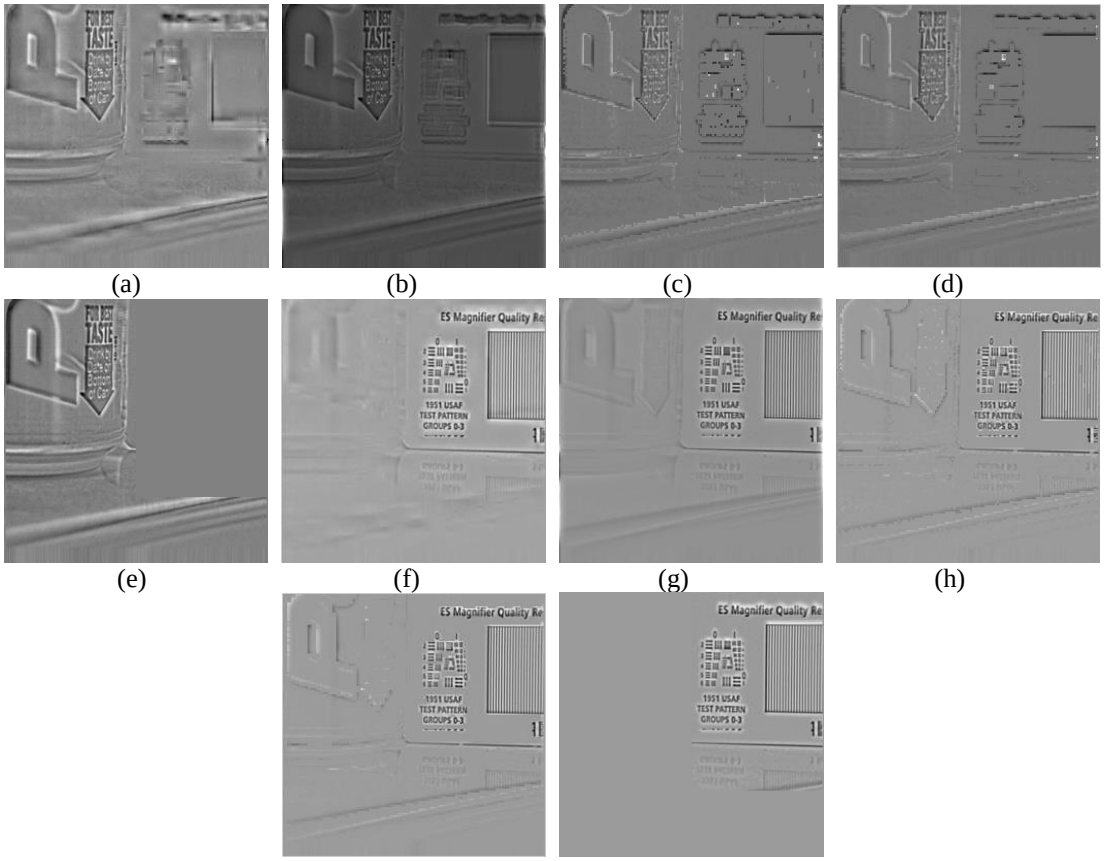


Fig. 8. The differences between the fused images in Fig.7(e)-(i) and the source images in Fig.7(a), (b). (a) The difference between the fused image using DWT (Fig.7(e)) and the source image in Fig.7(a); (b) The difference between the fused image using SIDWT (Fig.7(f)) and the source image in Fig.7(a); (c) The difference between the fused image using the method proposed in the literature [24] (Fig.7(g)) and the source image in Fig.7(a); (d) The difference between the fused image using the method proposed in the literature [25] (Fig.7(h)) and the source image in Fig.7(a); (e) The difference between the fused image using the proposed method (Fig.7(i)) and the source image in Fig.7(a); (f) The difference between the fused image using DWT (Fig.7(e)) and the source image in Fig.7(b); (g) The difference between the fused image using SIDWT (Fig.7(f)) and the source image in Fig.7(b); (h) The difference between the fused image using the method proposed in the literature [24] (Fig. 7(g)) and the source image in Fig.7(b); (i) The difference between the fused image using the method proposed in the literature [25] (Fig. 7(h)) and the source image in Fig.7(b); (j) The difference between the fused image using the proposed method (Fig.7(i)) and the source image in Fig.7(b).

