

## Research Article

# Image Fusion with Contrast Improving and Feature Preserving

Din-Chang Tseng,<sup>1</sup> Yu-Shuo Liu,<sup>1</sup> and Chang-Min Chou<sup>2</sup>

<sup>1</sup> Department of Computer Science and Information Engineering, National Central University, No. 300 Jhongda Road, Jhongli 32001, Taiwan

<sup>2</sup> Department of Electronic Engineering, Chien Hsin University of Science and Technology, No. 229 Jianxing Road, Jhongli 32097, Taiwan

Correspondence should be addressed to Chang-Min Chou; [changmin@uch.edu.tw](mailto:changmin@uch.edu.tw)

Received 11 June 2014; Accepted 28 July 2014

Academic Editor: Teen-Hang Meen

Copyright © 2015 Din-Chang Tseng 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 goal of image fusion is to obtain a fused image that contains most significant information in all input images which were captured by different sensors from the same scene. In particular, the fusion process should improve the contrast and keep the integrity of significant features from input images. In this paper, we propose a region-based image fusion method to fuse spatially registered visible and infrared images while improving the contrast and preserving the significant features of input images. At first, the proposed method decomposes input images into base layers and detail layers using a bilateral filter. Then the base layers of the input images are segmented into regions. Third, a region-based decision map is proposed to represent the importance of every region. The decision map is obtained by calculating the weights of regions according to the gray-level difference between each region and its neighboring regions in the base layers. At last, the detail layers and the base layers are separately fused by different fusion rules based on the same decision map to generate a final fused image. Experimental results qualitatively and quantitatively demonstrate that the proposed method can improve the contrast of fused images and preserve more features of input images than several previous image fusion methods.

## 1. Introduction

Image fusion is a technology that has been widely used in many areas such as defense surveillance, remote sensing, medical imaging, and structure assessment. The goal of image fusion is to obtain a fused image which contains most significant information in all input images captured by different sensors from the same scene [1]. A typical fusion process is to fuse spatially registered visible and infrared images. Compared with infrared image, the visible image usually captures abundant object details and should be considered as the main information resource. However, the interested objects are not easy to be observed in visible images when they are under dark circumstance or with the same color of the background. In such cases, an infrared image can be used for providing extra information which cannot be easily observed in the visible images. The fused image should provide more information for human or machine perception as compared to a single input image; in other words,

a good fused image should interpret the real scene with more reliability.

The image fusion methods can be categorized into pixel-level, feature-level, and symbolic-level fusion methods [1, 2]. In pixel-level fusion methods, a fused image is generated by combining individual pixel or small regular regions of pixels from multiple input images based on fusion decision algorithms [3–5]. In feature-level fusion methods, multiple input images are initially segmented into regions which are then fused according to various properties of all related regions [6–8]. In symbolic-level fusion methods, abstractions are extracted from all input images followed by combining these abstractions to a fused image [2]. Feature-level fusion methods are also named region-based methods. Compared with pixel-level fusion methods, region-based fusion methods have a number of perceived advantages including reducing sensitivity to noise, decreasing artifacts or inconsistencies in the fused images, preserving significant features of the input images, and increasing flexibility in choosing intelligent

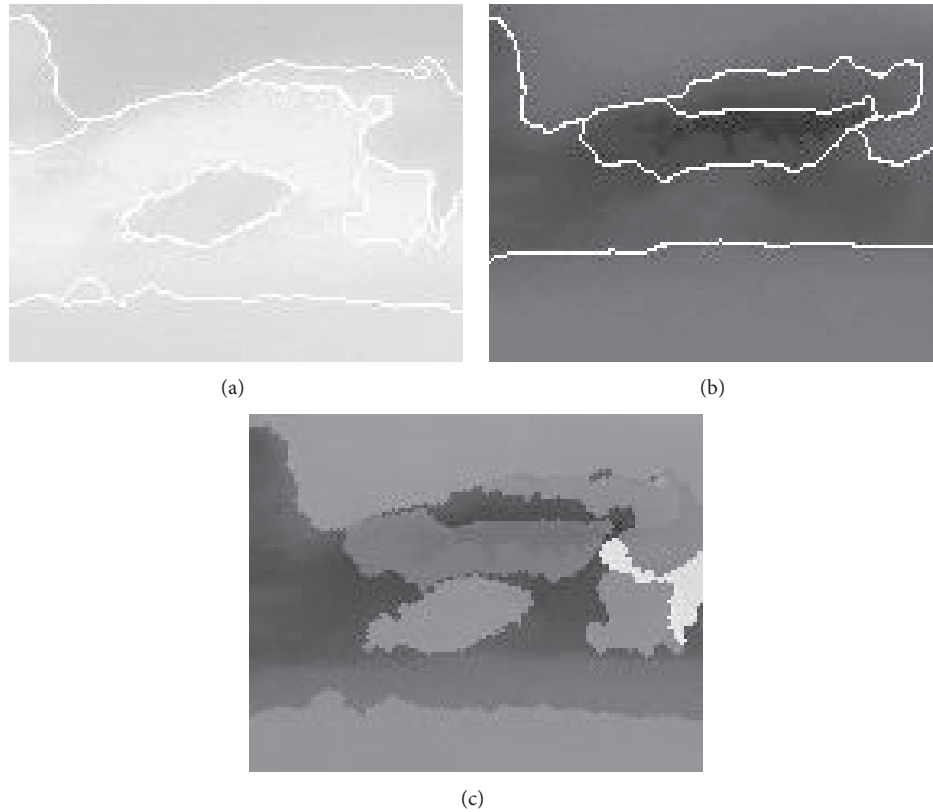


FIGURE 1: The inconsistency problem of maximum entropy method. (a) Visible image. (b) Infrared image. (c) Inconsistent result.

fusion rules [9]; thus we focus on region-based fusion methods in this study.

