





## Research Article

# Infrared and Visible Image Fusion Combining Interesting Region Detection and Nonsampled Contourlet Transform

Kangjian He <sup>1</sup>, Dongming Zhou <sup>1</sup>, Xuejie Zhang <sup>1</sup>, and Rencan Nie <sup>1,2</sup>

<sup>1</sup>School of Information, Yunnan University, Kunming 650500, China

<sup>2</sup>School of Automation, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Dongming Zhou; [zhoudm@ynu.edu.cn](mailto:zhoudm@ynu.edu.cn) and Xuejie Zhang; [xjzhang@ynu.edu.cn](mailto:xjzhang@ynu.edu.cn)

Received 14 August 2017; Revised 19 December 2017; Accepted 25 December 2017; Published 5 April 2018

Academic Editor: Calogero M. Oddo

Copyright © 2018 Kangjian He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The most fundamental purpose of infrared (IR) and visible (VI) image fusion is to integrate the useful information and produce a new image which has higher reliability and understandability for human or computer vision. In order to better preserve the interesting region and its corresponding detail information, a novel multiscale fusion scheme based on interesting region detection is proposed in this paper. Firstly, the MeanShift is used to detect the interesting region with the salient objects and the background region of IR and VI. Then the interesting regions are processed by the guided filter. Next, the nonsampled contourlet transform (NSCT) is used for background region decomposition of IR and VI to get a low-frequency and a series of high-frequency layers. An improved weighted average method based on per-pixel weighted average is used to fuse the low-frequency layer. The pulse-coupled neural network (PCNN) is used to fuse each high-frequency layer. Finally, the fused image is obtained by fusing the fused interesting region and the fused background region. Experimental results demonstrate that the proposed algorithm can integrate more background details as well as highlight the interesting region with the salient objects, which is superior to the conventional methods in objective quality evaluations and visual inspection.

## 1. Introduction

Image fusion is an important branch of information science, which has been widely used in many fields, such as bioinformatics, medical image processing, and military target visualization. Especially in military field, infrared (IR) and visible (VI) image fusion is important to military science technology, such as automatic military target detection and localization. As a hot image fusion field, it has attracted the attention of many researchers [1–7]. The key problem of IR and VI image fusion is to integrate and extract the feature information of the source images to produce a new image which is more reliable and understandable, and the fused image not only has the detailed texture information of VI image but also can highlight the target area in an IR image.

There are many different algorithms for the IR and VI image fusion that have been proposed and developed over the past few decades. The early fusion methods such as intensity-hue-saturation (IHS) and principal component analysis (PCA) were to process pixel values on spatial

domain, which were traditional classical methods, but the fusion effect was limited compared with other excellent fusion methods [8–10]. Many fusion methods based on multiscale transform (MST) have become popular in recent years, such as Laplacian pyramid (LP), wavelet transform (WT), discrete wavelet transform (DWT), and nonsampled contourlet transform (NSCT) [11–16]. Due to the excellent characteristics of the multiscale decomposition method, the MST-based method could get a good fusion effect compared with early fusion methods, such as NSCT-PCNN [17]. However, these methods usually failed to highlight the target information in the fused image. IR image target detection-based method is another popular IR and VI image fusion method; these methods detected the target region of the IR image firstly, then fused the background regions using other methods to get the fused background image, and finally fused the target region and background regions directly to get a new image. The advantages of these methods can fully retain the infrared target information in the fused image, but commonly, these infrared target regions

of the fused image will lack the corresponding detail information in the VI image. Our previous work proposed a fusion algorithm which was based on target extraction; it was useful to highlight the target in the infrared image due to the target region which was directly fused into the final image [8]. Taking into account the shortcomings of these algorithms, in order to overcome these problems, a novel IR and VI image fusion method is proposed in this paper. Compared with our previous work, we improved the accuracy of interesting region detection where it contains highlighting target and heat sources. In addition, in order to enrich the visible information in the interesting region, we also adopt fusion strategy to fuse them.

The first step of the proposed method is to detect the interesting region which contains significant target in the IR image by the MeanShift method. The MeanShift has many applications, such as clustering, discontinuity preserving smoothing, object contour detection, image segmentation, and nonrigid object tracking [15, 18]. We will use it to detect the interesting region with a significant infrared target from the background regions in the IR image. In order to fully retain and highlight the interesting region and significant target information in the fused image, the interesting region will be taken as a separate component and directly fused into the finally image. But the interesting region extracted from the IR image will lose the details of the corresponding region of the VI image. To solve this problem, we use the guided filter to fuse the interesting regions of IR and its corresponding VI image [19–21]. The interesting region of the IR image serves as the guidance image and the interesting region of VI image as the input image. The guided image filter was proposed in 2013 by He et al. [19]; the guided filter has many good characteristics, such as edge-preserving and image smoothing. So we use it to preserve the edge of the VI image; the produced interesting region contains the significant target information as well as the detail information.

Next, the background regions will be decomposed by nonsubsampling contourlet transform to get a low-frequency and a series of high-frequency layers. NSCT as an effective decomposition tool was proposed by Da Cunha et al. [16]. NSCT has many good properties of time-frequency localization, multidirection, and multiscale; therefore, it has been widely used in image fusion compared with other multiscale-based methods [22–24]. For the low-frequency layer, we proposed an improved weighted average method based on per-pixel weighted average. Due to the characteristics of the low-frequency layer (hazy image), the per-pixel weighted average based method is effective and will be detailedly described in the fusion rule section. For the high-frequency layers, the pulse-coupled neural network (PCNN) will be used to process each high-frequency layer. PCNN was proposed by Eckhorn et al. [25]. Since it was introduced, it has been widely used in the field of image processing, such as image segmentation, image enhancement, image edge detection, and image fusion [26, 27]. In the proposed method, the spatial frequency (SF) metric of the high-frequency layers will be used as external incentive information of the PCNN model, which makes it better to deal

with overexposed or weak exposure images and make the fusion result more suitable for human visual inspection.

The remaining sections of this paper are organized as follows: the related work and proposed methods are introduced in Section 2, including the interesting region detection and fusion, the background region fusion, and concrete fusion steps. Experimental result comparisons and analysis are given in Section 3. The conclusions are shown in Section 4.

## 2. Related Work and Proposed Methods

### 2.1. Related Work

**2.1.1. MeanShift Algorithm.** The most important function of the MeanShift is as a tool for computing probability density function in a set of data samples [28]. It has been widely used in discontinuity preserving smoothing, object contour detection, and image segmentation.

Given a finite number of data points  $x_1, \dots, x_n$  in the  $d$ -dimensional space  $R^d$ , a multivariate kernel density function is defined as

$$P(x) = \frac{1}{n} \sum_{i=1}^n K(x - x_i), \quad (1)$$

where

$$K(x) = c \prod_{i=1}^d k(x_i). \quad (2)$$

With the kernel  $K(x)$  being a bounded function with the following properties,

$$\begin{aligned} \int_{R^d} K(x) dx &= 1, \\ \int_{R^d} xK(x) dx &= 0, \\ \lim_{\|x\| \rightarrow \infty} \|x\|^d K(x) &= 0, \\ \int_{R^d} xx^T K(x) dx &= cI, \end{aligned} \quad (3)$$

where  $c$  is a constant and those of equations mean normalized, symmetric, and exponential weight decay, respectively. The normal kernel  $K(x)$  is computed by

$$K_N(x) = c \times \exp\left(-\frac{1}{2} \|x\|^2\right). \quad (4)$$

Estimate the kernel density gradient by

$$\nabla P(x) = \frac{1}{n} \sum_{i=1}^n \nabla K(x - x_i), \quad (5)$$

and using the normal kernel form, (5) can be rewritten as

$$\nabla P(x) = \frac{c}{n} \sum_{i=1}^n \nabla k_i = \frac{c}{n} \left[ \sum_{i=1}^n g_i \right] \cdot \left[ \frac{\sum_{i=1}^n x_i g_i}{\sum_{i=1}^n g_i} - x \right], \quad (6)$$

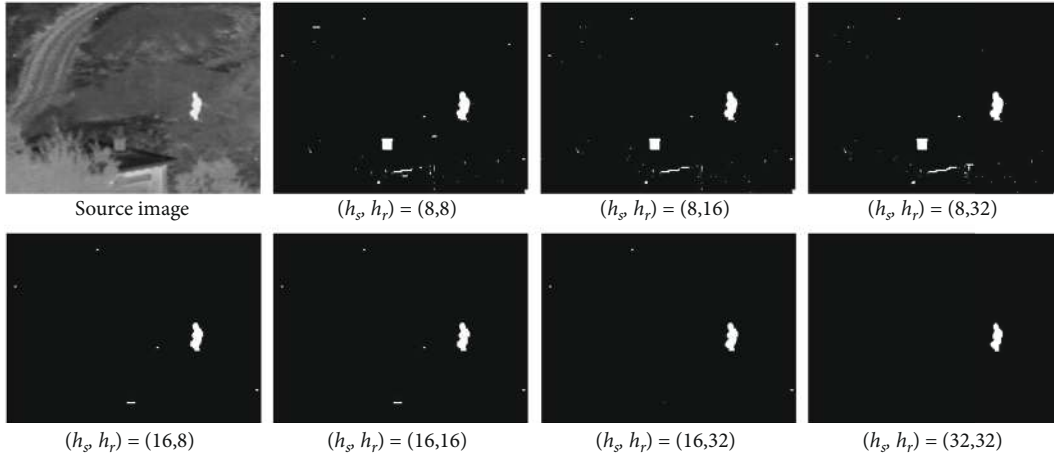


FIGURE 1: Interesting region detection with different bandwidths.

