Optics Communications 284 (2011) 80–87

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Multi-focus image fusion using a bilateral gradient-based sharpness criterion

Jing Tian [a,*], Li Chen [a], Lihong Ma [b], Weiyu Yu [c]

[a] School of Computer Science and Technology, Wuhan University of Science and Technology, 430081, PR China
[b] Guangdong Key Lab. of Wireless Network and Terminal, School of Electronic and Information Engineering, South China University of Technology, Guangzhou, 510641, PR China
[c] School of Electronic and Information Engineering, South China University of Technology, Guangzhou, 510641, PR China

ARTICLE INFO

Article history:
Received 21 May 2010
Received in revised form 16 August 2010
Accepted 23 August 2010

Keywords:
Image fusion
Gradient
Sharpness

ABSTRACT

The aim of multi-focus image fusion is to combine multiple images with different focuses for enhancing the perception of a scene. The challenge is to how evaluate the local content (sharp) information of the input images. To tackle the above challenge, a new bilateral sharpness criterion is proposed to exploit both the strength and the phase coherence that are evaluated using the gradient information of the images. Then the proposed bilateral sharpness criterion is further exploited to perform weighted aggregation of multi-focus images. Extensive experimental results are provided to demonstrate that the proposed bilateral sharpness criterion outperforms conventional seven sharpness criterions.

1. Introduction

Multi-focus image fusion aims to fuse two or more images that are captured using different camera settings (e.g., at different focuses) of the same scene to form another superior image with uniform focus and sharp content [1]. Due to the limited depth-of-focus of optical lenses, it is usually impossible to acquire an image that contains all relevant objects in-focus. Therefore, the multi-focus image fusion technique is desirable to create a single image where all objects are in focus. It has been used as an effective tool for many important applications, such as medical imaging and microscopic imaging.

The basic idea of performing image fusion is to choose the clearer image pixels (or blocks/regions) from source images, or adaptively average observed images according to their respective sharpness, to construct the fused image, since the optical out-of-focus is the major source of image quality degradations in multi-focus image fusion [2]. Various techniques have been developed in the literature, including spatial frequency [3,4], energy of gradient [5], phase coherence [6,7], etc. A detailed review will be provided in Section 3. The other kind of popular approach is to exploit multi-resolution analysis [8–10]. Its idea is to perform multi-resolution decomposition on each input image, integrate the decompositions to form a composite representation based on certain fusion rules, and then reconstruct the fused image via an inverse multi-resolution transform. This scheme is usually complicated and time-consuming to implement [11].

The key challenge of multi-focus image fusion is how to evaluate the blur of each image and then select information from the most informative (sharp) image [12–14]. To tackle the above challenge, the gradient information of the images is examined in this paper to propose a new criterion to measure the local sharpness of the image by considering bilateral statistics of gradient information–the strength and the phase coherence. To be more specific, an image with higher strength and higher phase coherence is considered as sharper and more informative; consequently, it should contribute more to the fused image. Furthermore, the proposed criterion is exploited to develop a weighted aggregation approach to perform image fusion.

The paper is organized as follows. Section 2 presents the formulation for the multi-focus image fusion problem. Section 3 reviews various sharpness criterions developed in the literature. Then a bilateral sharpness criterion using statistics of image's gradient information is proposed in Section 4. Extensive experimental results are presented in Section 5 to compare the proposed bilateral sharpness criterion with conventional criterions. Finally, Section 6 concludes this paper.

2. Problem formulation

Given a set (say, N) of 2-D images $I_1(r,c), I_2(r,c), \ldots, I_N(r,c)$, which have been acquired using different imaging setting and aligned well, the goal of multi-focus image fusion is to integrate the information content of the individual images into a single fused image $f(r,c)$. A simple but effective method for image fusion is to perform a simple normalized aggregation of the images. This can be mathematically expressed as

$$f(r,c) = \frac{1}{N}\sum_{n=1}^{N} I_n(r,c). \tag{1}$$

The main drawback with this simple normalized aggregation of images is that all information content within the images are treated the same. Therefore, important image regions, which yield more detailed information (edge or high-frequency) and are more informative, are

* Corresponding author.
E-mail addresses: eejtian@gmail.com (J. Tian), chenli@ieee.org (L. Chen), eelhma@scut.edu.cn (L. Ma), yuweiyu@scut.edu.cn (W. Yu).

0030-4018/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.optcom.2010.08.085

treated no differently than unimportant regions. To overcome this drawback, a normalized weighted aggregation approach to image fusion can be used. It can be mathematically expressed as

$$f(r,c) = \frac{\sum_{n=1}^{N} w_n(r,c) I_n(r,c)}{\sum_{n=1}^{N} w_n(r,c)}. \quad (2)$$

where $w_n(r,c)$ is the weight assigned to information content at (r,c) in the n-th image.

The choice of the weighting scheme (i.e., the determination of $w_n(r,c)$ in Eq. (2)) is very crucial to the performance of the image fusion algorithm. Since the optical out-of-focus is the major source of image quality degradations in multi-focus image fusion [12–14], the weighting scheme is required to be sensitive to the blur of each image and be locally adaptive to image's content. In view of this, a brief review on the existing sharpness measures will be provided in Section 3, followed by the development of the proposed bilateral sharpness criterion using two kinds of statistics of image's gradient information in Section 4.