Multiscale analysis plays a basic role for image fusion using different fusion rules to integrate an image with different scales. Multiscale analysis can be classified into multiscale transform and multiscale geometrical analysis methods [10]. Multiscale transform methods such as Laplacian pyramid transform, discrete wavelet transform (DWT) [3], and dual-tree complex wavelet transform (DT-DWT) [9] can effectively extract important information of images, such as edges and details; thus multiscale transform methods have been widely used for image fusion. However, multiscale transform methods are usually accompanied with complex computations, which will make these methods become inefficient. Multiscale geometrical analysis methods such as curvelet filter [11], contourlet filter [12], weighted least squares [13], guided filter [14], and bilateral filter [15, 16] have been successfully applied on several image fusion methods for their ability of capturing intrinsic geometrical structure of images [10, 14, 17]. Moreover, the bilateral filter is a spatial-domain filter which can preserve significant edges while smoothing images. In this study, we use the bilateral filter to decompose input images into base and detail layers for fusion.

Many region-based fusion methods have been proposed by using multiscale analysis to decompose an input image into one or several detail layers and a base layer in intensity [9, 18, 19]. Then the detail layers and base layers of input images are separately processed to construct a fused detail

layer and a base layer. At last, the fused base and detail layers are composed to obtain a fused image. In conventional methods, the construction of the fused detail layer is achieved by selecting maximum coefficients or calculating weighted averages in a region-based manner by comparing regional properties such as normalized Shannon entropy [9] or alpha-stable modeling [19]. The construction of the fused base layer is usually achieved by averaging the gray levels in a pixel-based manner. However, fusing base layers by averaging the gray levels may weaken significant contrast of the fused image. Saeedi and Faez [20] proposed a region-based method using particle swarm optimization and genetic algorithm to find the optimal weights for fusing the base layer with maximum entropy. Most region-based fusion methods suffer from oversegmentation problem. Oversegmentation may result in inconsistency problems. The inconsistency problem usually occurs on the oversegmented regions with low gray levels in a visible image and high gray levels in an infrared image. As one example shown in Figure 1, the fused image (sky) in Figure 1(c) is composed of regions of highly different gray levels and becomes unnatural. In this study, we proposed a method to alleviate the inconsistency problem. The goal of the proposed method is to improve the contrast of the fused image and preserve the significant features of input images.

The remainder of this paper is organized as follows. In Section 2, we review the principle of bilateral filters. Section 3 explains the proposed region-based fusion method in detail. The experiments and their results are presented and discussed

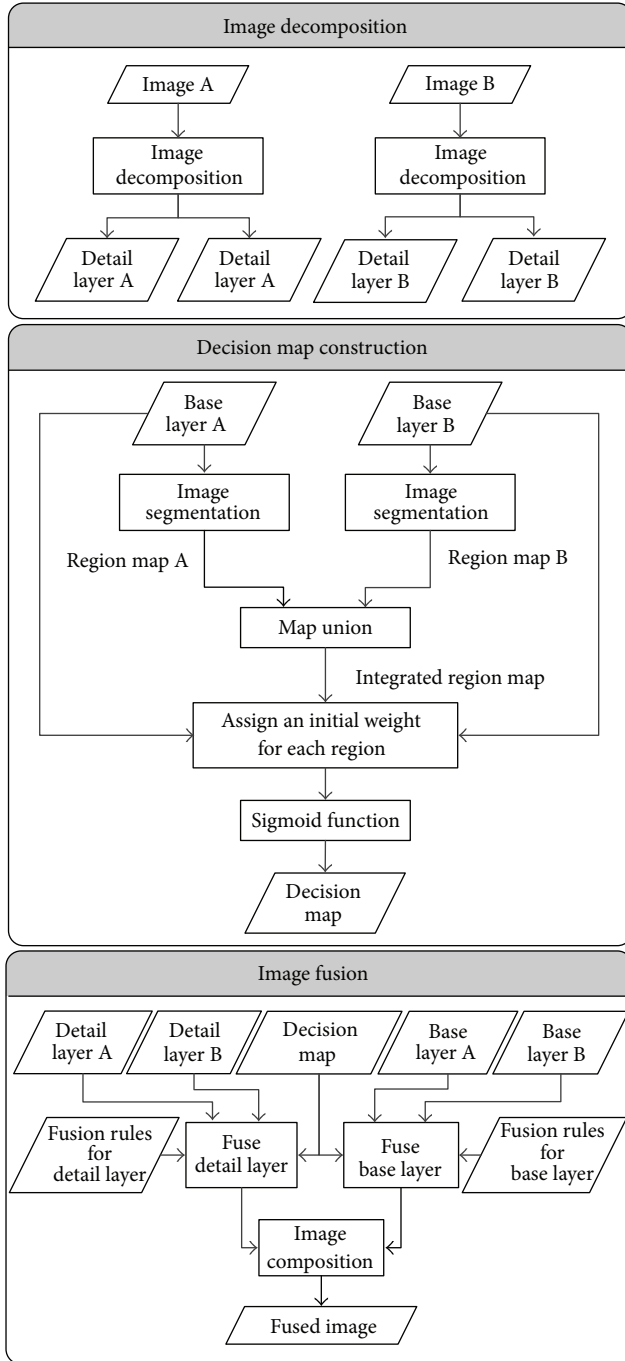


FIGURE 2: Procedure diagram of the proposed image fusion method consists of image decomposition, decision map construction, and image fusion stages.

in Section 4. Finally, concluding remarks and future works are given in Section 5.