To better evaluate the above fusion methods, a quantitative assessment of the performance of the seven methods is then conducted. It should be noted that because the reference image is

not available for real multi-focus images, only the two evaluation criteria, the $Q^{AB/F}$ and MI are used to objectively compare the fusion results. A quantitative comparison of their performance is shown in **Fig. 9**. As shown in **Fig. 9**, all of the quality indices of the proposed method are larger than those for the pixel averaging, the gradient pyramid, DWT, SIDWT, and the methods proposed in the literatures [24], [25], which means that the performance of the proposed method is the best among the seven fusion methods.

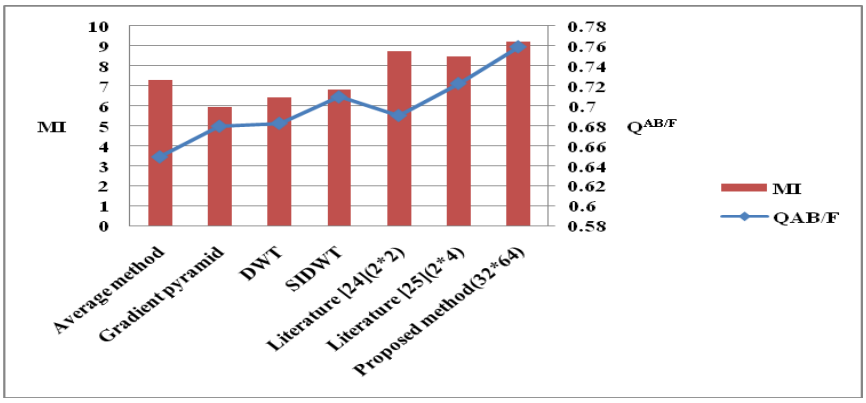


Fig. 9. Performance comparison of different fusion algorithms in Fig.7

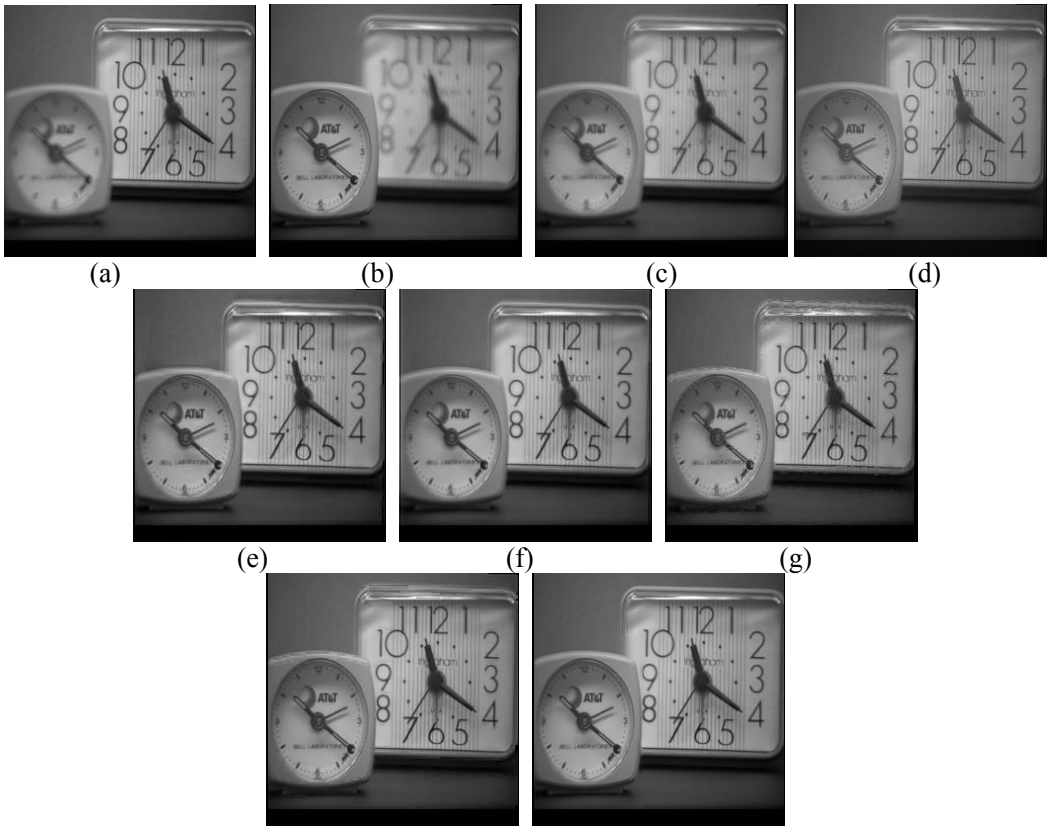


Fig. 10. Original and fused image of 'clock': (a) focus on the right; (b) focus on the left; (c) fused image using the average; (d) fused image using the Gradient pyramid; (e) fused image using DWT; (f) fused image using SIDWT; (g) fused image using the method proposed in the literature [24]; (h) fused image using the method proposed in the literature [25]; (i) fused image using the proposed method.

The last experiments are performed on another pair of *clock* images, as illustrated in **Fig. 10 (a)** and **(b)**, respectively. The suitable size of the sub-block is automatically found by the methods proposed in the literatures [24], [25], and our proposed method is 1*4, 2*8 and 64*16, respectively. The fused images obtained by using the above seven fusion methods are shown in **Fig. 10(c)-(i)**, respectively. The difference images between the fused images and the source images are given in **Fig. 11(a)-(j)**, respectively. The quantitative comparison of their performance is shown in **Fig. 12**. From the **Fig. 10(c)-(i)**, **Fig. 11 (a)-(j)** and **Fig. 12**, we can see that the fusion results of the experiment are identical to the previous experiment.

Hence, based on these experimental results on the two real sets of multi-focus images, we can conclude that the proposed method provides better performance and outperforms a series of traditional multi-focus fusion algorithms and the existing GA-based methods.

