Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Multi-focus Image Fusion Based on Salient Edge Information within Adaptive Focus-Measuring Windows

P. W. Huang, Cheng-I Chen
Department of Computer Science and Engineering
National Chung Hsing University
Taichung 40227, Taiwan
powhei.huang@msa.hinet.net

Phen-Lan Lin
Department of Computer Science and
Information Management, Providence University
Shalu, Taichung 433, Taiwan
lan@pu.edu.tw

***Abstract*—In this paper, we propose a new fusion scheme for multi-focus images based on some special edge information called discriminative edge points. A new focus measure is also proposed in this paper to identify which pixels in the source images will be included in the resultant image by evaluating the sum of neighborhood energy of the discriminative edge points within an adaptive focus-measuring window whose center is the location of the pixel under current consideration. Experimental results show that our fusion scheme outperforms other methods in the literature.**

***Keywords*—multifocus image fusion, discriminative edge points, focus measure, false edge**

I. INTRODUCTION

Image fusion process is a technique of integrating several source images of the same scene captured by different sensors or under different conditions into a composite image, which has a better description of the scene than any single source image. The resultant image acquired from image fusion technique is more informative and provides a good media for human perception and advanced image-processing tasks such as image segmentation, feature extraction, and target detection. Image fusion technique is not only capable of providing a better interpretation, but also improves the reduction of data storage as well as the reliability of image content.

When taking account of the objectives and benefits of image fusion, the following requirements must be imposed [1]:

1. The image fusion process should retain all relevant and salient information of source images while discarding irrelevant parts and noise in the fused result.

2. The image fusion process should not yield any artificial information such that the human observer or advanced image-processing tasks would be distracted.

3. The image fusion process should be capable of avoiding imperfections such as noise or mis-registration.

In an image captured by using a camera with finite depth of field, only objects within the depth of field are in-focus and sharp, whereas others are out-of-focus and blurred. The technique of acquiring an image, which contains all in-focus objects from different source images, is called multi-focus image fusion technique. In the literature, many methods for pixel-wise multi-focus image fusion have been developed. The computation of focus measure for a considered pixel in these methods is based on the information about that pixel and its surrounding ones, and the fusion rule is to select the pixel with largest focus-measuring value. Such an approach has the problem of dealing with false edges in blurred images. In other words, false edges may be included in the fused image accidentally. To reduce this problem, another approach is to compute the focus measure of a block of image by summarizing the focus-measuring values of all pixels within a block, and then preserve the whole block with largest focus-measuring value in the fused image. Although reducing the false edge problem, this kind of approach encounters the block-size selection problem, in which blocks with inappropriate block sizes may contain blurred area pixels and lead to a poorly fused result.

II. RELATED WORKS

Many techniques for image fusion have been proposed so far. Among them, averaging fusion method is the simplest one, which averages all source images pixel by pixel, and thus has the advantage of suppressing noise existing in some source image, but inevitably has serious side effects like reducing contrast or introducing blurred pixels. More complex and flexible multi-resolution fusion methods, such as image-pyramid fusion and wavelet-transform-based fusion, were also proposed [2][5][6][9]. They perform fusion process on the transformed images instead of directly on the raw images. Although multi-resolution fusion methods can preserve local details in the fusion result, they are still incapable of handling the false edge problem. False edge points are extra edge points appearing in an out-of-focus image due to blurring effect and the pixels around those false edge points usually have larger focus-measuring values and will be falsely identified as salient information. Wavelet-transform-based methods have several serious drawbacks. Firstly, underflow and overflow values may be introduced in fused image due to the combination of different wavelet coefficients from different source images. Secondly, as the result of the down-sampling process,

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

traditional wavelet-transform-based methods are no longer shift invariant. This will result in a problem that the activity level of a coefficient may incorrectly reflect the transformed content of that point and thus has an influence on the determination of wavelet coefficients. Thirdly, most wavelet-transform-based methods are linear, and some edge points in source images are apt to be washed out because of the operation of low-pass filtering. Fourthly, the false edge problem still occurs in the fused result. To remedy the second drawback, a shift-invariant wavelet transform was proposed and further improved by the dual tree complex wavelet transform [7] [8].

De and Chanda presented a fusion method based on multi-resolution morphological wavelet [3]. However, their method not only introduces underflow values and position errors, but also still has the false edge problem. Later, Lin and Huang proposed a dynamic-segmented morphological wavelets method which is modified from De and Chanda's method by slimming down the occurrence of underflow values and position errors [4]. Lin and Huang also proposed a fusion method called dynamic-segmented cut and paste (or DSCP for short), which completely overcomes the problem of underflow values and position errors and the result is better than that of De and Chanda's method. However, the performance of Lin and Huang's method depends on the block selection algorithm. Blocks with improper sizes may contain blurred area pixels. Consequently, their method may generate a falsely fused result.