## 2. Bilateral Filters

A bilateral filter is a nonlinear filter; it extends the concept of Gaussian smoothing filter by weighting the mask coefficients according to their corresponding pixel relative intensities. A

bilateral filter consists of a spatial filter in spatial domain and a range filter in intensity domain. When applying on image decomposition, the bilateral filter is an edge-preserving smoothing technique which can effectively maintain the sharpness of edges while smoothing images [16]. The entries of a bilateral-filter mask completely follow the Gaussian functions of these two filters. The bilateral filter is defined by

$$\text{bf}(p_{ij}) = \frac{1}{\sum_{(s,t) \in \Omega} G(\|(s,t) - (i,j)\|) F_r(|I_{st} - I_{ij}|)} \times \sum_{(s,t) \in \Omega} G(\|(s,t) - (i,j)\|) F_r(|I_{st} - I_{ij}|) I_{st}, \quad (1)$$

where  $(i, j)$  is the coordinates of pixel  $p_{ij}$ ,  $I_{ij}$  is the gray level of  $p_{ij}$ ,  $\|(s, t) - (i, j)\|$  is the Euclidean distance between pixels  $p_{st}$  and  $p_{ij}$ ,  $\Omega$  is a window centered at  $(i, j)$ ,  $G$  is a Gaussian spatial kernel for smoothing differences in location and defined as  $G(\|\mathbf{x}\|) = e^{-(1/2)(\|\mathbf{x}\|/\sigma_s)^2}$ ,  $\sigma_s$  is the standard deviation of  $G$ ,  $F_r$  is a Gaussian range kernel for smoothing differences in intensities and defined as  $F_r(|x|) = e^{-(1/2)(|x|/\sigma_r)^2}$ , and  $\sigma_r$  is the standard deviation of  $F_r$ .

## 3. The Proposed Fusion Method

The procedure diagram of the proposed method is shown in Figure 2. The proposed method consists of three essential stages: image decomposition, decision map construction, and image fusion.

- (1) In the image decomposition stage, input images are decomposed into detail layers and base layers using a bilateral filter.
- (2) In the decision map construction stage, the base layers of visible image and infrared image are segmented into regions to produce individual region maps. These two region maps are unionized into a single region map, named as integrated region map. Based on the integrated region map, the initial weights of regions in the input images are calculated as difference of neighboring regions (DNRs) in the base layers. At last, the contrast of DNRs is enhanced by using a sigmoid function to produce a decision map for recording the final weights of all regions.
- (3) In the image fusion stage, a set of fusion rules are first defined; then according to the rules, the base layers and the detail layers of all input images are fused based on the same decision map. At last, the fused base and detail layers are composed to obtain a fused image.

**3.1. Image Decomposition.** The bilateral filter is used to decompose an image into a base layer and a detail layer [21]. At first, the base layer  $I_{\text{base}}$  is obtained by applying the bilateral filter to the input image  $I$ :

$$I_{\text{base}} = \text{bf}(I), \quad (2)$$

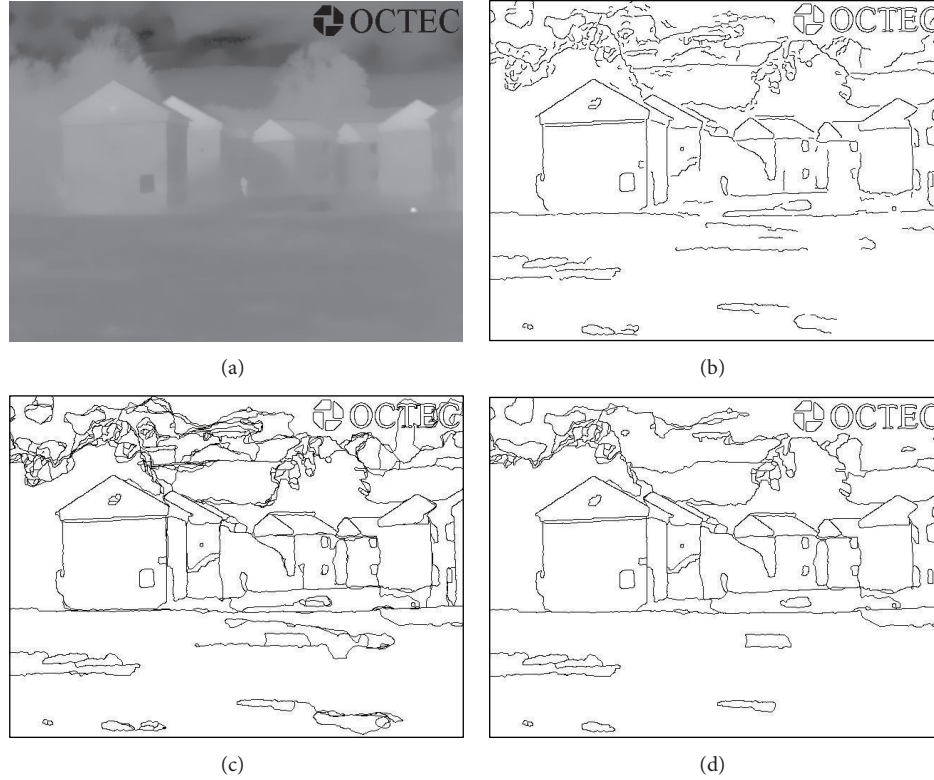


FIGURE 3: Image segmentation. (a) The base layer of input image. (b) The result of canny edge detection. (c) After edge linking. (d) The result of segmentation after region merging.