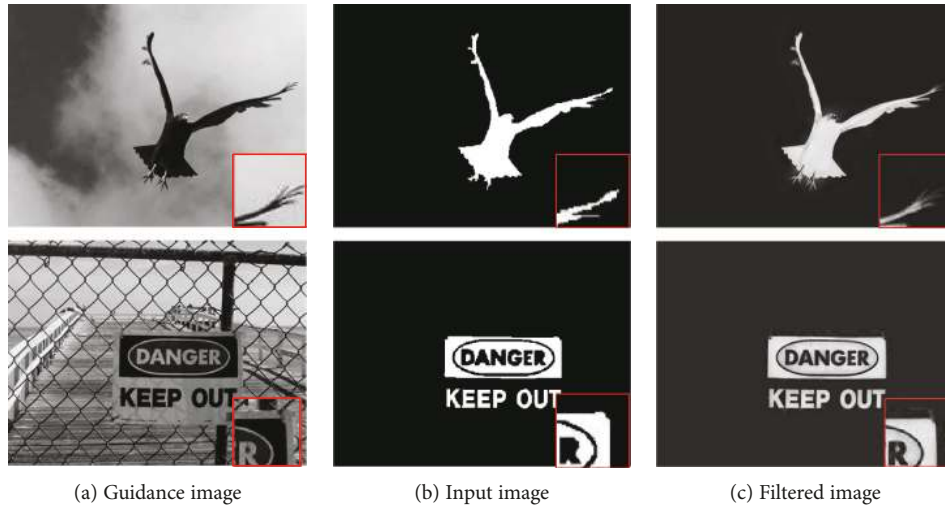


FIGURE 2: Two examples of the guided filter.

where  $g(x) = -k'(x)$ . We use the MeanShift to process the IR image, cluster the infrared target pixels, and obtain the interesting region of IR. For image clustering and segmentation, we treat the image as data points in the spatial and gray level domains; two radially symmetric kernels will be used which are defined as follows:

$$K(x) = C \cdot k_s \left( \left\| \frac{x^s}{h_s} \right\| \right) \cdot k_r \left( \left\| \frac{x^r}{h_r} \right\| \right), \quad (7)$$

where  $x^s$  is the spatial coordinate,  $x^r$  is the range of a feature vector in color space, and  $h_s$  and  $h_r$  are the employed kernel bandwidths. An example of the interesting region detection of IR with different bandwidths is given in Figure 1.

We can see from Figure 1 that the MeanShift method can effectively extract the IR image region which we interested in and the infrared target information accurately. Compared with IR, the interesting region of VI contains more detailed information. In order to highlight the interesting region of IR and enrich the details of the corresponding region of VI in the fused image, when the interesting regions of IR and VI are determined, the guided filter is used to fuse ones.

**2.1.2. Guided Filter.** The guided filter is an edge-preserving filter and can compute the filter output by considering the content of the guidance image. There are many good characteristics of the guided filter, especially in edge detail preservation [18, 19, 29]. The filtered output image is very similar to the input image, and it also contains both texture and detail information of the guidance image, as shown in Figure 2.

Supposing that the guidance image is  $I$ , the detail description of the guided filter is given as follows:

$$O_i = a_k I_i + b_k, \quad \forall i \in \omega_k, \quad (8)$$

where  $O$  is the linear transformation of  $I$ ,  $\omega_k$  is a local window, in which the pixel  $k$  is the center, and the coefficients  $a_k$  and  $b_k$  are constant, to make the input image and the output image as similar as possible; we minimize the variance between the output image  $O$  and the input image  $P$  as follows:

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k I_i + b_k - P)^2 + \epsilon a_k^2), \quad (9)$$

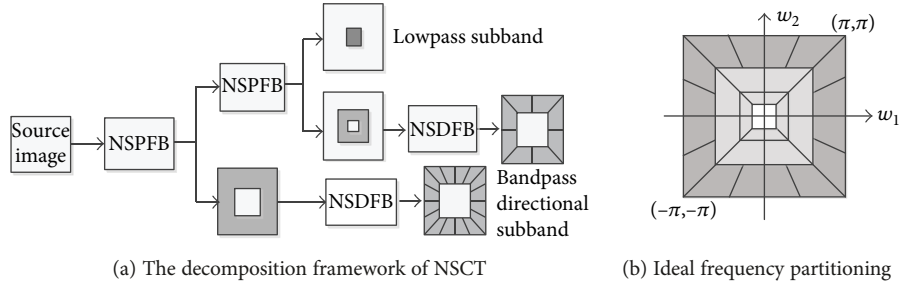


FIGURE 3: The construction of NSCT.

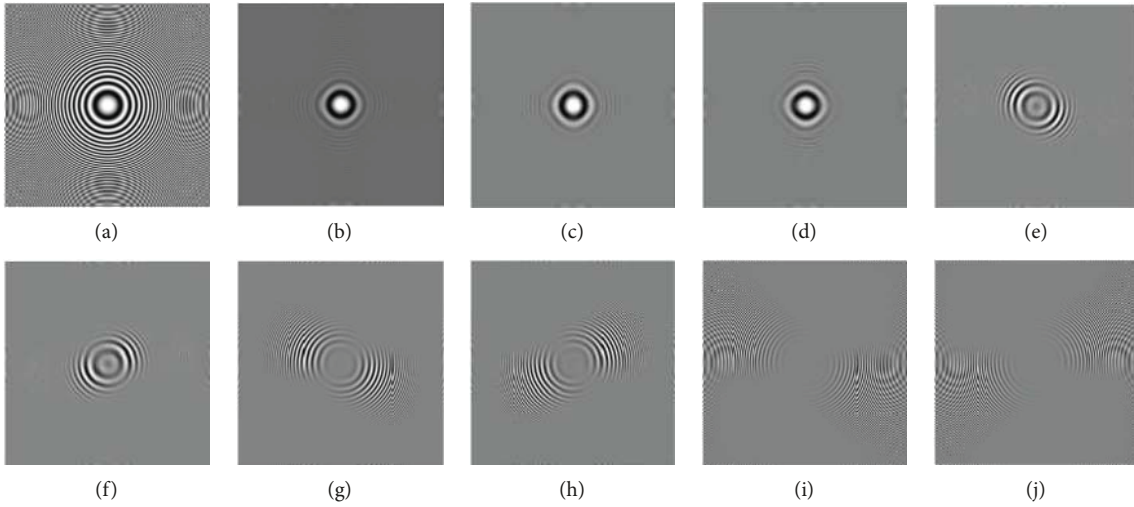


FIGURE 4: An example of NSCT decomposition. (a) Source image. (b) Low-frequency layer. (c)-(d) Level 1. (e)-(f) Level 2. (g)-(h) Level 3. (i)-(j) Level 4.

where  $\varepsilon$  is the regularization parameter and  $a_k$  and  $b_k$  are computed by

$$a_k = \frac{x(1/|\omega|) \sum_{i \in \omega_k} I_i P_i - \mu_k \bar{P}_K}{\delta_k + \varepsilon}, \quad (10)$$

$$b_k = \bar{P}_K - a_k \mu_k,$$

where  $a_k$  is the mean and  $b_k$  is the variance of the local window  $\omega_k$  in the image  $I$ ,  $|\omega|$  is the total number of pixels in the local window  $\omega_k$ , and  $\bar{P}_K$  is the mean of the input image  $P$  in the local window  $\omega_k$ . Figure 2 shows a set of examples of the guided filter.

It can be seen from Figure 2 that the guidance image contains a large number of detail and texture information. And the input image just contains significant regional information but lacks of detail, texture, and edge information. As can be seen from Figure 2(c), the output image of the guided filter is consistent with the input image, but it only contains detailed texture information in the corresponding region of the guidance image. This is also suitable to process the saliency target in the IR image and its corresponding region in the VI image. Through the guided filter, we can fuse the detail information into the interesting region in the IR image;

in this way, the produced new interesting region contains both salient object and detail information.

**2.1.3. Nonsampled Contourlet Transform (NSCT).** NSCT is a new two-dimensional image decomposition and analysis tool, which is derived from contourlet transform (CT) [16]. The construction of NSCT contains nonsampled pyramid filter banks (NSPFBB) and nonsampled directional filter banks (NSDFBB), which are shown in Figure 3.

It can be seen from Figure 3 that the source image can be decomposed by NSCT to get a low-frequency layer and a series of high-frequency layers. All obtained layers are the same size with the source image. Figure 4 shows an example of NSCT decomposition. In Figure 4, we decompose the source image into four layers, each of which is decomposed into four images in four directions; we select two images from each layer as shown in Figure 4. Figure 4(b) is the low-frequency layer; it can be seen that it contains only the low frequency information of the source image without high frequency details. From level 1 to level 4 are the high-frequency layers, which show the detail information from different levels to different levels.

