Multi-focus image fusion based on spatial frequency and morphological operators

Bin Yang (杨 斌) and Shutao Li (李树涛)

College of Electrical and Information Engineering, Hunan University, Changsha 410082

Received January 26, 2007

A new multi-focus image fusion method using spatial frequency (SF) and morphological operators is proposed. Firstly, the focus regions are detected using SF criteria. Then the morphological operators are used to smooth the regions. Finally the fused image is constructed by cutting and pasting the focused regions of the source images. Experimental results show that the proposed algorithm performs well for multi-focus image fusion.

OCIS codes: 100.2000, 350.2660.

The camera, like human eyes, has a limited depth of field. In other words, it focuses on one level and the objects in front or behind are often blurred. An image which is in focus everywhere contains more information than one which is focused on one object and is useful in many fields such as biomedical imaging, microscopic imaging, remote sensing, computer vision, and robotics. A possible way to attain images which are in focus everywhere is by image fusion where one acquires a series of pictures with different focus settings and fuses them to produce an image with an extended depth of field[1]. For most pixel-level image fusion algorithms, the pixels are treated independently and these algorithms split the relationship among pixels[2−4]. Some region-based image fusion algorithms perform well, but suffer from complex implementation[5]. We propose an effective algorithm suitable for combining multi-focus images of a scene. The spatial frequency (SF) of a pixel's neighbor block is used to judge its sharpness[6,7] and morphological opening and closing are used to correct during post-processing. Finally, the fused image is obtained by combining the focused regions.

We assume two registered images, $\mathbf{I}_1$ and $\mathbf{I}_2$, with different focuses to be fused. The algorithm consists of following steps:

Step 1. Calculating the SF of each pixel within a 5×5 window in $\mathbf{I}_1$ and $\mathbf{I}_2$, denoted by SF_1 and SF_2, respectively. SF is calculated from

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

where RF and CF are the row frequency and column frequency,

$$\mathrm{RF}(m,n) = 1/5 \times \sqrt{\sum_{i=-2}^{2}\sum_{j=-2}^{2}(F(m+i,n+j) - F(m+i,n+j-1))^2},$$

$$\mathrm{CF}(m,n) = 1/5 \times \sqrt{\sum_{i=-2}^{2}\sum_{j=-2}^{2}(F(m+i,n+j) - F(m+i,n+j-1))^2},$$

and F is the gray image of size $M \times N$ and $F(m,n)$ denotes the gray value at each pixel position (m,n). The values of SF are affected by image sharpness[7].

Step 2. Comparing the values SF_1 and SF_2 to determine which pixel is in focus. The logical matrix Z (essentially a binary image) is constructed as

$$Z(m,n) = \begin{cases} 1 & \mathrm{SF}_1(m,n) > \mathrm{SF}_2(m,n) \\ 0 & \text{otherwise} \end{cases}, \quad (2)$$

'1' in Z indicates that the pixel at position (m,n) in image $\mathbf{I}_1$ is in focus otherwise the pixel in $\mathbf{I}_2$ is in focus.

Step 3. However, determinating by SF alone is insufficient to discern all the focused pixels. There are thin protrusions, narrow breaks, thin gulfs, small holes etc. in Z. To correct these defects morphological opening and closing, constructed by combining dilation and erosion, are employed[8]. Opening, denoted as $Z \circ B$, is simply erosion of Z by structure element B, followed by dilation of the result by B. It removes thin connections and thin protrusions. Closing, denoted as $Z \bullet B$, is dilation followed by erosion. It joins narrow breaks and fills long thin gulfs. Holes larger than B cannot be removed simply using opening and closing operators. In practice, small holes are always judged incorrectly therefore a threshold is set to remove the holes smaller than the threshold. Opening and closing are again performed to smooth object contours.

Step 4. The fusion image is then constructed as

$$F(m,n) = \begin{cases} I_1(m,n) & Z(m,n) = 1 \\ I_2(m,n) & \text{otherwise} \end{cases}. \quad (3)$$

The proposed method is compared with the wavelet-based method[3], which is implemented in the following way. Firstly, the scaled images and detail images are obtained by using the wavelet transform. The wavelet basis 'db1', together with a decomposition of 3, is used. Scaled images and detail images are then combined by choosing the pixel with the maximum absolute value. Consistency verification is implemented in this step. Specifically, if the center pixel value comes from image $\mathbf{I}_1$ while the majority of the surrounding pixel values come from image $\mathbf{I}_2$, the center pixel value is switched to that of image $\mathbf{I}_2$. Finally, the inverse wavelet transform is implemented to recover the fused image[3]. In

1671-7694/2007/080452-02

our algorithm, the structure element B is a logical '1' 5×5 matrix and the threshold is set to 1000. The experiment is done in the environment of AMD Sempron CPU 2.17 GHz with a 512 MB RAM PC operating under Windows XP and Matlab 6.5. Four pairs of multi-focus images are used to test our algorithm against