where  $\text{bf}(\cdot)$  represents the bilateral function as described in (1). The detail layer  $I_{\text{detail}}$  is obtained by substituting the base layer from the original image:

$$I_{\text{detail}} = I - I_{\text{base}}. \quad (3)$$

**3.2. Image Segmentation.** The proposed segmentation method classifies the neighboring pixels with similar gray levels into a region. There are three main stages in the image segmentation process: edge detection, edge linking, and region merge. An example is illustrated in Figure 3. In the edge detection stage, we use the canny edge detector [22] to extract edges as shown in Figure 3(b), where the edges with length less than 10 pixels are eliminated. In general, the edges extracted by an edge detector are not completely connected; the edges should be linked to form regions. The edge linking stage is achieved by extending the existing edges along their end directions to the other existed edges, as one example shown in Figure 3(c).

An edge pixel extracted in the previous step is assigned to the region of its neighboring pixel with the closest gray level. Thus all pixels of an image are classified into regions. The edge linking may generate oversegmented results; thus we need to merge neighboring regions with similar gray levels. In the region merge stage, two adjacent regions will be merged if the difference of their gray-level averages is smaller than a predefined threshold  $T_m$ , as one example shown in Figure 3(d).

**3.3. Integrated Region Map.** For an image  $I$  segmented into a set of regions, its region map is defined as  $R = \{r_1, r_2, \dots, r_n\}$ , where  $r_i$  represents the  $i$ th region of  $I$ . Let  $I_{\text{base}}^{\text{visible}}$  and  $I_{\text{base}}^{\text{infrared}}$  denote the base layers of input visible and infrared images, respectively. Each pixel in  $I_{\text{base}}^{\text{visible}}$  and  $I_{\text{base}}^{\text{infrared}}$  is assigned to a region according to their region maps  $R^{\text{visible}}$  and  $R^{\text{infrared}}$  respectively, where  $R^{\text{visible}} = \{r_1^{\text{visible}}, r_2^{\text{visible}}, \dots, r_p^{\text{visible}}\}$  and  $R^{\text{infrared}} = \{r_1^{\text{infrared}}, r_2^{\text{infrared}}, \dots, r_q^{\text{infrared}}\}$ . These two region maps are unionized to form an integrated region map  $R = \{r_1, r_2, \dots, r_n\}$ , where  $r_k$  denotes the  $k$ th region. Pixels belonging to the same regions in both region maps will be assigned to the same region in the integrated region map. The idea of the integrated region map is illustrated in Figure 4. In Figure 4, pixels  $p_1$ ,  $p_2$ , and  $p_3$  belong to the same region in the visible image,  $p_1$  and  $p_2$  belong to the same region, but  $p_3$  belongs to the other region in the infrared image. Thus  $p_1$  and  $p_2$  are assigned to the same region and  $p_3$  is assigned to another region in the integrated region map.

**3.4. Decision Map.** A decision map is a map of weights to indicate the importance of all regions in the input images. Let  $b(i, j)$  denote the gray level of boundary pixel  $(i, j)$  in  $I_{\text{base}}$ , where a boundary pixel is defined as a pixel whose eight neighboring pixels belong to more than one regions. The average gray-level difference  $d(i, j)$  of a boundary pixel  $(i, j)$  is calculated by averaging the difference between the boundary