3. Existing sharpness criterions

In view of that the determination of the weighting scheme is the key issue of designing multi-focus image fusion techniques, a review on the existing sharpness criterions is provided in this section. Since the degree of de-focus varies inversely with the amount of high spatial frequency energy present in the spatial frequency spectrum, the amount of high-frequency information (corresponding to edge information in images) is usually used as the basis to measure the degree of image's blur [14]. More specifically, the well focused image has sharper edges and is expected to have higher frequency content than those that are blurred. In the following analysis, denote $I(r,c)$ be the intensity value at the position (r,c) of the image I.

- Variance [12]. For an $M\times N$ block of the image, its variance is defined as [12]

$$S_{VAR} = \frac{1}{M\times N}\sum_r \sum_c (I(r,c)-\mu)^2, \quad (3)$$

where μ is the mean intensity value of the image block and it is defined as $\mu = \frac{1}{M\times N}\sum_r \sum_c (I(r,c))$.

- Energy of image gradient [12]. For an $M\times N$ block of the image, it is measured as [12]

$$S_{EG} = \sum_r \sum_c \left(I_r^2 + I_c^2\right), \quad (4)$$

where I_r and I_c represent image gradients at the row and column directions, respectively. They are usually defined as $I_r = I(r+1,c) - I(r,c)$ and $I_c = I(r,c+1) - I(r,c)$.

- Tenenbaum [12]. For an $M\times N$ block of the image, it is measured as [12]

$$S_{TNG} = \sum_r \sum_c (\nabla I(r,c))^2, \quad (5)$$

where $\nabla I(r,c) = \sqrt{I_r^2 + I_c^2}$, in which I_r and I_c are gradients (obtained using Sobel operators) along the row and column directions, respectively.

- Energy of Laplacian [12]. For an $M\times N$ block of the image, it is measured as [12]

$$S_{EL} = \sum_r \sum_c \left(\nabla^2 I(r,c)\right)^2, \quad (6)$$

where $\nabla^2 I(r,c)$ represents image gradient obtained by Laplacian operator $[-1,-4,-1;-4,20,-4;-1,-4,-1]$.

- Sum-modified-Laplacian [12]. It differs from the usual Laplacian operator in that the absolute values of the partial second derivatives are summed instead of their actual values. That is, it can be mathematically expressed as [12]

$$S_{SML} = \sum_r \sum_c \left(\nabla^2 I(r,c)\right)^2, \quad (7)$$

where $\nabla^2 I(r,c) = |2I(r,c) - I(r+1,c) - I(r-1,c)| + |2I(r,c) - I(r,c+1) - I(r,c-1)|$.

- Frequency selective weighted median filter [14]. It measures the sharpness of the image as

$$S_{FSWM} = \sum_r \sum_c \left(I_r^2 + I_c^2\right), \quad (8)$$

where

$$\begin{aligned} I_r = {} & \mathrm{med}\{I(r-1,c), I(r,c), I(r+1,c)\} - \tfrac{1}{2}\mathrm{med}\{I(r-3,c), I(r-2,c), I(r-1,c)\} \\ & - \tfrac{1}{2}\mathrm{med}\{I(r+1,c), I(r+2,c), I(r+3,c)\}, \\ I_c = {} & \mathrm{med}\{I(r,c-1), I(r,c), I(r,c+1)\} - \tfrac{1}{2}\mathrm{med}\{I(r,c-3), I(r,c-2), I(r,c-1)\} \\ & - \tfrac{1}{2}\mathrm{med}\{I(r,c+1), I(r,c+2), I(r,c+3)\}. \end{aligned}$$

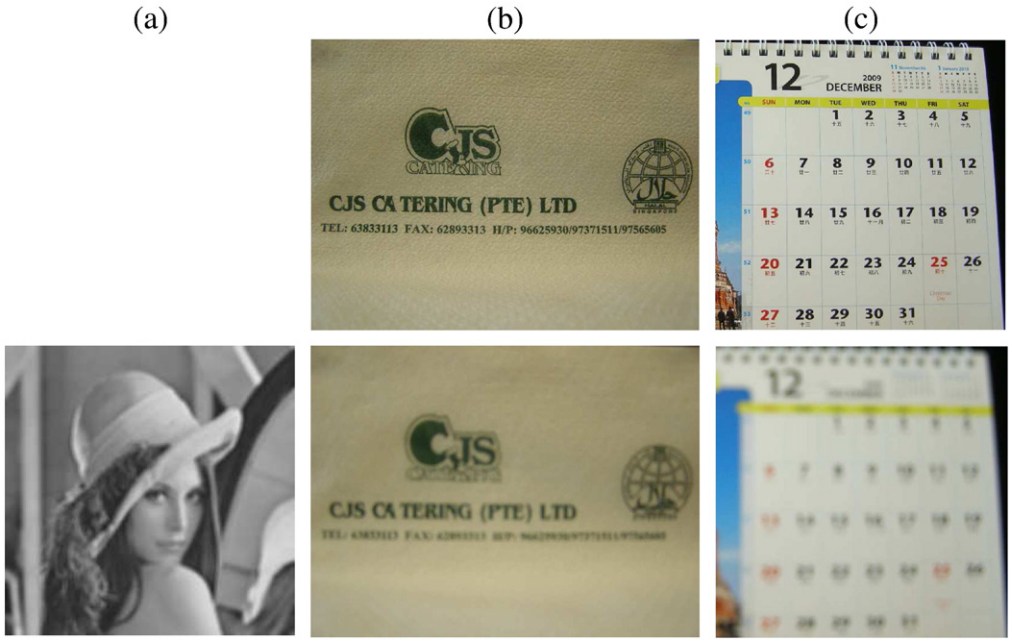

Fig. 1. Three sets of test images: (a) Lena, (b) Tissue and (c) Calendar.