III. A New Focus Measure and Window Selection Algorithm

Our new focus measure is based on two observations: (1) false edges usually exist in blurred images; (2) it is difficult to justify if a pixel is in an in-focus image without considering its surrounding area containing edge information.

False edge problem is that extra edge points may exist in out-of-focus images due to the blurring effect. By using the focus-measuring methods reported in the literature, those pixels around a false edge in an out-of-focus image will have high activity levels or large energy values, and thus will be preserved in the fused result. To better understand what we call "false edge problem", we take Fig. 1 as an example. Fig. 1(a) and Fig. 1(b) are two images of the same scene and focused in different parts. The corresponding gray-level values within the white-line enclosed areas are shown in Figs. 1(c) and 1(d), respectively. Notice that the gray-level values of the edge points detected by Sobel operator are enclosed by bold lines. As we can see that there is an edge line appearing in the third column of the out-of-focus area 1(d) while there is no edge line appearing in the same position of the in-focus area 1(c). If the design of a focus measure is based on the pixel currently being considered and its eight neighboring pixels, the focus-measuring value of the pixel in defocused image will be larger than that of the pixel in well-focused image. Therefore, the pixel in defocused image will be falsely preserved in the fused image. Our first observation is that edge information about defocused objects must be excluded from consideration for focus measuring.

Our second observation is that edge information about well-focused objects is important to image fusion. For example, as shown in Fig. 2(a) and 2(b), image 2(a) is out-of-focus while image 2(b) is in-focus of the same scene. If we crop the same area from the two images as shown in Figs. 2(c) and 2(d), it is very difficult to identify which portion is cropped from the in-focus image and which one is cropped from the defocused image because there is no edge information for helping judgment. If we expand the cropped areas in the two images at the same pace until one of the areas contains some edge information, then we can easily identify which cropped area belongs to the in-focus image. The above observation leads to an important idea, that is, the size of a focus-measuring window should not be fixed and should be large enough to include edge information in well-focused image to achieve a good fusion result.

Notice that not all edge points are useful for judging whether an area of image is in-focus or not, especially for those edge points that appear in two or more different source images. An edge point which exists in only one source image is called a "discriminative edge point." Given a set of multi-focus images of the same scene, we can find a set of discriminative edge points. The discriminative edge points for a set of multi-focus images of the same scene may include false edge points as shown in Fig. 1 (d) and the edge points which originally exist in the focused area of some image but do not exist in the same area of other defocused images as shown in Fig. 1 (c). A set of discriminative edge points S_d for a set of multi-focus images of the same scene can be constructed as follows:

$$S_d = \bigcup_{k=1}^{N} S_d^k \tag{1}$$

$$S_d^k = I_{EP}^k - \bigcap_{i=1}^{N} I_{EP}^i \tag{2}$$

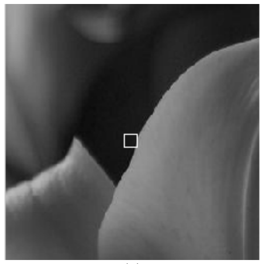
where N is the total number of source images, and I_{EP}^i represents the set of edge points detected from source image I_i.

Let $S_z(x, y)$ represent the window with center at (x, y) and size of $z*z$; $S_w(u, v)$ represents the window with center at (u, v) and size $w*w$; $f(u',v')$ and $f(u,v)$ are the pixel gray-values at (u',v') and (u,v), respectively. Given a discriminative edge point (u, v) in image I_i, (u, v) is associated with a salient weight $E_w^i(u,v)$ within window $S_w(u, v)$. Salient weight $E_w^i(u,v)$ can be computed by averaging the non-zero squared differences between $f(u',v')$ and $f(u,v)$ for all (u',v') in $S_w(u, v)$ with $(u',v') \neq (u,v)$. We define the focus measure $M_z^i(x,y)$ at a point (x, y) in image I_i using window $S_z(x, y)$ as the sum of the salient weights associated with all discriminative edge points within window $S_z(x, y)$. That is, $M_z^i(x,y)$ can be computed by using the following equations:

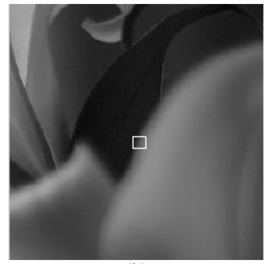
$$M_z^i(x,y) = \sum_{\forall (u,v)\in S_z(x,y)\cap S_d} E_w^i(u,v) \tag{3}$$

$$E_w^i(u,v) = (|D_w(u,v)|)^{-1} \sum_{\forall d\in D_w(u,v)} d \tag{4}$$

$$D_w(u,v) = \{d \mid [f(u',v') - f(u,v)]^2 = d \neq 0 \;\;,\;\; \forall (u',v') \in S_w(u,v)\} \tag{5}$$



(a)



(b)

23	23	23	24	24	24	25	25	27	24
22	23	23	23	24	24	24	24	26	23
22	22	23	23	23	24	24	24	25	21
22	22	22	23	23	24	24	24	25	19
23	23	22	22	22	23	24	25	31	17
22	22	22	22	23	24	25	25	31	20
21	21	21	22	23	24	25	26	28	25
21	21	21	22	23	24	26	27	21	33
21	21	21	22	23	24	25	26	15	34
23	23	22	22	22	23	24	25	15	83

(c)

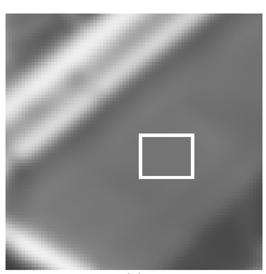
32	39	49	59	69	79	88	94	102	106
34	41	52	64	74	84	92	97	105	109
37	45	57	69	79	88	96	100	106	110
40	48	60	72	83	92	99	103	107	111
44	54	68	80	88	94	100	104	109	113
45	55	69	81	89	95	100	105	110	114
48	58	72	83	90	96	102	106	112	116
52	61	75	86	93	98	104	108	114	118
56	65	78	89	96	101	106	109	116	120
60	69	82	92	98	103	107	111	118	122

(d)

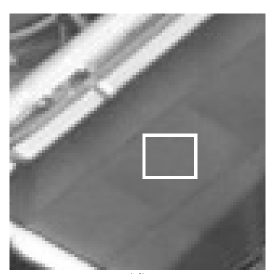
Figure 1. Pixel values in the designated rectangles of a focused area and its corresponding defocused area.

(a)

(b)



(c)



(d)

Figure 2. Two corresponding parts in defocused and focused images.

Based on the new focus measure defined above, we present our new fusion scheme as follows:

Algorithm: Image Fusion

Input: A set of multi-focus images $I_1, I_2, \ldots, I_N$

Output: A fused image I_{fused}

1. Find the set of discriminative edge points S_d from N multi-focus images using equations (1) and (2).
2. For each location (x, y) in the fused image, find the maximum window $S_z(x, y)$ having a window-size less than Z_{max} such that all pixels in $S_z(x, y)$ are contained in

the well focused area of a source image by using the window selection algorithm to be described later.

3. For every source image I_i, compute the focus measure $M_z^i(x, y)$ using equation (3) for pixel at location (x, y) using the window $S_z(x, y)$ found in step (2).

4. Assume that $M_z^k(x, y) \geq M_z^i(x, y)$ for all i =1, 2, …, N and $i \neq k$. Assume that the pixel at location (x, y) of source image I_k has the maximum focus measure among all pixels in N source images at this same location. Then, place the pixel value $f(x, y)$ at location (x, y) of source image I_k in the same location of the fused image. If maximum focus measure appears in several source images, then place the average pixel value of the pixels having maximum focus measure in the (x, y) location of the fused image.

5. Repeat steps (2), (3), and (4) until all pixel values of I_{fused} are found. Return I_{fused} .

The window selection algorithm is presented below.

Algorithm: Window selection

Input: A location (x, y) and a set of multi-focus images I_1, I_2, …, I_N

Output: An appropriate window $S_z(x, y)$ centered at (x, y)

1. Let Z_{max} be the maximum window size and s be the step size of each window expansion.

2. Find the smallest window $S_M(x, y)$ containing at least one discriminative edge point. If $M \geq Z_{max}$, then return $S_M(x, y)$ and exit.

3. Expand the current window by a step size s to become a new current window denoted by $S_z(x, y)$. If $Z \geq Z_{max}$, then return $S_{z-s}(x, y)$ and exit.

4. Compute the gradient magnitudes for all pixels within the current window in all source images. The gradient magnitudes within the current window of source image I_i form a matrix M_i whose dimension is the same as the size of current window. For every pair of matrices M_i and M_j, compute the matrix difference M_i-M_j . If the elements in M_i-M_j are all positive or all negative, then go to step (3).