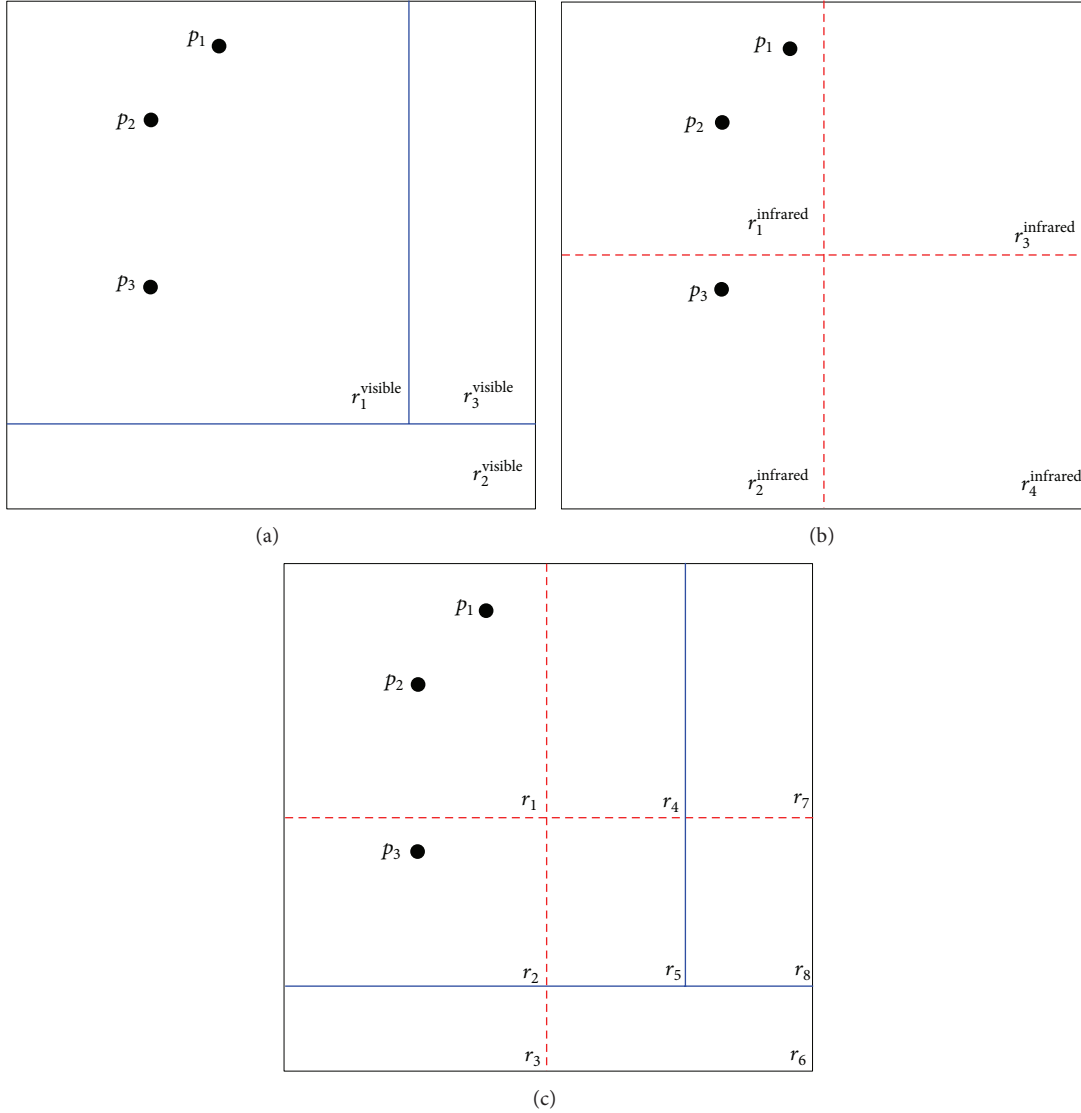


FIGURE 4: A schematic diagram of the integrated region map. (a) The region map  $R^{\text{visible}}$ . (b) The region map  $R^{\text{infrared}}$ . (c) Integrated region map  $R$ .

pixel  $(i, j)$  and pixels of other regions in its 8 neighbors. Consider the following:

$$d(i, j) = \frac{1}{\text{NP}(i, j)} \sum_{s=-1,0,1} |b(i, j) - b(i + s, j + s)|, \quad (4)$$

where  $\text{NP}(i, j)$  is the number of the 8-neighboring pixels which belong to other regions. A difference of neighboring region (DNR) value is used to represent the gray-level difference of a region and its neighboring regions. The  $\text{DNR}(r_k)$  is calculated by averaging all  $d(i, j)$  in region  $r_k$  of the integrated region map,

$$\text{DNR}(r_k) = \frac{1}{\text{NR}_k} \sum_{(i,j) \in \text{boundary pixels of } r_k} d(i, j), \quad (5)$$

where  $\text{NR}_k$  is the number of boundary pixels in region  $r_k$ . An example of calculating DNR is illustrated in Figure 5, where

four regions are separated by bold lines. Take  $p(3, 3) = 43$  as an example,  $d(3, 3) = (|17| + |19| + |21| + |22| + |21|)/5 = 20$ , and  $\text{DNR}(r_3) = (18 + 18.6 + 20 + 21.6 + 23.6 + 24)/6 = 21$ .

A modified sigmoid function  $\text{sig}(x)$  is used to enhance the contrast of DNR. Consider

$$\text{sig}(x) = \frac{1}{1 + e^{-(x-\alpha)/\beta}}, \quad (6)$$

where  $x$  is an input value,  $\alpha$  is the center of the function curve, and  $\beta$  is the width of the curve. The DNR should be normalized to  $[0, 1]$  before applying to the modified sigmoid function. At last,  $\text{sig}(\text{DNR}(r_k))$  should be uniformly normalized to  $[0, 1]$ , which is denoted by  $\text{SDNR}(r_k)$ .  $\text{SDNR}(r_k)$  is calculated by

$$\text{SDNR}(r_k) = \frac{\text{sig}(1) - \text{sig}(\text{DNR}(r_k))}{\text{sig}(1) - \text{sig}(0)}. \quad (7)$$