This criterion is fairly robust to against the noise incurred in the observed images.

- Phase coherence model [6]. It is consistent to the perceptual significance of the image, and it can be determined at a particular position (r,c) as

$$S_{PCM}(r,c) = \frac{1}{2}\sum_{\theta} |(h(r,c,\theta)sin(\theta))^2 + (h(r,c,\theta)cos(\theta))^2| + \frac{1}{2}\sqrt{4\left(\sum_{\theta}(h(r,c,\theta)sin(\theta)h(r,c,\theta)cos(\theta))\right)^2 + \left(\sum_{\theta}[(h(r,c,\theta)cos(\theta))^2-(h(r,c,\theta)sin(\theta))^2]\right)^2}, \tag{9}$$

where

$$h(r,c,\theta) = \frac{\sum_n W(r,c,\theta)|A_n(r,c,\theta)\Delta\varphi_n(r,c,\theta)|}{\sum_n A_n(r,c,\theta) + \xi}, \tag{10}$$

$$\Delta\varphi_n(r,c,\theta) = cos(\varphi_n(r,c,\theta)-\overline{\varphi}_n(r,c,\theta))-|sin(\varphi_n(r,c,\theta)-\overline{\varphi}_n(r,c,\theta))|, \tag{11}$$

in which W represents the frequency spread weighting factor, A_n and φ_n represent the amplitude and phase at the wavelet scale n, respectively, $\overline{\varphi}_n$ represents the weighted mean phase, ξ is a small constant used to avoid the division by zero. All of these parameters are as same as that used in [15].

4. Proposed bilateral gradient-based sharpness criterion

The statistics of image's gradient is examined in this section to propose a new sharpness measurement criterion, which exploits bilateral statistics of image's gradient information. Image structure can be measured effectively by using the image gradients. Consider an image of interest $I(r,c)$. The gradient covariance matrix of a region within an $M\times N$ local window is defined as [16]

$$C = \begin{pmatrix} \sum_W I_r^2(r,c) & \sum_W I_r(r,c)I_c(r,c) \\ \sum_W I_r(r,c)I_c(r,c) & \sum_W I_c^2(r,c) \end{pmatrix}, \tag{12}$$

where $I_r(r,c)$ and $I_c(r,c)$ represent image's gradient at the row and column directions, respectively. Furthermore, the above gradient covariance matrix can be decomposed as

$$C = VDV^T = (v_1 \quad v_2)\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}\begin{pmatrix}v_1^T\\ v_2^T\end{pmatrix}, \tag{13}$$

where V represents a 2×2 matrix whose column vectors are eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, D denotes the 2×2 diagonal matrix whose diagonal elements are eigenvalues λ_1 and λ_2 ($\lambda_1>\lambda_2$) that correspond to eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively, and the superscript T denotes the transpose.

The geometrical structure at a pixel in an image can be described by the eigenvalues λ_1 and λ_2 of the above gradient covariance matrix [17]. Motivated by this, the first criterion is proposed to measure the strength of the image's gradient, which is defined as

$$A(r,c) = \lambda_1-\lambda_2. \tag{14}$$

On the other hand, local phase coherence is consistent with the perceptual significance of image's characteristics. This has been supported by the physiological evidence that showed high human perception response to signal characteristics with high local phase coherence [15]. Another advantage is the fact that it is insensitive magnitude variations caused by illumination conditions or noise

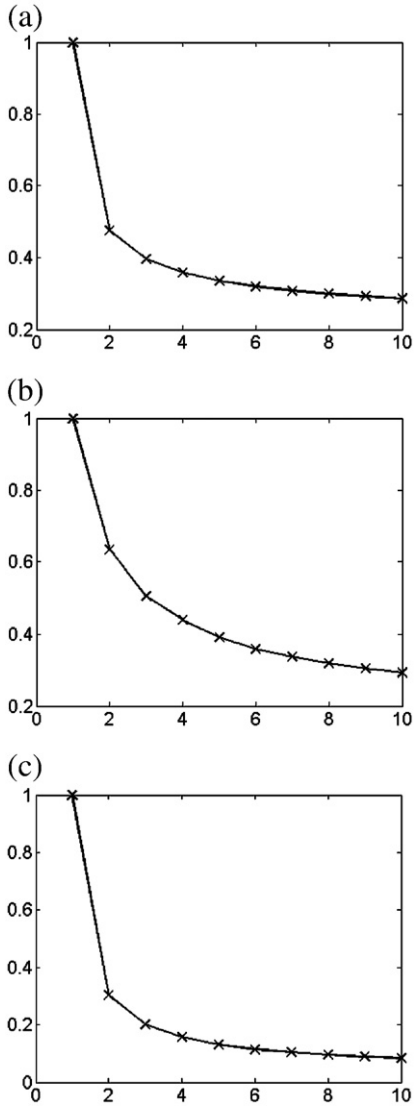

Fig. 2. Plots of responses of Eqs. (14))–((16)) for iteration numbers, where a 5×5 average blur is iteratively applied to the test image Lena. (a) criterion (14); (b) criterion (15); (c) criterion (16).