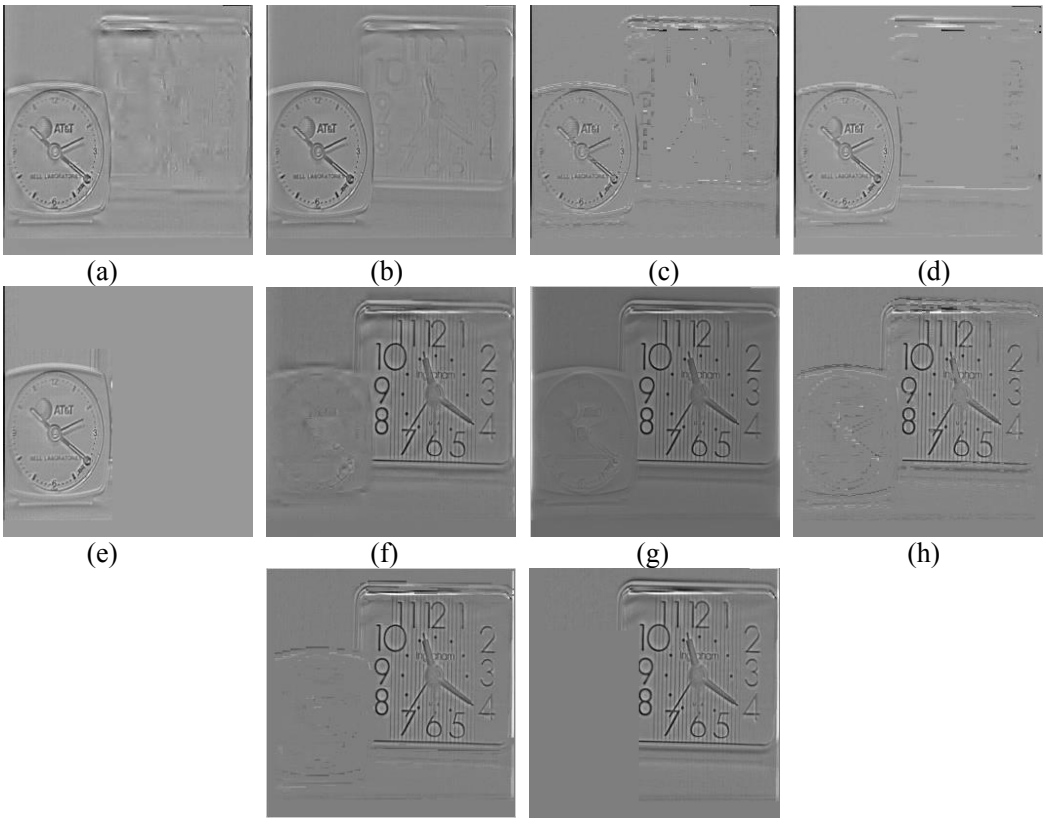


Fig. 11. Differences between the fused images in Fig.10(e)-(i) and the source images in Fig.10(a), (b). (a) The difference between the fused image using DWT (Fig.10(e)) and the source image in Fig.10(a); (b) The difference between the fused image using SIDWT (Fig.10(f)) and the source image in Fig.10(a); (c) The difference between the fused image using the method proposed in the literature [24] (Fig.10(g)) and the source image in Fig.10(a); (d) The difference between the fused image using the method proposed in the literature [25](Fig. 10(h)) and the source image in Fig.10(a); (e) The difference between the fused image using the proposed method (Fig.10(i)) and the source image in Fig.10(a); (f) The difference between the fused image using DWT (Fig.10(e)) and the source image in Fig.10(b); (g) The difference between the fused image using SIDWT (Fig.10(f)) and the source image in Fig.10(b); (h) The difference between the fused image using the method proposed in the literature [24] (Fig.10(g)) and the source image in Fig.10(b); (i) The difference between the fused image using the method proposed in the literature [25](Fig.10(h)) and the source image in Fig.10(b); (j) The difference between the fused image using the proposed method (Fig.10(i)) and the source image in Fig.10(b).