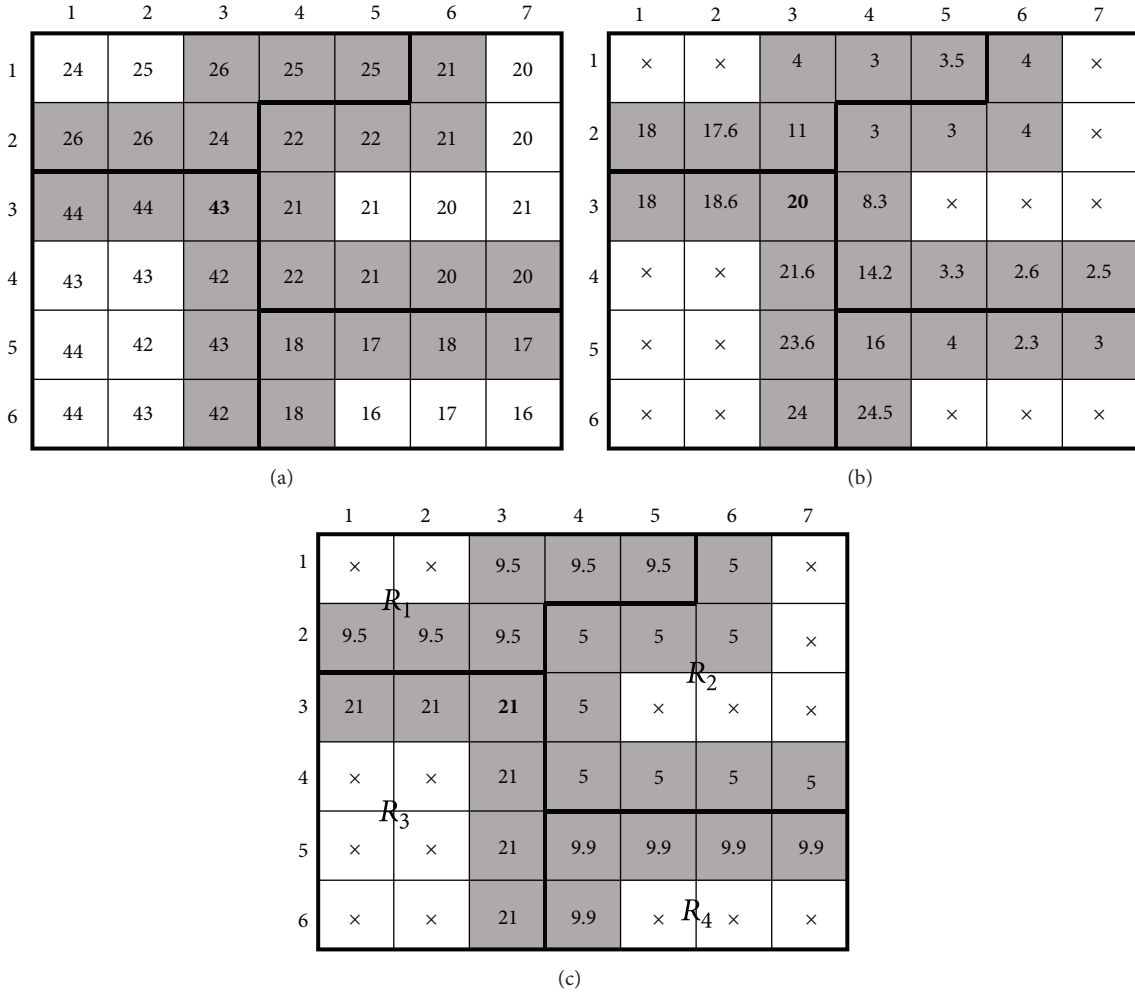


FIGURE 5: The idea of calculating DNR. (a) An input image. (b)  $d(i, j)$ . (c)  $DNR(r_k)$ .

Several curves of the normalized sigmoid function with different values of  $\alpha$  and  $\beta$  are illustrated in Figure 6. Different values of  $\alpha$  and  $\beta$  can control the degree of contrast enhancement.

A decision map consists of weights of regions to indicate the importance of all regions in the input images. Let  $w(r_k)$  be the weight of region  $r_k$  for image  $I_{\text{base}}^{\text{visible}}$ ;  $w(r_k)$  is calculated by

$$\text{if } \text{SDNR}^{\text{visible}}(r_k) < T_{\text{base}}^{\text{visible}}$$

$$\text{and } \text{SDNR}^{\text{infrared}}(r_k) < T_{\text{base}}^{\text{infrared}},$$

$$\text{then } w(r_k) = \frac{\text{SDNR}^{\text{visible}}(r_k)}{\text{SDNR}^{\text{visible}}(r_k) + \text{SDNR}^{\text{infrared}}(r_k)};$$

$$\text{if } \text{SDNR}^{\text{visible}}(r_k) \geq T_{\text{base}}^{\text{visible}}$$

$$\text{and } \text{SDNR}^{\text{infrared}}(r_k) < T_{\text{base}}^{\text{infrared}},$$

$$\text{then } w(r_k) = 1;$$

$$\text{if } \text{SDNR}^{\text{infrared}}(r_k) \geq T_{\text{base}}^{\text{infrared}}$$

$$\text{and } \text{SDNR}^{\text{visible}}(r_k) < T_{\text{base}}^{\text{visible}},$$

$$\text{then } w(r_k) = 0;$$

otherwise

$$\text{if } I_{\text{base}}^{\text{visible}} \text{ is desired, then } w(r_k) = 1,$$

$$\text{if } I_{\text{base}}^{\text{infrared}} \text{ is desired, then } w(r_k) = 0,$$

(8)

where  $T_{\text{base}}^{\text{visible}}$  and  $T_{\text{base}}^{\text{infrared}}$  are predefined threshold values.

**3.5. Fusion Rules.** The fusion rules are applied on the input images region by region. We define different fusion rules for fusing the base layers and detail layers of the input images, both based on the same decision map.