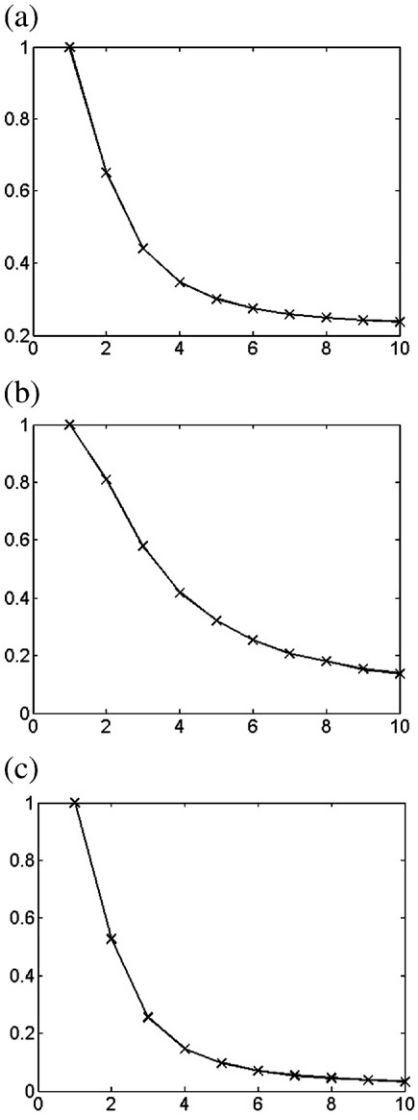

Fig. 3. Plots of responses of Eqs. (14)–(16) for iteration numbers, where an average blur, whose size starts from 5×5 and increases with a step of 2, is iteratively applied to the test image Lena. (a) criterion (14); (b) criterion (15); (c) criterion (16).

incurred in image signals. In view of this, the second criterion is to consider the phase coherence of the image's gradient, that is

$$P(r,c) = -cos\left(\theta(r,c) - \overline{\theta}(r,c)\right), \tag{15}$$

where $\theta(r,c)$ is the phase information at the position (r,c) determined by the principle eigenvector $\mathbf{v}_1$ associated with the largest eigenvalue λ_1 defined as (13), $\overline{\theta}_{r,c}$ is the average of phases of the neighboring positions. This measure achieves the maximal value when the local phase coherence is worst, which is usually caused by an edge.

Finally, the above two criterions (14) and (15) are jointly considered to develop a bilateral sharpness criterion as

$$S_{BSC} = A^{\alpha}(r,c)P^{\beta}(r,c), \tag{16}$$

where α and β are two factors to adjust contributions of two criterions.

Experiments are conducted to justify that the proposed sharpness criterions (14), (15) and (16) are fairly good to measure the sharpness of the image. Therefore, they can be used as the weighting criterion of the image fusion approach in Eq. (2).

Table 1
The responses of criterions (14)–(16) for test images with different manually-adjusted focus levels.

Test image		Criterion (14)	Criterion (15)	Criterion (16)
Tissue	Sharp	6067	0.4375	2654
	Blurred	2738	0.3726	1020
Calendar	Sharp	17,784	0.4375	7781
	Blurred	2168	0.3285	712

- In the first experiment, a 5×5 average blur is iteratively applied to the well-known test image Lena (as shown in Fig. 1(a)), the resulting criterion values are recorded, normalized with their respective maximum values, and presented in Fig. 2, where one can see that the sharpness values drop steadily, when the images are more blurred.
- In the second experiment, an average blur, whose size starts from 5×5 and increases with a step of 2, is iteratively applied to the well-known test image Lena (as shown in Fig. 1(a)). The purpose is change the size of the averaging blur filter to generate a number of