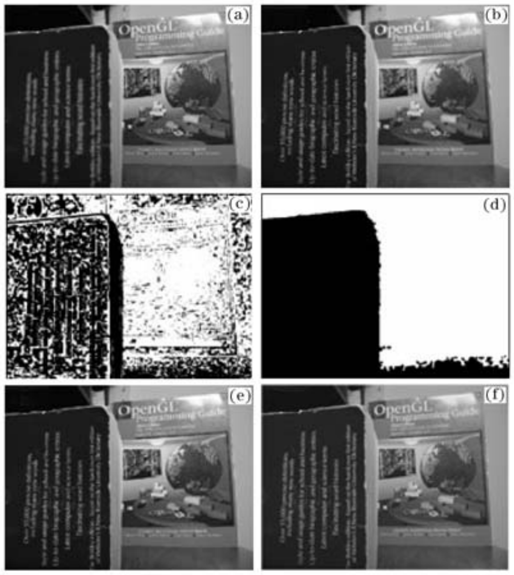

Fig. 1. Fusion example 1. (a) Focus on the right book (size 480×640); (b) focus on the left book (size 480×640); (c) Z matrix in step 2; (d) Z matrix in step 3; (e) fusion result using our algorithm; (f) fusion result using wavelet-based method.

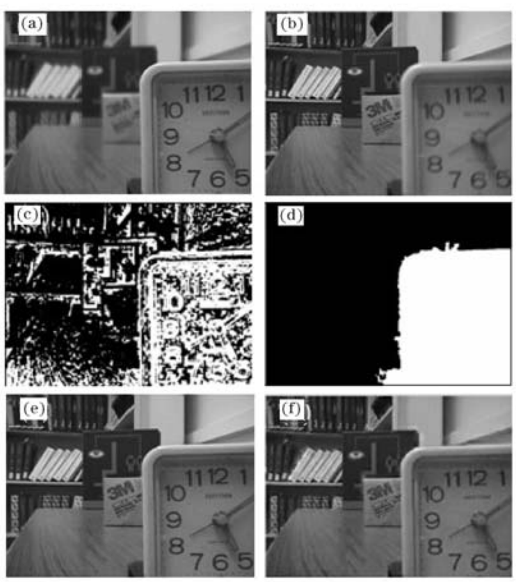

Fig. 2. Fusion example 2. (a) Focus on the right clock (size 480×640); (b) focus on the left bookshelf (size 480×640); (c) Z matrix in step 2; (d) Z matrix in step 3; (e) fusion result using our algorithm; (f) fusion result using wavelet-based method.

Table 1. Performance of Different Fusion Methods. EN: Entropy; STD: Standard Difference; SIM: Similarity; DWT: Wavelet-Based Method; SF-MOR: Proposed Method

Test Image		Pair 1	Pair 2	Pair 3	Pair 4
EN	DWT	7.3436	7.1876	7.4536	7.4450
	SF-MOR	7.3494	7.2769	7.4583	7.6248
STD	DWT	59.9199	44.1591	47.3949	69.6140
	SF-MOR	61.4038	46.8276	48.3748	71.2814
SIM	DWT	0.8810	0.8661	0.8801	0.8347
	SF-MOR	0.9196	0.9128	0.9542	0.8986

the wavelet-based method. The first two pairs of source images and the fused results are shown in Figs. 1 and 2. Their sizes are both 480×640. Carefully comparing the results, we can see that the wavelet method loses sharpness and exhibits prominent blocking artifacts (the left books in Figs. 1(e) and (f), and the left bookshelves in Figs. 2(e) and (f)). To evaluate the performance of the proposed method objectively, three criteria, entropy (EN), standard deviation (STD), and similarity (SIM)[9] are used. For those criteria, larger values indicate better fusion results. From Table 1, we can observe that the proposed algorithm outperforms the wavelet-based method.

The proposed image fusion algorithm resembles the manual cut-and-paste method, which is often used to obtain a standard fused image. From the experimental results, we conclude that the proposed algorithm performs better on multi-focus image fusion than the conventional wavelet-based method. However, the generality of the parameters should be investigated further because of the diversity of the image content.

This work was supported by the National Natural Science Foundation of China (No. 60402024) and the Program for New Century Excellent Talents in University (No. NCET-2005). S. Li is the author to whom the correspondence should be addressed, his e-mail address is shutao_li@yahoo.com.cn.

References

1. W. B. Seales and S. Dutta, Proc. SPIE **2905,** 227 (1996).
2. G. Pajares and J. M. Cruz, Pattern Recogn. **37,** 1855 (2004).
3. Z. Zhang and R. S. Blum, Proc. IEEE **87,** 1315 (1999).
4. G. Liu, Z. Jing, S. Sun, J. Li, Z. Li, and H. Leung, Chin. Opt. Lett. **2,** 386 (2004).
5. G. Piella, Inform. Fusion **4,** 259 (2003).
6. Z. Li, Z. Jing, and S. Sun, Chin. Opt. Lett. **2,** 82 (2004).
7. S. Li, J. Kwok, and Y. Wang, Inform. Fusion **2,** 169 (2001).
8. S. Mukhopadhyay and B. Chanda, Pattern Recogn. **34,** 1939 (2001).
9. I. De and B. Chand, Signal Process. **86,** 924 (2006).