**3.5.1. The Fusion Rules for the Base Layers.** Let  $I_{\text{base}}(r_k)$  represent the  $k$ th regions of the base layer. Based on the

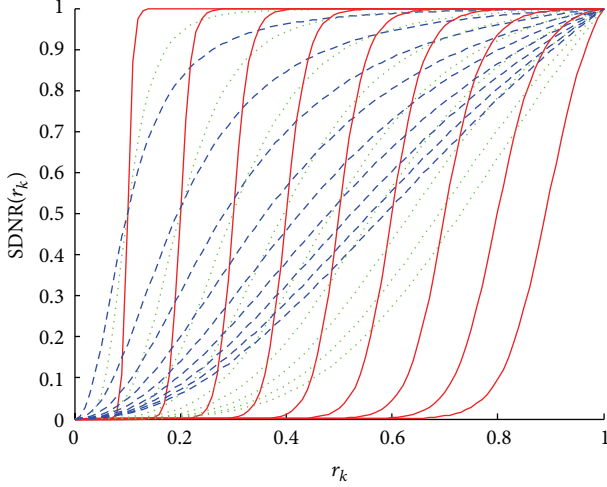


FIGURE 6: The curves of sigmoid functions, when  $\alpha = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$  (from left to right);  $\beta = 0.05$  (solid line),  $0.25$  (dotted line), and  $0.45$  (dashed line).

decision map  $w(r_k)$ , the fusion rules for the base layers are defined as

$$I_{\text{base}}^{\text{fused}}(r_k) = w(r_k) I_{\text{base}}^{\text{visible}}(r_k) + (1 - w(r_k)) I_{\text{base}}^{\text{infrared}}(r_k). \quad (9)$$

The purpose of this rule is to preserve the gray-level difference between a region and its neighboring regions. An example of illustrating the advantage of this strategy is given in Figure 7. The base layers of image  $I^{\text{visible}}$  and image  $I^{\text{infrared}}$  are, respectively, shown in Figures 7(a) and 7(b). The fused results by using the equal-weighted average and the proposed weighting are separately shown in Figures 7(c) and 7(d). In Figure 7(d), the regions of the solid contours keep significant appearance coming from  $I^{\text{visible}}$ , and the region of the dashed contour keeps significant contrast coming from  $I^{\text{infrared}}$ ; the objects in Figure 7(d) are much clearer than that in Figure 7(c).

**3.5.2. The Fusion Rules for the Detail Layers.** Let  $I_{\text{detail}}(r_k)$  represent the  $k$ th regions of the detail layer. In conventional region-based fusion methods, the fusion method is achieved by selecting maximum coefficients or calculating weighted averages in the detail layers. The detail layer contains small-scale details in an image; the entropy of a region can reflect the quantity of information in it. A common way is to set the entropy of each region as its weight for fusion. We think that a region with higher weight in the decision map contains more useful information. Therefore, a region with higher weight in the decision map should be considered in the fused image even if it has smaller entropy in the detail layer; thus, we fuse the detail layers by referring the entropy and the

TABLE 1: The setting of parameters for image fusion.

Input images	$\alpha$	$\beta$	$T_{\text{base}}^{\text{visible}}$	$T_{\text{base}}^{\text{infrared}}$	$\lambda$
UN Camp	0.4	0.3	0.6	0.7	0
Dune	0.3	0.2	0.7	0.7	0
Trees	0.3	0.2	0.5	0.6	0
OCTEC	0.5	0.4	0.6	0.7	0

decision map. The fusion rules for the detail layer are defined as follows:

$$\text{if } E(I_{\text{detail}}^{\text{visible}}(r_k)) > E(I_{\text{detail}}^{\text{infrared}}(r_k)) \text{ and } w(r_k) > 0.5,$$

$$I_{\text{detail}}^{\text{fused}}(r_k) = I_{\text{detail}}^{\text{visible}}(r_k);$$

$$\text{if } E(I_{\text{detail}}^{\text{visible}}(r_k)) < E(I_{\text{detail}}^{\text{infrared}}(r_k)) \text{ and } w(r_k) < 0.5,$$

$$I_{\text{detail}}^{\text{fused}}(r_k) = I_{\text{detail}}^{\text{infrared}}(r_k);$$

otherwise,

$$I_{\text{detail}}^{\text{fused}}(r_k) = w(r_k) I_{\text{detail}}^{\text{visible}}(r_k) + (1 - w(r_k)) I_{\text{detail}}^{\text{infrared}}(r_k), \quad (10)$$

where

$$E(I) = - \sum_{i=0}^{L-1} p(i) \log p(i), \quad (11)$$

and  $L$  is the number of gray levels.  $p(i)$  is the probability of gray level  $i$  in image  $I$ . The proposed detail layer fusing method is consistent with the region importance of the base layer while preserving the features of detail layers.

**3.6. Image Construction.** The fused image  $I^{\text{fused}}$  is obtained by composing the fused base layer  $I_{\text{detail}}^{\text{fused}}$  and the fused detail layer  $I_{\text{base}}^{\text{fused}}$ :

$$I^{\text{fused}} = I_{\text{detail}}^{\text{fused}} + I_{\text{base}}^{\text{fused}}. \quad (12)$$

## 4. Experiments

The proposed method was compared with three previously published methods: discrete wavelet transform (DWT) [3], multiscale directional bilateral filter (MDBF) [10], and visual weight map extraction (VWM) [23]. All test images were downloaded from <http://www.imagefusion.org/> site. The setting of parameters for fusion image is shown in Table 1.