In step (4) of the above window selection algorithm, we compute the matrix difference M_i-M_j. If the elements in M_i-M_j are all positive, it implies that the pixels within the window associated with M_i in image I_i are in a well-focused area. If the elements in M_i-M_j are all negative, it implies that the pixels within the window associated with M_j in image I_j are in a well-focused area. The purpose of our window selection algorithm is to find the maximum window centered at (x, y) within which all pixels belong to a well-focused area in some source image.

IV. EXPERIMENTAL RESULTS

In our experiments, we use three images of size 512*512, as shown in Fig 3, to compare performance of our fusion scheme with those of Chanda's method (MMWF method for short), and Lin's method (DSCP method for short), respectively [3][4]. Two partly-blurred images were created from each reference image. We use PSNR of the fused images with respect to its reference image as the performance measure of fusion results. A larger PSNR value represents a fusion result with better quality. The PSNR is computed as follows:

$$PSNR = 10 \times \log(\frac{255^2}{MSE}) \tag{6}$$

$$MSE = \frac{\sum_{n=1}^{image\ \ size}(R_n - F_n)^2}{image\ \ size} \tag{7}$$

where MSE is mean square error, R_n and F_n are the gray-level intensities of the nth pixel in the reference image and the fused image, respectively. Fig. 4 shows the fused results for three sets of images. The images in the first row are out-of-focus in the upper parts. The images in the second row are out-of-focus in the lower parts. The third, fourth, and fifth rows are fused results obtained from Chanda's MMWF, Lin's DSCP, and our fusion schemes, respectively. Table I shows the PSNR values of the fused results obtained from our fusion scheme by using fixed window-size and adaptive window-size with maximum window size ranging from 21 to 101 pixels for the three test images shown in Fig. 3. As we can see that adaptive window-size approach has much better performance than the fixed window-size approach. In addition, the performance of adaptive window-size approach in our fusion scheme seems very stable despite of the variations in window-size. Table II shows the comparisons of performances of the three fusion schemes. Chanda's MMWF methos is better than Lin's DSCP method in image 1. However, Lin's DSCP method is better than Chanda's MMWF method in both image 2 and image 3. Our fusion scheme outperforms Lin's and Chanda's methods in all three images.

V. CONCLUSION

In this paper, we propose a new fusion scheme for multi-focus images based on a new focus measure. Our proposed focus measure evaluates the salient weights of discriminative edge points within a focus-measuring window so that the false edge problem caused by blurring effect can be reduced. In our fusion scheme, an automatic window selection method is also proposed to select feasible focus-measuring windows for focus measure computation. Experimental results show that our fusion scheme outperforms Chanda's MMWF method and Lin's DSCP method.

Reference image 1 Reference image 2 Reference image 3

Figure 3. Reference images used in our experiment.