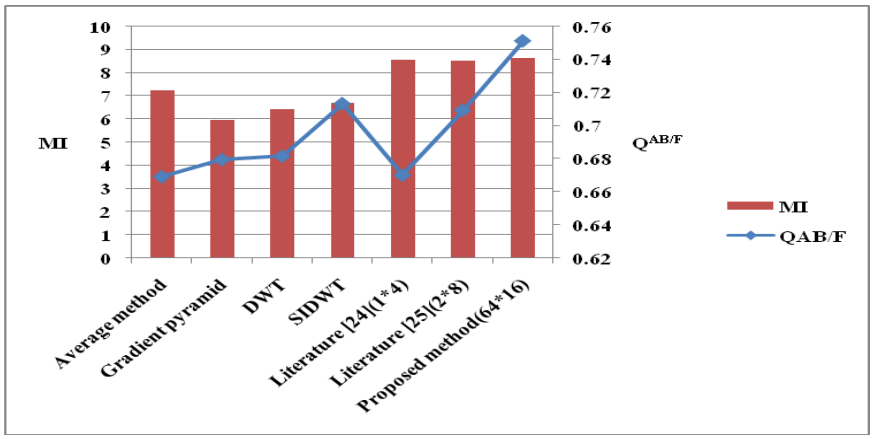


Fig. 12. Performance comparison of different fusion algorithms in Fig.10

5. Conclusions

The difficulty with block-based multi-focus image fusion is in deciding on the size of the sub-blocks. This paper presents a novel GA-based multi-focus image fusion method, with which the block size can be found automatically. The proposed method is composed of three steps. In the first step, the source images to be fused are divided into blocks of random size, and then, the clearest block measured by SML is selected to construct the preliminary fused image. In the second step, the consistency of the fusion result is verified, and the edge information retention is employed to calculate the fitness of each individual. Finally, through the selection, crossover and mutation procedures of the GA, the optimal solution for the sub-block size is found, which is ultimately used to fuse the images. The qualitative and quantitative experiments can clearly indicate the superiority and effectiveness of the proposed fusion method compared to a series of traditional fusion algorithms and two existing GA-based methods.