**2.1.4. Pulse-Coupled Neural Network (PCNN).** The PCNN is a single-layered artificial neural network [25]. A basic neuron

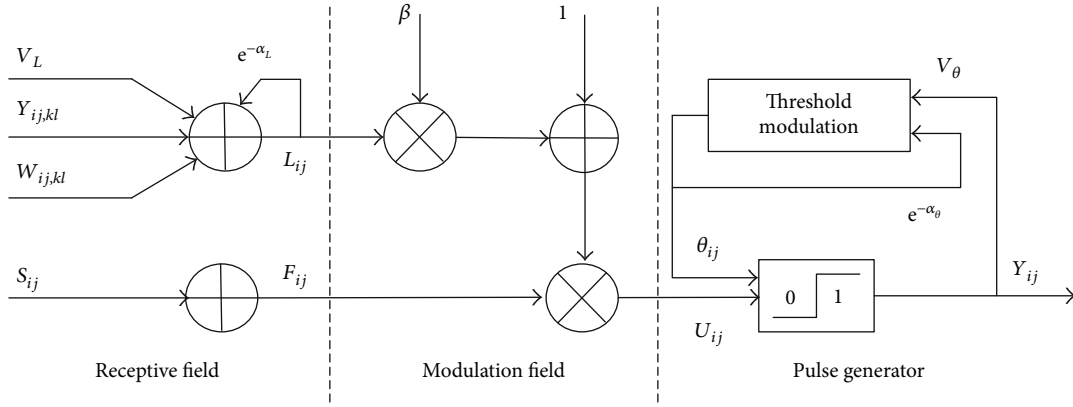


FIGURE 5: The typical structure of PCNN model.

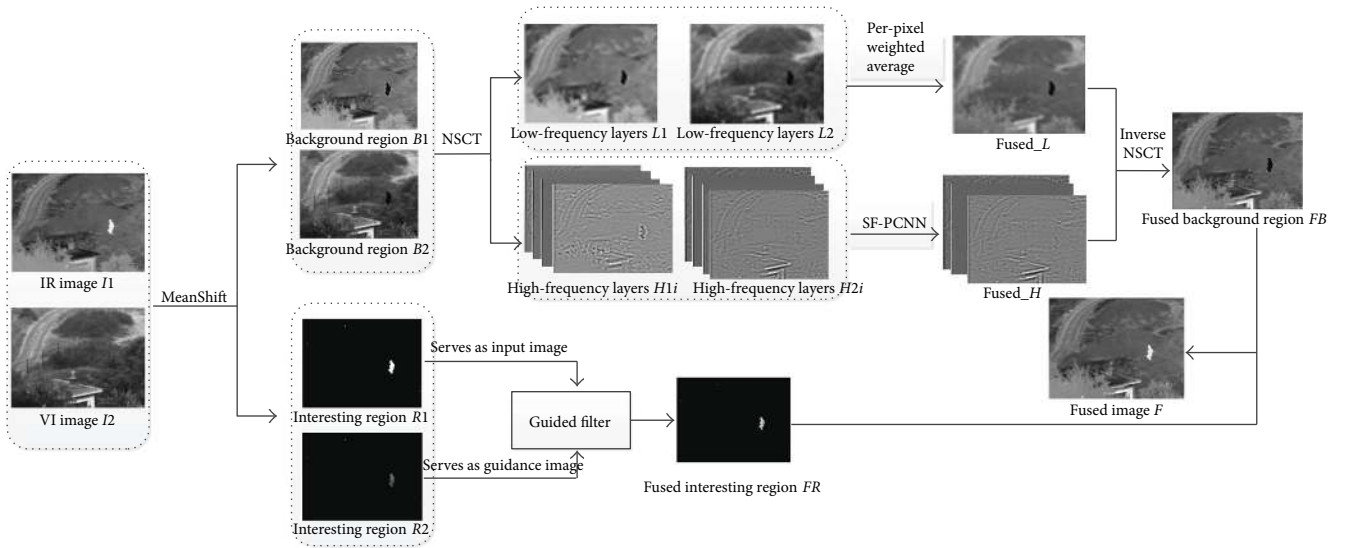


FIGURE 6: Schematic diagram of the proposed fusion method.

of PCNN contains the receptive field, the modulation field, and the pulse generator, which are shown in Figure 5.

The receptive field of PCNN can be described in detail as follows:

$$\begin{aligned} F_{ij}(n) &= V_F \sum_{k,l} \mathbf{W}_{kj} Y_{kl}(n-1) + F_{ij}(n) e^{-\alpha_F} + S_{ij}, \\ L_{ij}(n) &= e^{-\alpha_L} L_{ij}(n) + V_L \sum_{kl} \mathbf{M}_{kj} Y_{ijkl}(n-1), \end{aligned} \quad (11)$$

where  $S_{ij}$  is the input stimulus at pixel  $(i, j)$  of the source image,  $F_{ij}$  is the feeding input of it, matrices  $\mathbf{W}$  and  $\mathbf{B}$  are the constant synaptic weight,  $\alpha_F$  and  $\alpha_L$  are the time constants, and  $V_F$  and  $V_L$  are normalizing constants.

In modulation field, the internal state is controlled by linking strength  $\beta$ , which is given by

$$U_{ij}(n) = F_{ij}(n) [1 + \beta L_{ij}(n)], \quad (12)$$

where  $U_{ij}$  is the internal state of the neuron, which is created by modulating the feeding and linking channels.

The pulse generator field can be described as

$$\theta_{ij}(n) = e^{-\alpha_\theta} \theta_{ij}(n-1) + V_T Y_{ij}(n-1), \quad (13)$$

$$Y_{ij}(n) = \begin{cases} 1, & U_{ij}(n) > \theta_{ij}(n), \\ 0, & \text{otherwise,} \end{cases} \quad (14)$$

where  $Y_{ij}$  is the output of input  $S_{ij}$  and  $\theta_{ij}$  is the dynamic threshold of the neuron, which is used to compare with  $U_{ij}$ . It can be seen from (14), if  $U_{ij}$  is larger than  $\theta_{ij}$ , the output  $Y_{ij}$  of the neuron at  $(i, j)$  is 1, which we call the neuron is fired. The time matrix  $\mathbf{T}$  of the neuron fired can be described as follows:

$$\mathbf{T}_{ij}(n) = \begin{cases} n, & \text{if } Y_{ij}(n) = 1, \text{ for the first time,} \\ \mathbf{T}_{ij}(n-1), & \text{otherwise.} \end{cases} \quad (15)$$

**2.2. Proposed Method.** The proposed fusion algorithm framework is depicted in Figure 6. The first step in the proposed method is to detect the interesting region which



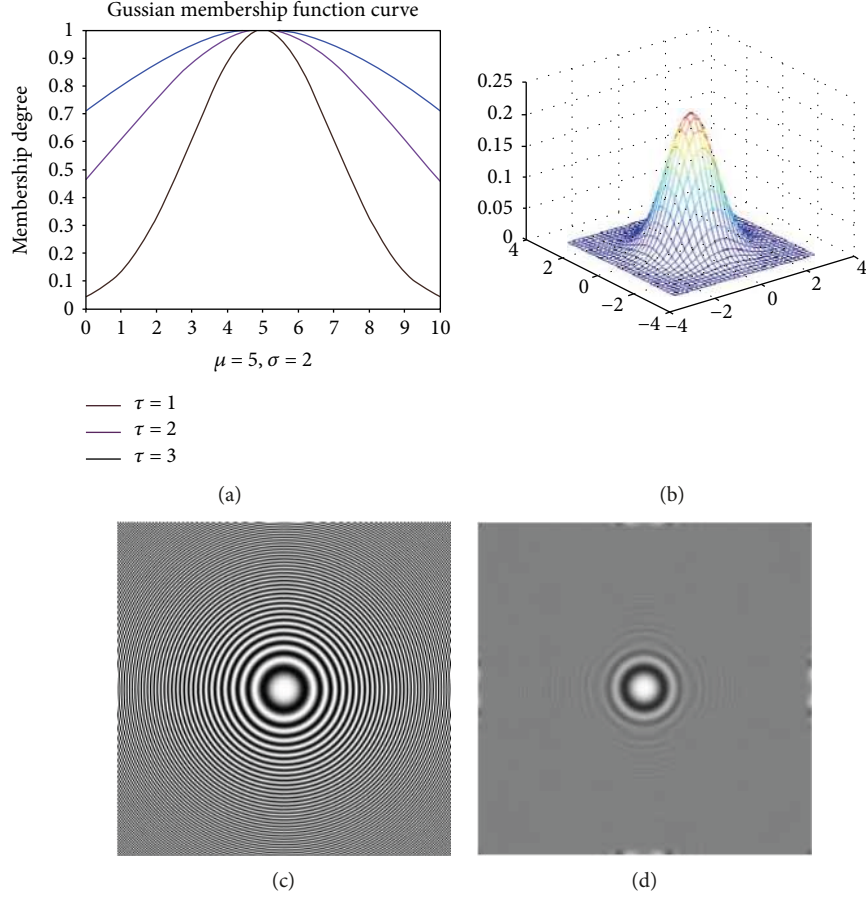


FIGURE 7: Gaussian function. (a) Gaussian curve with different  $\tau$ . (b) Gaussian surface. (c) Source image. (d) Produced image by Gaussian.

contains the significant target areas and then fuse the interesting regions of the IR and VI image. In our method, the MeanShift and guided filter are used to perform the first step in our algorithm.