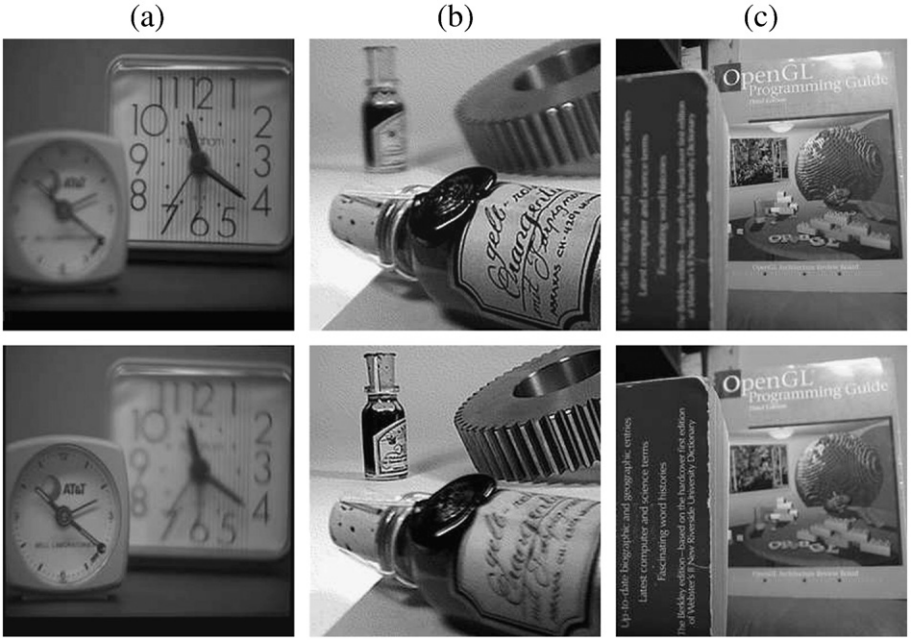

Fig. 4. Three sets of test images: (a) Clock, (b) Bottle, and (c) Book.

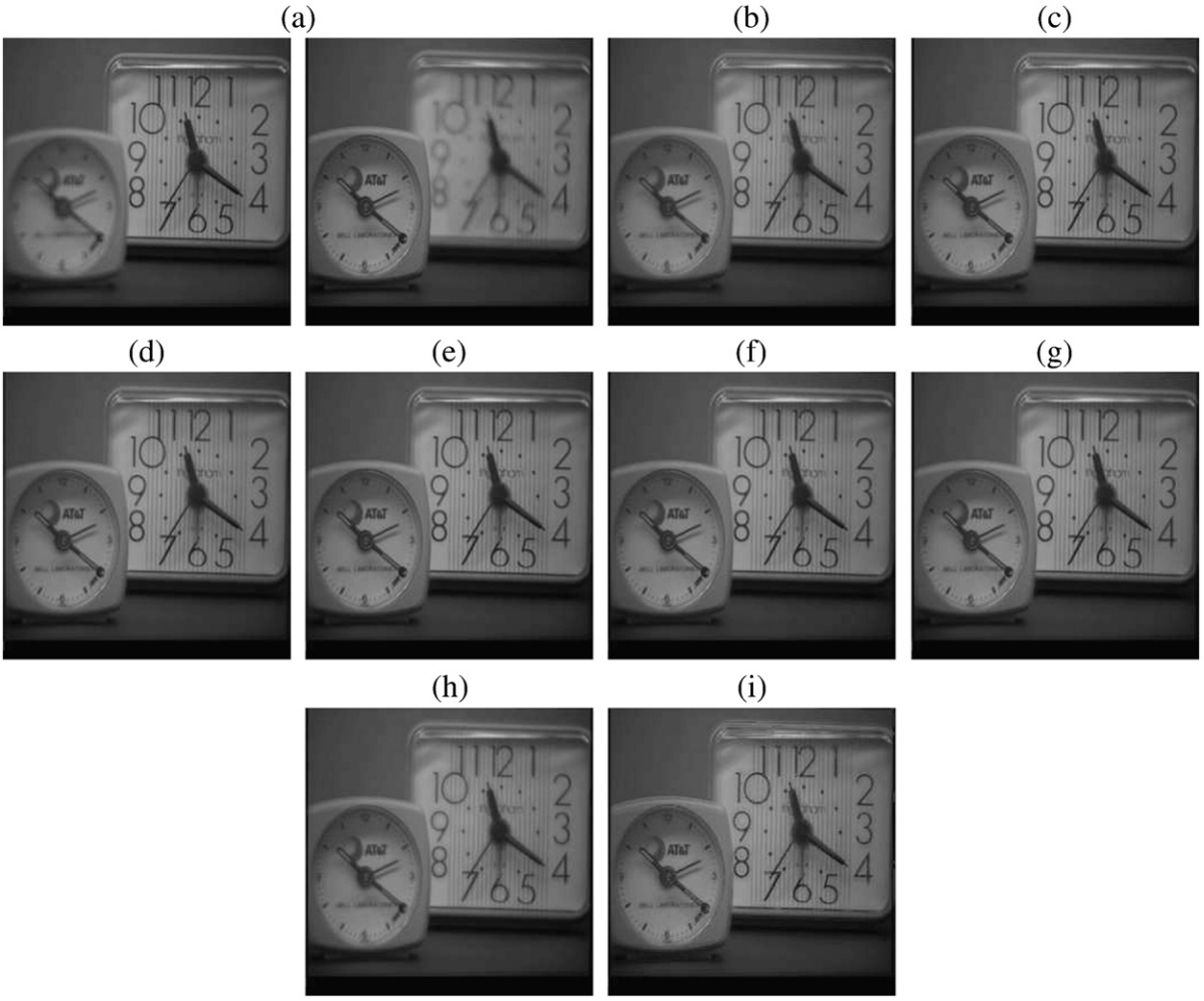

Fig. 5. A comparison of fused images (Clock) using different sharpness criterions: (a) two source images; (b)–(h) are results obtained using criterions defined in Eqs. (3)–(9), respectively; (i) proposed bilateral gradient-based sharpness criterion (16).