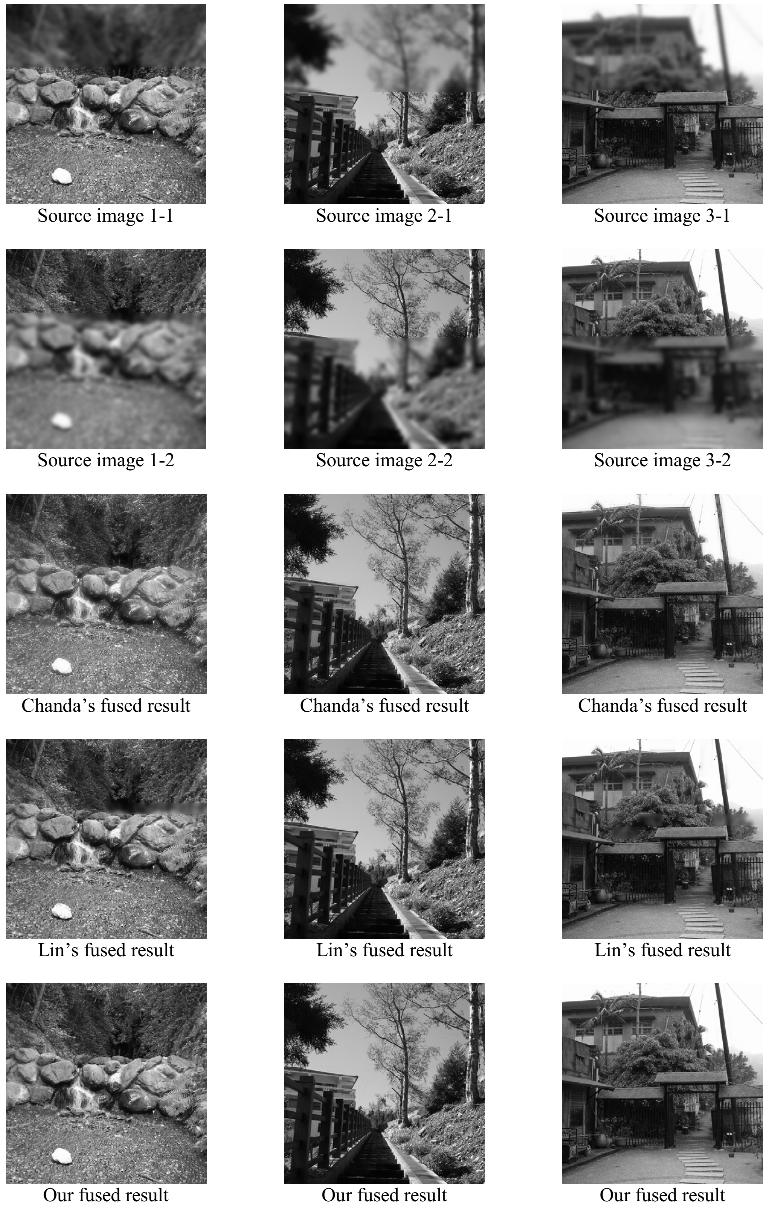


Figure 4. Test images and the fused results obtained from our method and other two methods.

TABLE I.

COMPARISON OF PSNR VALUES OF FUSED RESULTS OBTAINED FROM OUR FUSION SCHEME USING FIXED AND ADAPTIVE WINDOW-SIZE. THE MAXIMUM WINDOW-SIZE METHOD IS FROM 21 TO 101 PIXELS.

Image index	Image 1		Image 2		Image 3	
Scheme / Z_{max}	Fixed window	Adaptive window	Fixed window	Adaptive window	Fixed window	Adaptive window
21	51.34	55.19	46.36	52.61	49.39	51.62
41	46.15	55.19	42.62	52.64	42.48	51.79
61	43.87	55.19	41.72	52.65	42.75	51.79
81	41.07	55.19	41.70	52.65	42.42	51.79
101	39.66	55.19	41.59	52.65	40.61	51.79

TABLE II.
COMPARISON OF PSNR VALUES OF FUSED RESULTS OBTAINED FROM OUR FUSION SCHEME USING FIXED AND ADAPTIVE WINDOW-SIZE. THE MAXIMUM WINDOW-SIZE METHOD IS FROM 21 TO 101 PIXELS

Method / Images	MMWF (level 2)	DSCP (DCT)	Our scheme
Image 1	37.27	33.30	55.19
Image 2	34.51	45.83	52.65
Image 3	32.87	35.50	51.79

REFERENCES

[1] O. Rockinger, “Pixel - Level Fusion of Image Sequences using Wavelet Frames,” *In: Proceedings of the 16 th Leeds Applied Shape Research workshop*, Leeds University Press. 1996.

[2] P.J. Burt, R.J. Lolezynski, “Enhanced image capture through fusion,” *In: Proceedings of the Fourth International Conference on Computer Vision*, Berlin, Germany, pp. 173-182, 1993.

[3] I. De, B. Chanda, “A simple and efficient algorithm for multifocus image fusion using morphological wavelets,” *Signal Processing*, vol 86(5), pp 924-936, 2006.

[4] P.L. Lin, P.Y. Huang, “Fusion methods based on dynamic-segmented morphological wavelet or cut and paste for multifocus images,” *Signal Processing*, vol 88(6), pp. 1511-1527, 2008.

[5] A. Toet, L.J. van Ruyven, J.M. Valeton, “Merging thermal and visual images by a contrast pyramid,” *Optical Eng*, vol 28, pp. 789-792, 1989.

[6] A. Toet, “Multiscale contrast enhancement with applications to image fusion,” *Optical Eng*, vol 31, pp. 1026-1031,1992.

[7] O. Rockinger, “Image sequence fusion using a shift invariant wavelet transform,” *IEEE Trans. Image processing*, vol 3, pp. 288-291, 1997.

[8] P. Hill, N. Canagarajah, D. Bull, “Image fusion using complex wavelets,” *Complex Proc. 13th British Machine Vision Conference*, University of Cardiff, 2002.

[9] Pajares. Gonzalo, Manuel de la Cruz. Jesús, “A wavelet-based image fusion tutorial,” *Pattern Recognition*, vol 37, pp. 1855-1872, 2004.