References

[1] A. A. Goshtasby, S. G. Nikolov, "Image fusion: Advances in the state of the art," *Information Fusion*, vol. 8, no. 2, pp. 114-118, 2007. Article (CrossRef Link)

[2] J. Dong, D. F. Zhuang, Y. H. Huang and J. Y. Fu, "Advances in Multi-Sensor Data Fusion: Algorithms and Applications," *Sensors*, vol. 9, no. 10, pp. 7771-7784, 2009. Article (CrossRef Link)

[3] Y. Chai, H. F. Li, Z. F. Li, "Multifocus image fusion scheme using focused region detection and multiresolution," *Optics Communications*, vol. 284, no. 19, pp. 4376-4389, 2011. Article (CrossRef Link)

[4] J. W Hu, S. T. Li, "The multiscale directional bilateral filter and its application to multisensor image fusion," *Information Fusion*, vol. 13, no. 3, pp. 196-206, 2012. Article (CrossRef Link)

[5] R. S. Rosa, J. A. García, J. Fdez-Valdivia, "From computational attention to image fusion," *Pattern Recognition Letters*, vol. 32, no. 14, pp.1778-1795, 2011. Article (CrossRef Link)

[6] B. Yang, S. T. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Information Fusion*, vol. 13, no. 1, pp.10-19, 2012. Article (CrossRef Link)

[7] Q. Zhang, L. Wang, H. J. Li, Z. K. Ma, "Similarity-based multimodality image fusion with shiftable complex directional pyramid," *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1544-1553, 2011. Article (CrossRef Link)

[8] H. J. Zhao, Z. W. Shang, Y. Y. Tang, B. Fang, "Multi-focus image fusion based on the neighbor distance," *Pattern Recognition*, vol. 46, no. 3, pp. 1002-1011, 2013.Article (CrossRef Link)

[9] S. Zheng, Y. Q. Sun, J. W. Tian, J. Liu, "Support Value Based Fusing Images With Different Focuses," in *Proc. of the International Conference on Machine Learning and Cybernetics*, pp. 5249-5254, 2005. Article (CrossRef Link)

[10] S. T. Li, B. Yang, "Multi-focus image fusion by combining curvelet and wavelet transform," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1295-1301, 2008.Article (CrossRef Link)

[11] P. J. Burt, E. H. Andelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532-540, 1983. Article (CrossRef Link)

[12] P. J. Burt, "A gradient pyramid basis for pattern selective image fusion," in *Proc. of the Society for Information Display Conference*, pp. 467-470, 1992.

[13] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognition*, vol. 9, no. 4, pp. 245-253, 1989. Article (CrossRef Link)

[14] S. M. Mahbubur Rahman, M. Omair Ahmad, M. N. S. Swamy, "Contrast-based fusion of noisy images using discrete wavelet transform," *IET Image Processing*, vol. 4, no. 5, pp. 374-384, 2010. Article (CrossRef Link)

[15] Y. Yang, "A Novel DWT Based Multi-focus Image Fusion Method," *Procedia Engineering*, vol. 24, pp. 177-181, 2011. Article (CrossRef Link)

[16] J. Tian, L. Chen, "Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure," *Signal Processing*, vol. 92, no. 9, pp. 2137-2146, 2012. Article (CrossRef Link)

[17] B. Yang, S. T. Li, "Multi-focus Image Fusion and Restoration with Sparse Representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 884-892, 2010. Article (CrossRef Link)