The background region is obtained by removing the interesting region from the source image. For the background region, the multiscale transform-based method is used to process it. Firstly, nonsubsampling contourlet transform (NSCT) is used to decompose the background region of two source images and then to get a low-frequency layer and a series of high-frequency layers for each image. Next, we use an improved weighted average method based on per-pixel weighted average and pulse-coupled neural network (PCNN) to process the low-frequency and high-frequency layers, respectively.

**2.2.1. Low-Frequency Layer Fusion Rules.** In nature images, low-frequency information is the main component of an image; on the contrary, high-frequency information contains the details of the image [30]. It can be seen in Figure 6 that, compared with the image *B1* and *B2*, the low-frequency layers *L1* and *L2* are the main components without the details. Most low-frequency layer fusion methods are weighted averaging based methods, which do not consider the membership relationship between pixels and only weigh the independent pixel values. These methods cannot fully

fuse the details of the low-frequency layers. In order to have a better fusion effect, we proposed an improved weighted average method which is based on per-pixel weighted average, which can be described as follows:

$$C_F^L(i, j) = w_A(i, j) \times C_A^L(i, j) + (1 - w_A(i, j)) \times C_B^L(i, j), \quad (16)$$

where  $C_F^L(i, j)$  denotes the final result of the low-frequency layer,  $C_A^L(i, j)$  is the low-frequency layer of the background region in the source IR image A, and  $C_B^L(i, j)$  is the low-frequency layer of the background region in source VI image B.

$$w_A(i, j) = \exp \left[ -\frac{(C_B^L(i, j) - \mu)^2}{2(\tau\sigma)^2} \right], \quad (17)$$

where  $\mu$  and  $\sigma$  are the mean and variance of the background regions in source VI image B and  $\tau$  is the adjustment factor of Gaussian function. The Gaussian function curve and an example are shown in Figure 7. In the proposed method, we set  $\tau = 1$ . It can be seen in Figure 7(d) that, after the source image is processed by per-pixel weighted average, only the low-frequency information of the source image is reserved; to some extent, it is a low pass filter, and similar as Figure 4(b), the low-frequency layer is obtained by NSCT. Therefore, it is effective to process the

low-frequency layer by the weighted average method based on per-pixel weighted average.

**2.2.2. High-Frequency Layer Fusion Rules.** From Figure 4, it can be seen that most details of information, texture, and edge are included in the high-frequency layers. For the high-frequency layers, PCNN is used in the proposed method. In the modulation field, the linking coefficient  $\beta$  is a key parameter which value can directly affect the weighting of the linking channel. We use the spatial frequency (SF) of the high-frequency layer as the linking coefficient  $\beta$  in our proposed method. In Section 2.1.4, we have analyzed the PCNN model. The spatial frequency (SF) can reflect the overall definition level of an image; the SF of the source image is used to determine the linking strength  $\beta$ , which can be described as follows:

$$SF = \sqrt{RF^2 + CF^2}, \quad (18)$$

where RF is the spatial row frequency and CF is the spatial column frequency, which can be computed by

$$RF = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=2}^N [F(i, j) - F(i, j-1)]^2}, \quad (19)$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=2}^N [F(i, j) - F(i, j-1)]^2}.$$

The fused high-frequency layer  $C_{F,ij}$  can be determined as follows:

$$C_{F,ij} = \begin{cases} C_{A,ij}, & \mathbf{T}_{A,ij}(n) \geq \mathbf{T}_{B,ij}(n), \\ C_{B,ij}, & \mathbf{T}_{A,ij}(n) < \mathbf{T}_{B,ij}(n), \end{cases} \quad (20)$$

where  $\mathbf{T}_{A,ij}(n)$  and  $\mathbf{T}_{B,ij}(n)$  denote time matrices of each neuron obtained by (15) and  $C_{A,ij}$  and  $C_{B,ij}$  are the high-frequency layers of the background regions in source IR image A and VI image B.

**2.2.3. Fusion Steps.** The framework of the proposed method in this paper is shown in detail in Figure 6, and the concrete fusion steps are summarized as follows: input: source IR image A and VI image B.

- Step 1: detect the interesting region which contains the salient infrared objects of IR and corresponding VI image by the MeanShift, to get the interesting region and the background regions.
- Step 2: for the interesting regions of the source image, fuse them by the guided filter method which is described in Section 2.1.2, to produce the fused interesting region.
- Step 3: perform NSCT in the background regions and then obtain a low-frequency layer and a series of high-frequency layers for each source image.
- Step 4: for the low-frequency layer, an improved weighted average method based on per-pixel

weighted average algorithm is used to produce the fused low-frequency layer, which is shown in (16) and (17).

- Step 5: for the high-frequency layers, SF-PCNN-based method is used to produce the fused high-frequency layers, which are described in Section 2.2.2 in detail.
- Step 6: the fused background region is produced by NSCT reconstruction.
- Step 7: fuse the interesting region and the fused background region to produce the final fusion image.

### 3. Experimental Results and Analysis

In order to illustrate the effectiveness of the proposed fusion algorithm, several groups of IR and VI images fusion experiments will be described in detail in this section. These images are available at [http://figshare.com/articles/TNO\\_Image\\_Fusion\\_Dataset/1008029](http://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029). All simulations are conducted in MATLAB 2014a, on an Intel(R) Core (TM) i5-6400 @2.7GHz PC with 16GB RAM. Firstly, the experimental parameter setting is introduced; then the discussion of fusion results compared with other methods will be given.

**3.1. Experimental Introduction.** To show the improvement of the proposed method, the fusion results of “Jeep” by the proposed method and the method of [8] are shown in Figure 8. In [8], the target region was directly fused into the final image to highlight the target in an infrared image. And in this paper, we improve the accuracy of interesting region detection where it contains highlighting target and heat sources. In addition, we integrate the interesting regions of VI and IR for enriching the visible information in the interesting region. As shown in Figure 8 and Figure 9, the fusion result by the proposed method contains rich details while highlighting the target of IR image.

The proposed method will be compared with eight current fusion methods: principal component analysis- (PCA-) based method [10], discrete wavelet transform- (DWT-) based method [11], PCNN-based method [15], NSCT-based method [23], Laplacian pyramid transform- (LP-) PCNN-based method [14], NSCT-PCNN-based method [17], IFM based method [31], and MWGF-based method [32]. In all PCNN-based method experiments, through a large number of verification and comparison in experiments, the parameters of PCNN are set as  $\alpha_\theta = 0.2$ ,  $\alpha_L = 0.05$ ,  $V_L = 0.02$ ,  $V_\theta = 40$ ,  $N = 200$ , and  $M = W = [0.707, 1, 0.707; 1, 0, 1; 0.707, 1, 0.707]$ , where  $N$  is the number of iterations of PCNN. All NSCT-based methods, “pkva,” and “9-7” are set as the pyramid and the direction filter. For all multi-scale decomposition methods, the decomposition level is set to 3, and “averaging” is used to fuse the low-frequency layer, and the high-frequency layers are fused by “absolute maximum choosing.”

In order to evaluate the fusion results with different methods objectively, three most commonly used objective indicators will be used as the evaluation index: mutual

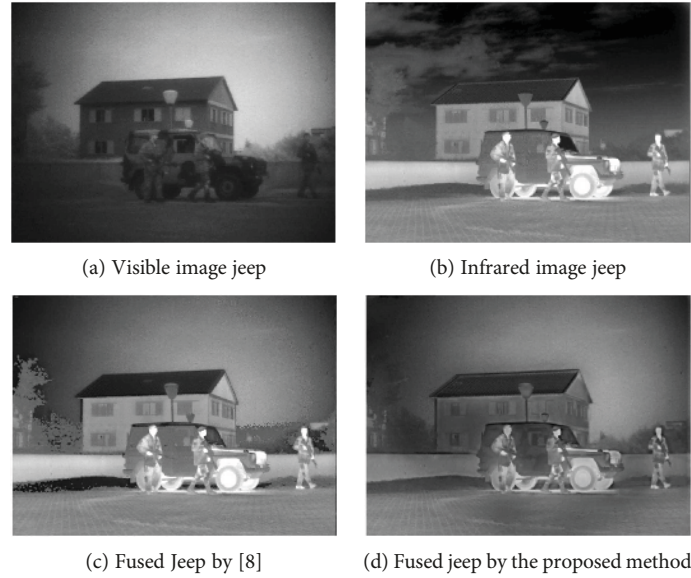


FIGURE 8: Experiment with “Jeep.”

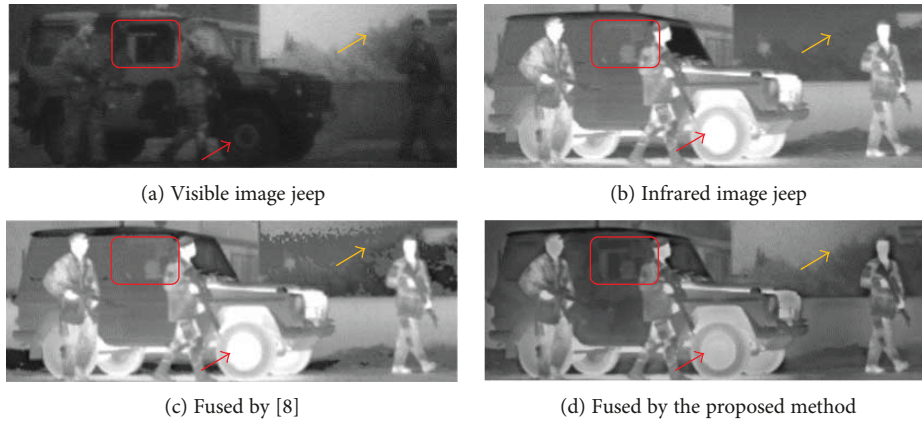


FIGURE 9: Detail with enlarged scale “Jeep.”

information (MI), pixel of visual information (VIF), and edge gradient operator ( $Q^{AB/F}$ ). MI is used to measure the amount of the source images' information retained in the fused image. VIF is an evaluation index for human visual system, which is based on natural scenes and image distortion [33].  $Q^{AB/F}$  is used to measure the edge information based on edge strength and orientation preserving from the source images. Commonly, the greater value of these evaluation metrics indicates that the fused image has a better quality [34].