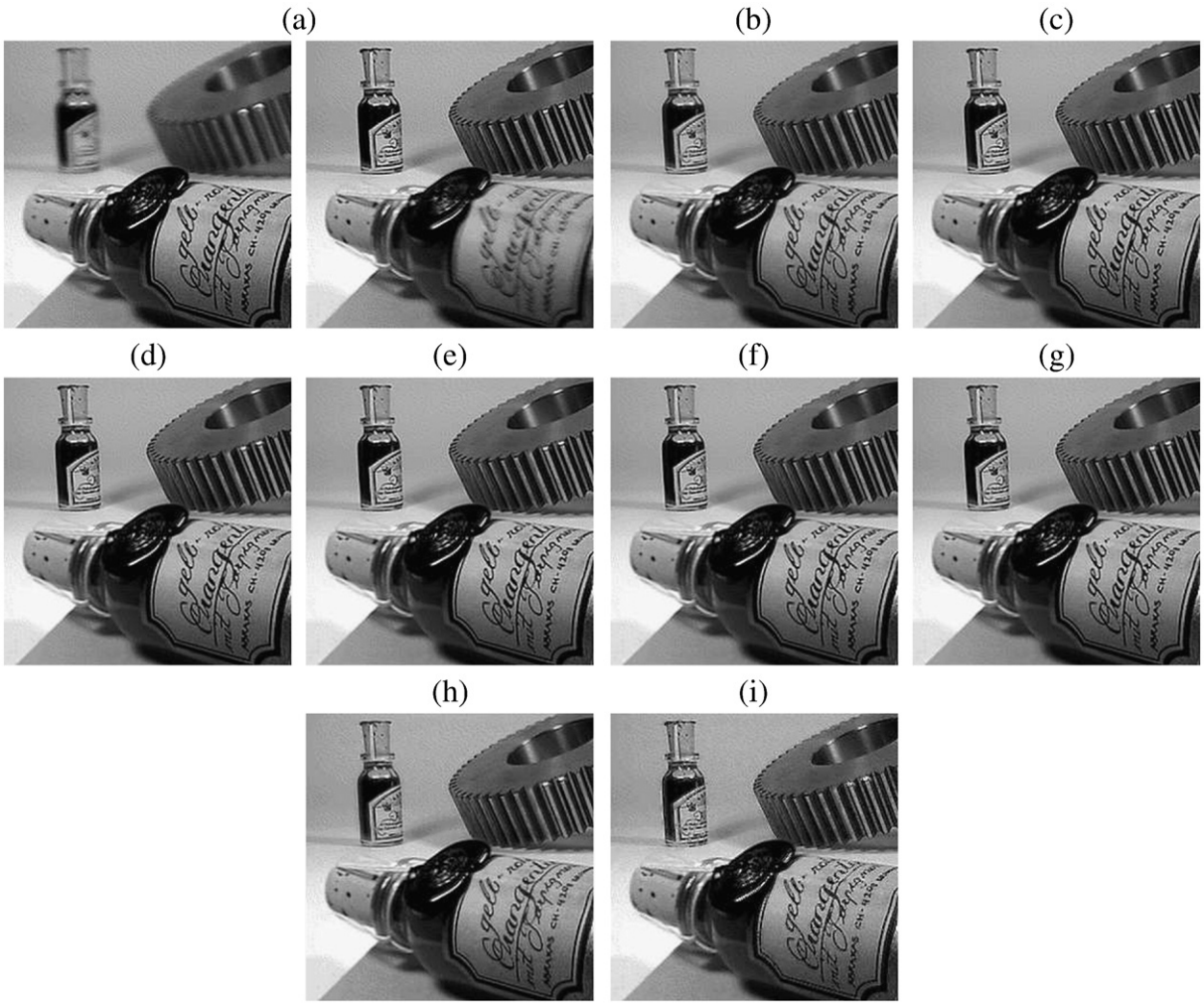

Fig. 6. A comparison of fused images (Bottle) using different sharpness criterions: (a) two source images; (b)–(h) are results obtained using criterions defined in (3)–(9), respectively; (i) proposed bilateral gradient-based sharpness criterion (16).

images with different blurring levels. More specifically, the image generated using a larger blur size is considered to be less sharper than that generated using a smaller blur size. Then the resulting sharpness values are recorded, normalized with their respective maximum values, and presented in Fig. 3, where one can see that the values drop steadily, when the images are more blurred.

- In the third experiment, two sets of images are captured using a camera Olympus SP-500 with different manually-adjusted focus levels; one is in focus, while the other one is out of focus, as shown in Fig. 1(b) and (c). Their respective sharpness values are compared in Table 1, where one can see that the values of blurred images are smaller than that of the sharp images.

5. Experimental results

Experiments are conducted to compare the performance of the proposed bilateral sharpness criterion with other seven sharpness criterions (3)–(9), by individually incorporating them as the weighting scheme of Eq. (2) to perform image fusion. The parameters of the proposed criterion (see Eq. (16)) are experimentally set as $\alpha=1$ and $\beta=0.5$. Also, the size of the neighborhood is set to be 5×5. The above parameter settings are experimentally selected. Our setting might not be globally optimal, but this setting yields fairly good performance in our simulations.

The first experiment is to conduct image fusion using three sets of images with different focus levels: 256×256 Clock, 256×256 Bottle, and 256×256 Book, as shown in Fig. 4. The comparison of various fused images are presented in Figs. 5–7, respectively. One can see that the fused images obtained using the proposed method yield better image quality than that of conventional approaches.