[18] W. Huang, Z. L. Jing, "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 493-500, 2007. Article (CrossRef Link)

[19] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Processing*, vol. 4, no. 11, pp.1549-560, 1995. Article (CrossRef Link)

[20] S. T. Li, J. T. Kwok, Y. N. Wang, "Multifocus image fusion using artificial neural networks," *Pattern Recognition Letters*, vol. 23, no. 8, pp. 985-997, 2002. Article (CrossRef Link)

[21] S. T. Li, B. Yang, "Multi-focus image fusion using region segmentation and spatial frequency," *Image and Vision Computing*, vol. 26, no. 7, pp. 971-979, 2008. Article (CrossRef Link)

[22] S. T. Li, Y. N. Wang, C. F. Zhang, "Feature of Human Vision System Based Multi-Focus Image Fusion," *Acata Electronica Sinica*, vol. 29, no. 12, pp. 1699-1701, 2001.

[23] J. Li, *The Research on Methods of Multi-focus Image Fusion*, Hunan University Press, 2006.

[24] X. M. Zhang, J. Q. Han, Y. Wang, "A Multifocus Image Fusion Algorithm for Adaptive Genetic Search," *Journal of Electronics & Information Technology*, vol. 28, no. 11, pp. 2054-2057, 2006.

[25] J. Kong, K. Y. Zheng, J. B. Zhang, X. Feng, "Multi-focus Image Fusion Using Spatial Frequency and Genetic Algorithm," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 220-224, 2008.

[26] S. K. Nayar, Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 8, pp. 824-831, 1994. Article (CrossRef Link)

[27] C. S. Xydeas, V. Petrović, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308-309, 2000. Article (CrossRef Link)

[28] S. T. Li, B. Yang, J. W. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Information Fusion*, vol. 12, no. 2, pp. 74-84, 2011. Article (CrossRef Link)

[29] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989.

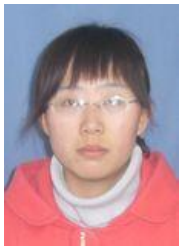
[30] M. T. Ayvaz, H. Karahan, M. M. Aral, "Aquifer parameter and zone structure estimation using kernel-based fuzzy c-means clustering and genetic algorithm," *Journal of Hydrology*, vol. 343, no. 3-4, pp. 240-253, 2007. Article (CrossRef Link)

[31] U. Maulik, "Medical Image Segmentation Using Genetic Algorithms," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 2, pp. 166-173, 2009. Article (CrossRef Link)

[32] A. G. Mahyari, M. Yazdi, "Fusion of panchromatic and multispectral images using temporal Fourier transform," *IET Image Process.*, vol. 4, no. 4, pp. 255-260. 2010. Article (CrossRef Link)

[33] Z. H. Li and H. Leung, "Fusion of Multispectral and Panchromatic Images Using a Restoration-Based Method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1482-1491, 2009. Article (CrossRef Link)

Yong Yang received his Ph.D. degree in Biomedical Engineering from Xi'an Jiaotong University, China, in 2005. From 2009 to 2010, he was a postdoctoral research fellow with the Chonbuk National University, Republic of Korea. He is currently a Full Professor with the School of Information Technology, Jiangxi University of Finance and Economics, China. He has published more than 70 research articles in international journals and conferences. His research interests include image processing, medical image analysis, and pattern recognition.



Wenjuan Zheng received her B.S. degree in Computer Science and Technology from Nanjing University of Science and Technology of Taizhou College, China, in 2011. Now she is pursuing her M.S. degree in Computer Application Technology of Jiangxi University of Finance and Economics, China. Her research interests include image processing, pattern recognition and information fusion.

Shuying Huang received her M.S. degree in Signal and Information Processing from Xi'an University of Technology, China, in 2005. She is currently a Ph.D. candidate with the Department of Electrical Engineering, Ocean University of China, China. She is also a Lecturer with the School of Software and Communication Engineering, Jiangxi University of Finance and Economics, China. Her research interests include image and signal processing, and pattern recognition.