**3.2. Fusion Results and Discussions.** The experimental images consisted of six pairs of IR and VI images, which are shown in Figure 10. The first line in Figure 10 is VI images, and the second line is IR images. A large number of details and texture are included in the VI images, while the IR image contains only significant information.

The fusion results obtained by different fusion algorithms of “Sand path” are given in Figure 11. Figure 11(a) is the fused image by PCA, Figure 11(b) is the fused image

by DWT, Figure 11(c) is the fused image by PCNN, Figure 11(d) is the fused image by NSCT, Figure 11(e) is the fused image by LP-NSCT, Figure 11(f) is the fused image by PCNN-NSCT, Figure 11(g) is the fused image by IFM, Figure 11(h) is the fused image by MWGF, and Figure 11(i) is the fused image by the proposed method. From Figure 11, we can see that the fused image by the proposed method contains more detail information of the VI image, as well as the highlighted infrared target information compared with other methods. In addition, the fused image by our method has advantages in visual effects and it is also superior to other algorithms in objective evaluation, which are shown in Table 1.

In order to illustrate the applicability of the proposed method, other groups of experiments are performed, which are given as follows.

It can be seen from Figure 12 and Figure 13 that the proposed method has more advantages in detail information integration. In order to better reflect their differences of the fused images obtained by different fusion methods,



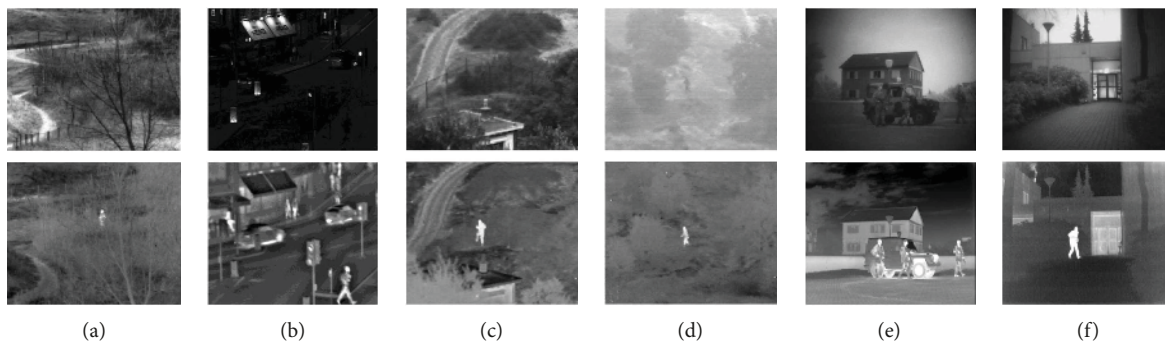


FIGURE 10: Experimental images. (a) “Sand path.” (b) “Bristol Queens Road.” (c) “UN Camp.” (d) “Trees.” (e) “Jeep.” (f) “Kaptein.”

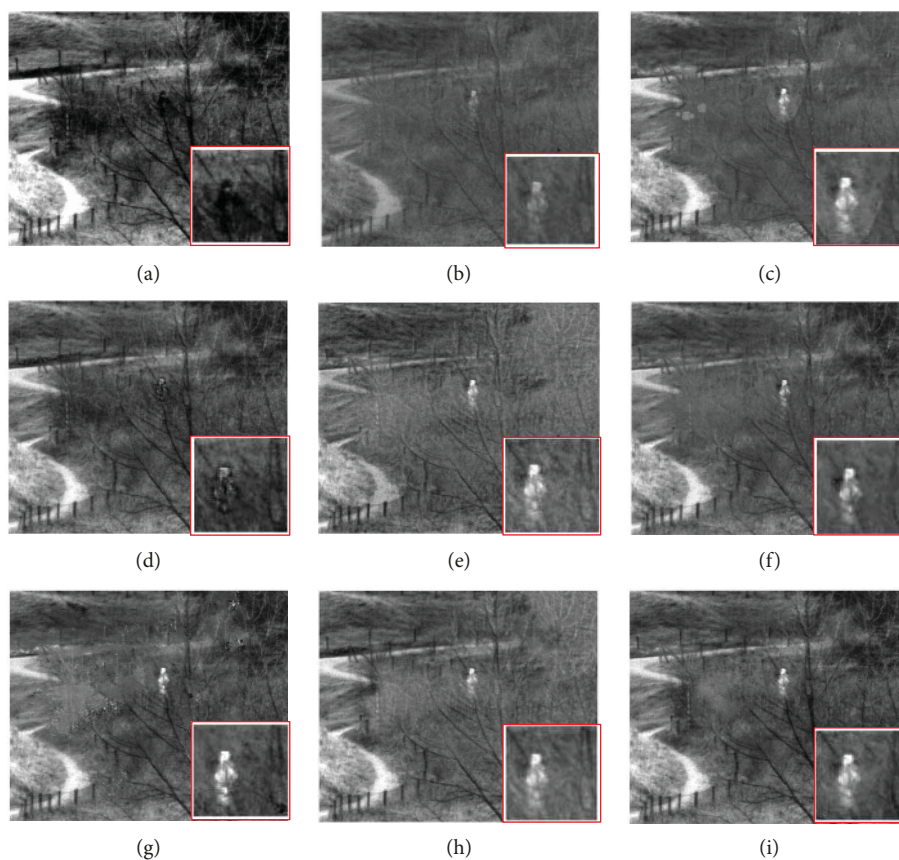


FIGURE 11: Experimental results of “Sand path”. (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.

Figure 13 shows the detail with an enlarged scale of Figure 12. In Figure 13(i), the red frame region is more suitable for human visual system, with more visible detail information. Compared with the same position in two source images, the fused region has higher readability and reliability.

Figure 14 and Figure 15 are the third and fourth groups of experiment of “UN Camp” and “Trees,” respectively. The objective evaluation matrices are given in Table 1. In order to reflect more directly out of their difference, line chart comparison of MI, VIF, and  $Q^{A/B/F}$  values of the experiments is given in Figure 18.

Figure 16 and Figure 17 are the experimental results of “Jeep” and “Kaptein.” The objective evaluation matrices are given in Table 1. Line chart comparison of MI, VIF, and  $Q^{A/B/F}$  values of the experiments is given in Figure 19. It can be seen that the fusion results of the proposed method have a better visual effect. Compared with the same position in two source images, the fused region has higher readability and reliability.

All fusion results by the proposed method and the method of [8] are shown in Figure 18. The first line in Figure 18 is the results of [8], and the second line is the results of the proposed method. The objective evaluation matrices

TABLE 1: Objective results for various fusion results.

Evaluation	Experiment	Fusion methods									
		PCA	DWT	PCNN	NSCT	LP-PCNN	NSCT-PCNN	IFM	MWGF	Reference [8]	Proposed
MI	Sand path	1.1117	0.5578	1.4473	1.5081	0.5468	1.0629	1.6325	1.7830	1.7935	1.8874
	Bristol Queens Road	2.4788	1.2387	2.3390	1.9813	1.1924	2.2035	1.6665	2.5665	2.3513	2.6238
	UN Camp	2.1191	0.7973	2.0342	1.5648	0.8595	1.0644	2.7810	2.9160	2.3461	2.1219
	Trees	1.3114	0.8653	0.8844	1.1424	0.7757	1.3761	1.0500	1.3480	1.3418	1.3598
	Jeep	2.1219	1.0000	3.1090	2.4130	1.9150	2.4980	1.8704	2.9550	2.9721	2.9613
	Kaptein	2.0190	1.0910	2.5090	1.8650	1.1960	1.7100	3.469	3.3830	3.2045	3.3921
	Average MI	1.8603	0.9250	2.0538	1.7457	1.0809	1.6524	2.0782	2.4919	2.3349	2.3910
VIF	Sand path	0.4201	0.2566	0.3048	0.3284	0.2788	0.3886	0.3599	0.4813	0.4821	0.5230
	Bristol Queens Road	0.4991	0.3797	0.4986	0.4906	0.4687	0.5196	0.4947	0.5715	0.4949	0.5202
	UN Camp	0.4599	0.2899	0.3107	0.3538	0.3156	0.3743	0.4419	0.4937	0.3873	0.4656
	Trees	0.3776	0.3238	0.3145	0.4128	0.3648	0.4372	0.5141	0.5013	0.5243	0.4311
	Jeep	0.4575	0.4100	0.5440	0.4841	0.5544	0.4788	0.4572	0.6023	0.5528	0.6107
	Kaptein	0.4182	0.3637	0.4332	0.4488	0.4308	0.4359	0.4871	0.5189	0.4947	0.5103
	Average VIF	0.4387	0.3372	0.4009	0.4197	0.4021	0.4390	0.4591	0.5281	0.4894	0.5101
$Q^{AB/F}$	Sand path	0.5013	0.3934	0.4188	0.4331	0.4136	0.5218	0.5104	0.5161	0.5072	0.5295
	Bristol Queens Road	0.5678	0.3841	0.6708	0.5963	0.5889	0.6661	0.6066	0.6746	0.6693	0.6719
	UN Camp	0.5038	0.3442	0.4785	0.4169	0.4356	0.4650	0.5265	0.5262	0.5296	0.5330
	Trees	0.4533	0.3909	0.3804	0.4306	0.4476	0.4423	0.6004	0.5815	0.4832	0.4524
	Jeep	0.4252	0.3436	0.6453	0.3022	0.5467	0.3142	0.3892	0.6769	0.6348	0.6818
	Kaptein	0.3249	0.3442	0.5280	0.3692	0.4759	0.4240	0.5853	0.6109	0.6118	0.6189
	Average $Q^{AB/F}$	0.4627	0.3667	0.5203	0.4247	0.4847	0.4722	0.5364	0.5977	0.5726	0.5813