Since there is no ground truth image to evaluate the performance of image fusion algorithms using PSNR, two image quality evaluation criterions (without need for ground truth) are used to provide objective performance comparison in this paper. These two metrics are: i) mutual information metric [18] and ii) spatial frequency metric [19], where larger metrics values indicate better image quality. The objective performance comparisons are presented in Tables 2 and 3, where one can see that the proposed approach always outperforms other seven conventional sharpness criterions by producing the best objective performance.

The second experiment is to compare the computational complexity of image fusion approach using various sharpness criterions. These image fusion approaches are implemented using the Matlab programming language and run on a PC with a Pentium 1.66 GHz CPU and a 2048 MB RAM. Ten experiments are conducted for each of the above-mentioned approaches, then their respective average run-times are compared in Table 4, where the computational complexity of the image fusion approach with the incorporation of the proposed sharpness criterion is comparable to that of other sharpness criterions.

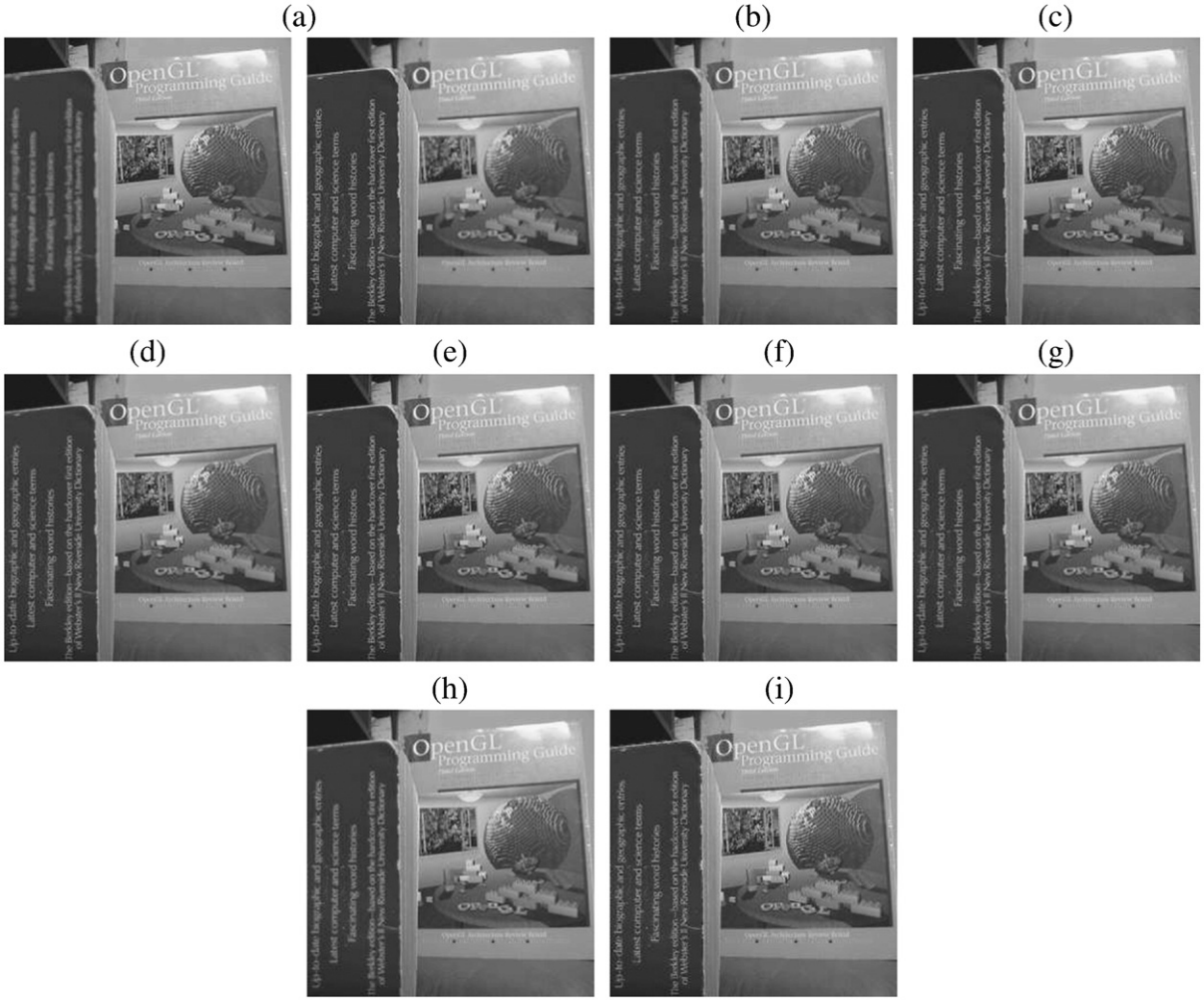

Fig. 7. A comparison of fused images (Book) using different sharpness criterions: (a) two source images; (b)–(h) are results obtained using criterions defined in (3)–(9), respectively; (i) proposed bilateral gradient-based sharpness criterion (16).

Table 2
The mutual information performance [18] comparison of image fusion using various sharpness criterions.