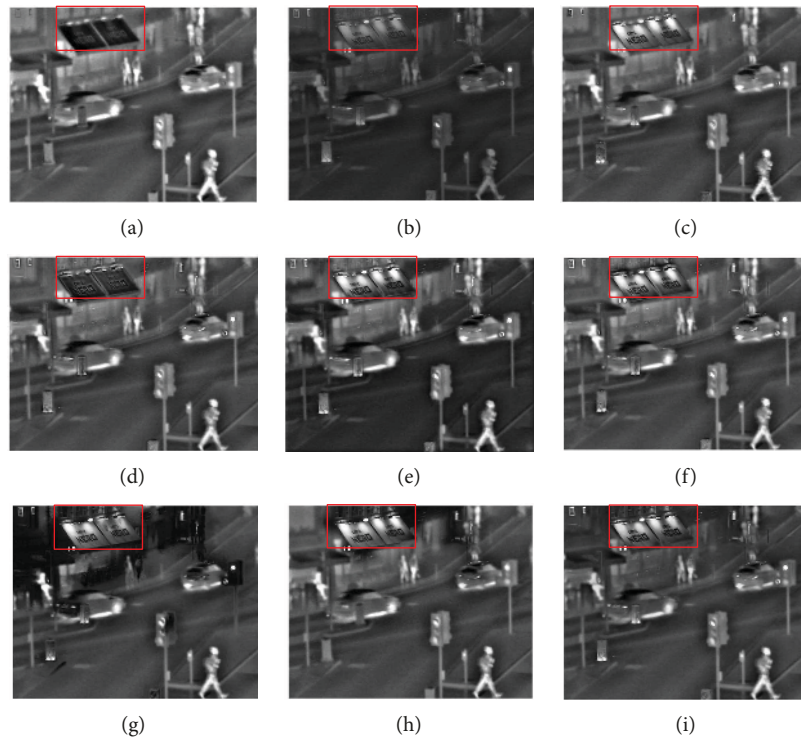


FIGURE 12: Second example experimental results of “Bristol Queens Road.” (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.

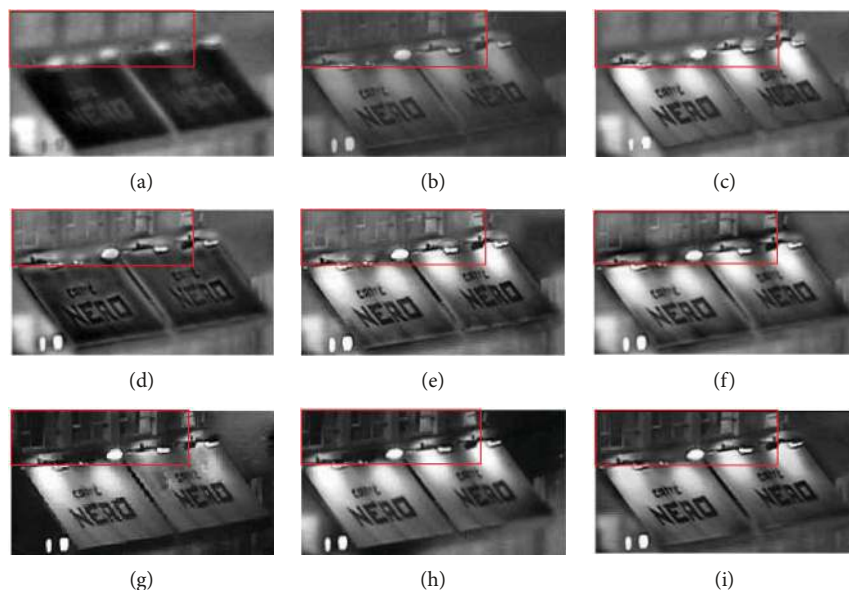


FIGURE 13: Detail with enlarged scale. (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.

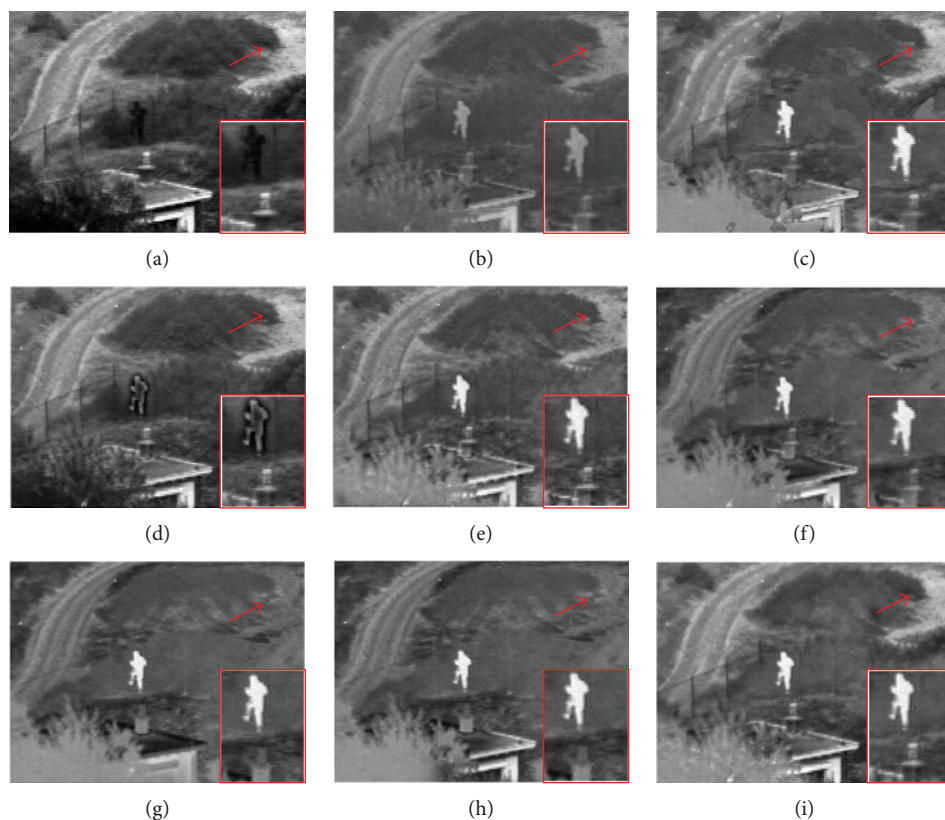


FIGURE 14: Experimental results of “UN Camp.” (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.

are given in Table 1. It can be seen from Figure 18 and Table 1 that the proposed method has better results in objective evaluation and visual effect. The average values of MI, VIF, and  $Q^{AB/F}$  for six pairs of IR and VI images are listed

in Table 1. It can be seen that the average value of the proposed method is only less than MWGF and greater than the other comparison algorithms, because the MWGF is more effective for some special images, so as to increase the



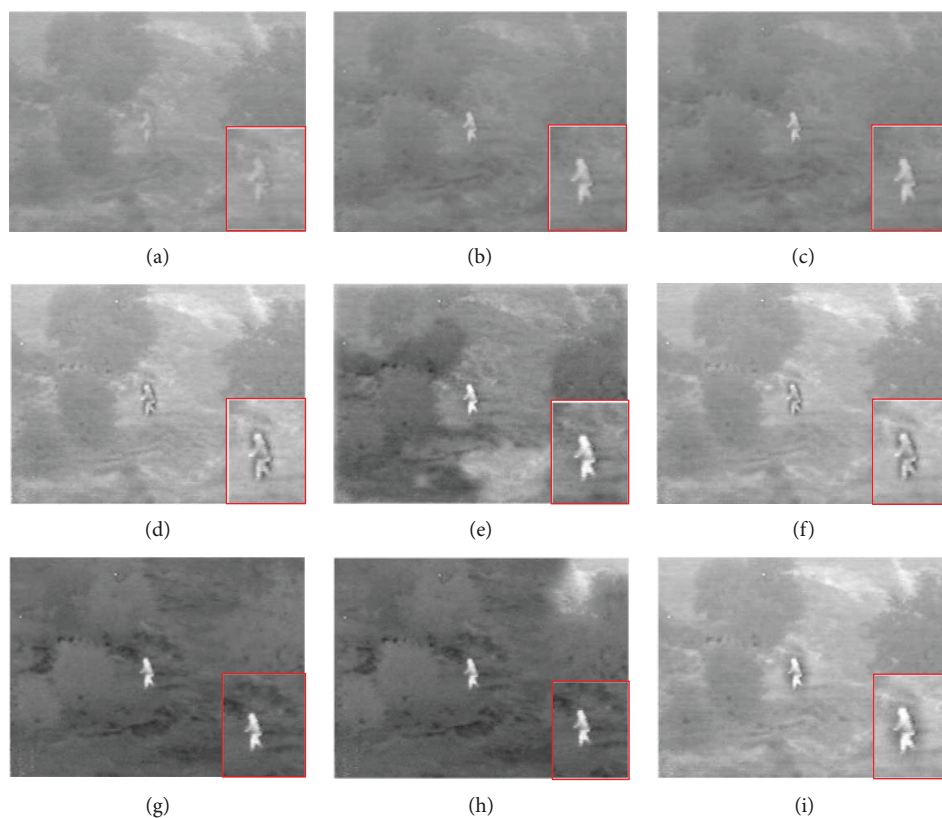


FIGURE 15: Experimental results of “Trees.” (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.

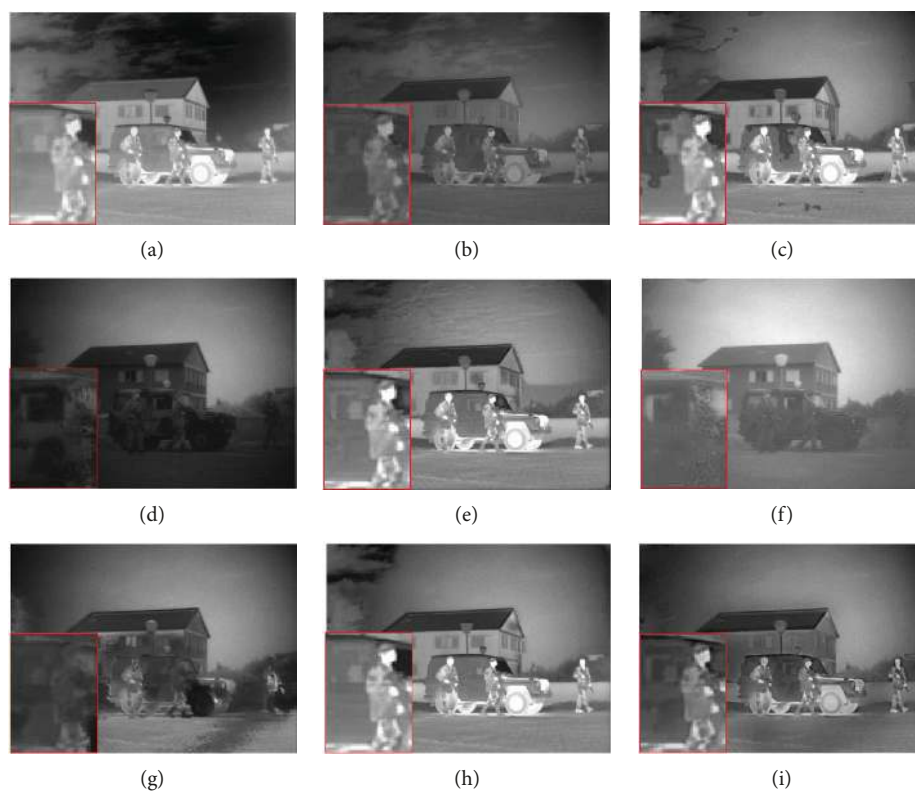


FIGURE 16: Experimental results of “Jeep.” (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.



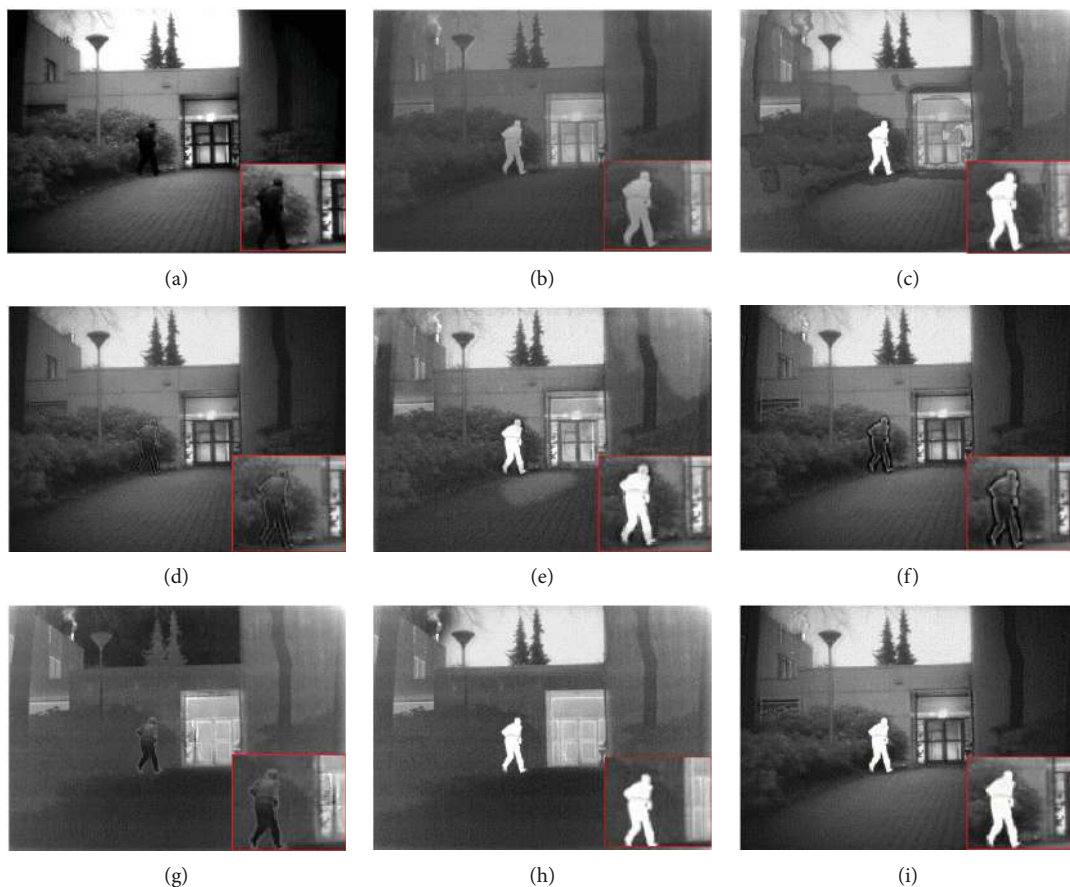


FIGURE 17: Experimental results of “Kaptein.” (a) PCA. (b) DWT. (c) PCNN. (d) NSCT. (e) LP-PCNN. (f) NSCT-PCNN. (g) IFM. (h) MWGF. (i) The proposed method.

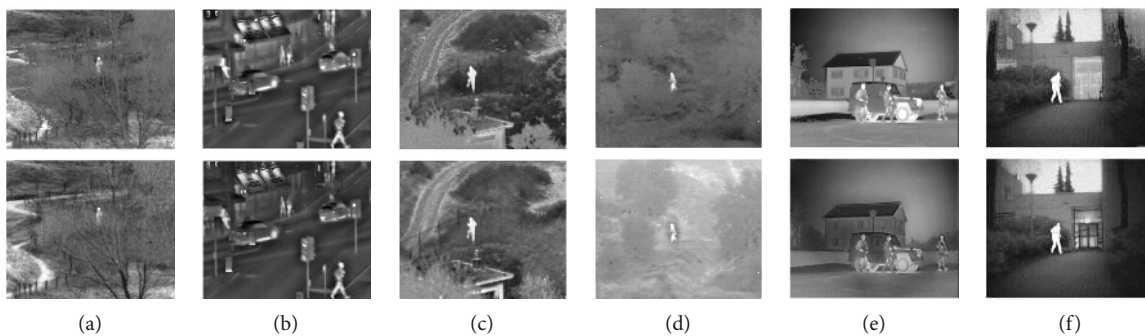


FIGURE 18: Experimental results by [8] and the proposed method. (a) “Sand path.” (b) “Bristol Queens Road.” (c) “UN Camp.” (d) “Trees.” (e) “Jeep.” (f) “Kaptein.”

average value, such as the MI of “UN Camp”, but the proposed method has higher evaluation metrics on most images than MWGF.