Test image	Criterion (3)	Criterion (4)	Criterion (5)	Criterion (6)	Criterion (7)	Criterion (8)	Criterion (9)	Proposed criterion (16)
Clock	7.18	7.49	7.38	7.65	7.66	7.56	7.05	8.52
Bottle	6.09	6.36	6.32	6.39	6.38	6.38	6.25	8.54
Book	7.92	8.23	7.96	8.43	8.45	7.99	7.93	9.36

Table 3
The spatial frequency performance [19] comparison of image fusion using various sharpness criterions.

Test image	Criterion (3)	Criterion (4)	Criterion (5)	Criterion (6)	Criterion (7)	Criterion (8)	Criterion (9)	Proposed criterion (16)
Clock	11.07	12.55	12.04	12.62	12.89	12.63	10.05	14.29
Bottle	33.50	38.53	37.99	38.95	38.85	38.61	33.14	41.92
Book	23.25	26.80	24.53	27.74	27.79	24.87	21.34	27.76

6. Conclusions

A multi-focus image fusion approach using a new sharpness criterion that depends on statistics of image's gradient information is proposed in this paper. The proposed approach exploits a bilateral sharpness criterion to adaptively perform image fusion by selecting most informative (sharp) information from the input images. The proposed bilateral sharpness criterion outperforms seven conventional sharpness criterions, as verified in our extensive experiments using four sets of test images under two objective metrics.

There are a few directions for future research. First, the proposed approach is conducted in spatial domain only in this paper, it could be further exploited into multi-resolution analysis [9]. Second, the noises incurred in observed images is neglected in this paper. However,

Table 4
The run-time (in seconds) comparison of image fusion using various sharpness criterions.

	Criterion (3)	Criterion (4)	Criterion (5)	Criterion (6)	Criterion (7)	Criterion (8)	Criterion (9)	Proposed criterion (16)
Run time	0.09	0.04	0.04	0.03	0.24	9.08	4.50	3.96

images are usually corrupted with noises in image acquisition or image communication. Such noise may cause miscalculation of sharpness values, which consequently introduce significant errors in the results of image fusion. Therefore, how to perform image fusion for noisy images would be a challenge, where the sharpness criterion is required to be robust to handle the noisy images [14,20]. Third, the neighborhood used in the proposed approach to calculate the sharpness value is experimentally selected to be a square-shape window. Certainly, it should be adaptive to local image content [21]. All of the above issues need future investigation to further improve the approach proposed in this paper.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. 60705012, 60972133).

References

[1] A.A. Goshtasby, S. Nikolov, Information Fusion 8 (Apr. 2007) 114.
[2] H.B. Mitchell, Image fusion: theories, techniques and applications, Springer, 2010.
[3] S. Li, J.T. Kwok, Y. Wang, Information Fusion 2 (Sep. 2001) 169.
[4] S. Li, B. Yang, Image Vision Computing 26 (Jul. 2008) 971.
[5] H. Li, B.S. Manjunath, S.K. Mitra, Graphical Models and Image Processing 57 (May 1995) 235.
[6] A. Wong, W. Bishop, Pattern Recognition Letters 29 (Feb. 2008) 173.
[7] R. Hassen, Z. Wang, M. Salama, Multifocus image fusion using local phase coherence measurement, Proc. Int. Conf. on Image Analysis and Recognition, Jul. 2009, p. 54.
[8] G. Pajares, J.M. Cruz, A wavelet-based image fusion tutorial, Pattern Recognition vol. 37 (Sept. 2004) 1855.
[9] S. Arivazhagan, L. Ganesan, T.G. Subash Kumar, Signal Image and Video Processing 3 (Jun. 2009) 137.
[10] J. Tian, L. Chen, Multi-focus image fusion using wavelet-domain statistics, Proc. IEEE Int. Conf. on Image Processing, 2010.
[11] Z. Wang, Y. Ma, J. Gu, Pattern Recognition (Jun. 2010) 2003.
[12] W. Huang, Z. Jing, Pattern Recognition Letters 28 (Apr. 2007) 493.
[13] Y. Zhang, L. Ge, Digital Signal Processing 19 (Mar. 2009) 186.
[14] V. Aslantas, R. Kurban, Optics Communications 282 (Aug. 2009) 3231.
[15] P. Kovesi, Videre: A Journal of Computer Vision Research 1 (1999) 2.
[16] A. Agrawal, R. Raskar, R. Chellappa, Edge suppression by gradient field transformation using cross-projection tensors, Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, 2006, p. 2301.
[17] C.-Y. Wee, R. Paramesran, Information Sciences 177 (Jun. 2007) 2533.
[18] G. Qu, D. Zhang, P. Yan, Electronics Letters 38 (Mar 2002) 313.
[19] S. Li, B. Yang, Pattern Recognition Letters 29 (Jul. 2008) 1295.
[20] V. Aslantas, R. Kurban, Evaluation of criterion functions for the fusion of multi-focus noisy images, Proc. Int. Conf. on Signal Processing and Communications Applications, Apr. 2009, p. 492.
[21] P.W. Huang, C.-I. Chen, P.-L. Lin, Multi-focus image fusion based on salient edge information within adaptive focus-measuring windows, Proc. Int. Conf. on Systems, Man and Cybernetics, Oct. 2009, p. 2589.