#### 4. Conclusion

A novel multiscale fusion scheme for the IR and VI image based on the interesting region detection is proposed in this paper, which can integrate more background details as well as highlight the interesting region with the salient objects. This method combines the advantages of the MeanShift

and guided filter, which are used to detect the interesting significant target region and fuse the interesting regions of the IR and VI image. Next, the background regions are fused in the NSCT domain. An improved weighted average method based on per-pixel weighted average is used to fuse the low-frequency layers, and for the high-frequency layers, SF-PCNN-based method is used to produce the fused new layers. Then the fused background regions are produced by NSCT reconstruction. The fused image is produced by fusing the fused interesting and background regions. Experimental results show that the

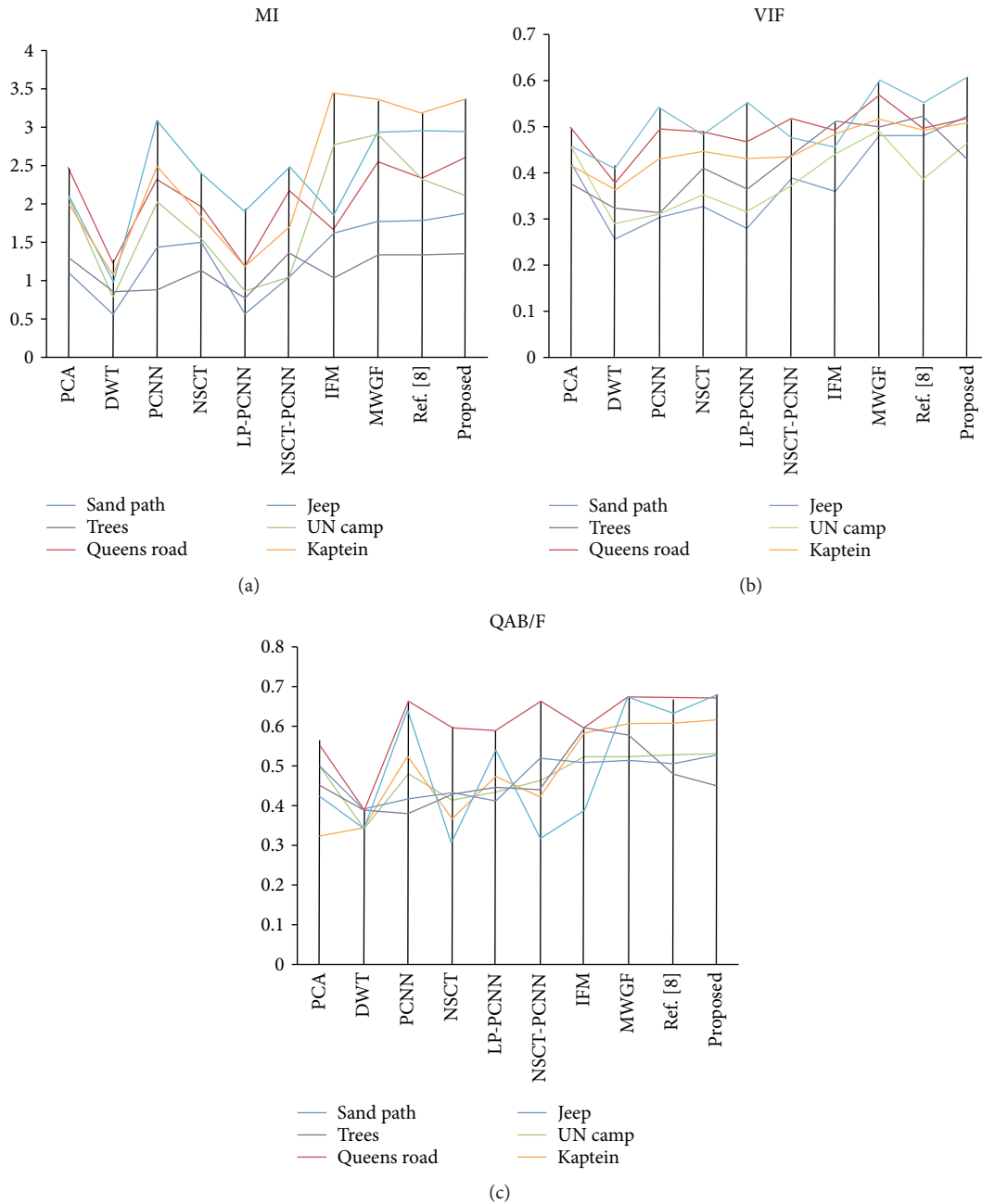


FIGURE 19: Line chart comparison of MI, VIF, and  $Q^{AB/F}$  values for four groups of the experiments.

proposed fusion scheme can achieve superior results in visual inspection and objective evaluations.

### Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (Grant nos. 61365001 and 61463052).

### References

- [1] H. Jin, Q. Xi, Y. Wang, and X. Hei, "Fusion of visible and infrared images using multiobjective evolutionary algorithm based on decomposition," *Infrared Physics & Technology*, vol. 71, pp. 151–158, 2015.
- [2] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8–17, 2017.
- [3] C. H. Liu, Y. Qi, and W. R. Ding, "Infrared and visible image fusion method based on saliency detection in sparse domain," *Infrared Physics & Technology*, vol. 83, pp. 94–102, 2017.

- [4] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Information Sciences*, vol. 432, pp. 516–529, 2018.
- [5] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191–207, 2017.
- [6] H. Li, X. Li, Z. Yu, and C. Mao, "Multifocus image fusion by combining with mixed-order structure tensors and multiscale neighborhood," *Information Sciences*, vol. 349–350, pp. 25–49, 2016.
- [7] Y. Ma, J. Chen, C. Chen, F. Fan, and J. Ma, "Infrared and visible image fusion using total variation model," *Neurocomputing*, vol. 202, pp. 12–19, 2016.
- [8] K. He, D. Zhou, X. Zhang, R. Nie, Q. Wang, and X. Jin, "Infrared and visible image fusion based on target extraction in the nonsubsampling contourlet transform domain," *Journal of Applied Remote Sensing*, vol. 11, no. 1, pp. 1–14, 2017.
- [9] H. Zhang, Q. Chen, D. Yuan, Y. H. You, and M. Sun, "Fusion of infrared and visible images using 2DPCA bases," in *2013 2nd IAPR Asian Conference on Pattern Recognition*, pp. 596–600, Naha, Japan, November 2013.
- [10] Y. Hu and Y. Yang, "Study on feature fusion for target recognition based on PCA with infrared and visible light images," *International Journal of Digital Content Technology and its Applications*, vol. 7, no. 3, pp. 436–444, 2013.
- [11] Y. Niu, S. Xu, L. Wu, and W. Hu, "Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform," *Mathematical Problems in Engineering*, vol. 2012, Article ID 275138, 10 pages, 2012.
- [12] Z. Fu, X. Wang, J. Xu, N. Zhou, and Y. Zhao, "Infrared and visible images fusion based on RPCA and NSCT," *Infrared Physics & Technology*, vol. 77, pp. 114–123, 2016.
- [13] H. Li, H. Qiu, Z. Yu, and Y. Zhang, "Infrared and visible image fusion scheme based on NSCT and low-level visual features," *Infrared Physics & Technology*, vol. 76, pp. 174–184, 2016.
- [14] M. Unser, "An improved least squares Laplacian pyramid for image compression," *Signal Processing*, vol. 27, no. 2, pp. 187–203, 1992.
- [15] T. Xiang, L. Yan, and R. Gao, "A fusion algorithm for infrared and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain," *Infrared Physics & Technology*, vol. 69, pp. 53–61, 2015.
- [16] A. L. Da Cunha, J. Zhou, and M. N. Do, "The nonsubsampling contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.
- [17] S. Jianhui, G. Jing, and L. Yanju, "Fusion of infrared and visible images based on pulse coupled neural network and nonsubsampling contourlet transform," *The Open Cybernetics & Systems Journal*, vol. 9, no. 1, pp. 17–22, 2015.
- [18] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [19] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [20] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2864–2875, 2013.
- [21] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 120–129, 2015.
- [22] H. Li, Y. Chai, and Z. Li, "Multi-focus image fusion based on nonsubsampling contourlet transform and focused regions detection," *Optik-International Journal for Light and Electron Optics*, vol. 124, no. 1, pp. 40–51, 2013.
- [23] C. Zhao, Y. Guo, and Y. Wang, "A fast fusion scheme for infrared and visible light images in NSCT domain," *Infrared Physics & Technology*, vol. 72, pp. 266–275, 2015.
- [24] Q. Zhang and B. L. Guo, "Multifocus image fusion using the nonsubsampling contourlet transform," *Signal Processing*, vol. 89, no. 7, pp. 1334–1346, 2009.
- [25] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, "Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex," *Neural Computation*, vol. 2, no. 3, pp. 293–307, 1990.
- [26] X. Jin, R. Nie, D. Zhou et al., "A novel DNA sequence similarity calculation based on simplified pulse-coupled neural network and Huffman coding," *Physica A: Statistical Mechanics and its Applications*, vol. 461, pp. 325–338, 2016.
- [27] X. Jin, Q. Jiang, S. Yao et al., "A survey of infrared and visual image fusion methods," *Infrared Physics & Technology*, vol. 85, pp. 478–501, 2017.
- [28] K. Fukunaga and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 32–40, 1975.
- [29] S. Hao, D. Pan, Y. Guo, R. Hong, and M. Wang, "Image detail enhancement with spatially guided filters," *Signal Processing*, vol. 120, pp. 789–796, 2016.
- [30] A. van der Schaaf and J. H. van Hateren, "Modelling the power spectra of natural images: statistics and information," *Vision Research*, vol. 36, no. 17, pp. 2759–2770, 1996.
- [31] S. Li, X. Kang, J. Hu, and B. Yang, "Image matting for fusion of multi-focus images in dynamic scenes," *Information Fusion*, vol. 14, no. 2, pp. 147–162, 2013.
- [32] Z. Zhou, S. Li, and B. Wang, "Multi-scale weighted gradient-based fusion for multi-focus images," *Information Fusion*, vol. 20, pp. 60–72, 2014.
- [33] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.
- [34] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.





**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