The visible and infrared images of “UN Camp” and the fused results are shown in Figure 8. The challenge of fusing “UN Camp” image is to keep the fence appeared in the visible image (the solid contour in Figure 8(a)) and the person appeared in the infrared image (the dashed contour in Figure 8(b)) remaining in the fused image. These two objects are not clear enough in the fused images by the

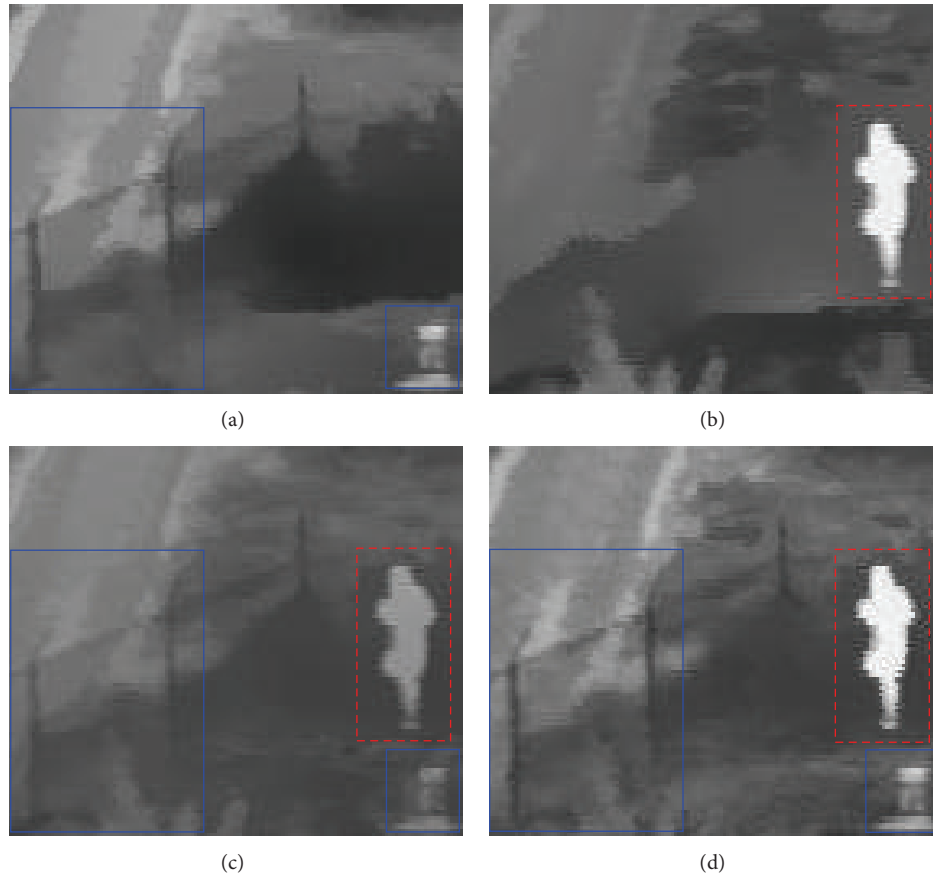


FIGURE 7: An example of fused base layer. (a) The base layer of a visible image  $I^{\text{visible}}$ . (b) The base layer of the corresponding infrared image  $I^{\text{infrared}}$ . (c) The fused result by taking equal-weighted average. (d) The fused result of the proposed method.

weighted averaging methods (DWT and MDBF) and the VWM method. The DWT and MDBF methods keep the detail as much as possible, but they enhance noise in the fused image, too. The meaningful detail and saliency of the subject have better transferred into the fused image in the proposed method than in the other three methods.

The visible and infrared images of the low-contrast “dune” image and the fused results are shown in Figure 9. The fused image of the proposed method preserves the hill in the visible image (the solid contour in Figure 9(a)) and a pedestrian in the infrared image (the dashed contour in Figure 9(b)). The VWM method obtains a better result than DWT and MDBF methods, but the pedestrian obtained by the VWM method is less clear than that by the proposed method. The proposed method preserves the most saliency of the two objects in the fused image.

The visible and infrared images of “trees” and the fused results are shown in Figure 10. In the visible image, a pedestrian is not clear due to the low contrast to the background (the dashed contour in Figure 10(a)). In the infrared image, the pedestrian is clear (the dashed contour in Figure 10(b)), but the background is less separable than that of the visible image. The challenge of fusing these two images is to preserve the clear pedestrian and the detail background in the fused image. The other three methods preserve less detail

background in their fused images. The proposed method can simultaneously highlight the pedestrian and preserve the detail background in the fused image.

The visible and infrared images of “OCTEC” and the fused results are shown in Figure 11. The challenge of the fusion is to highlight the pedestrian appearing in the infrared image (the dashed contour in Figure 11(b)) while keeping the detail background appearing in the visible image. The pedestrian is not highlighted enough in the fused images of averaging methods (DWT and MDBF) as shown in Figures 11(c) and 11(d). The detail of the background in the visible image is not kept enough in the fused images of the other three methods. Conversely, the proposed method preserves the most saliency on both pedestrian and background.

Not only the above qualitative evaluation, but also the two objective indices, entropy [24] and weighted fusion quality index [25], were used to quantitatively evaluate the performance of the mentioned fusion methods.

The entropy measures the quantity of information in an image. Entropy is defined as in (11). The weighted fusion quality index  $Q_w$  is calculated between the fused image and the input images to measure how well the input images are fused. The index  $Q_w$  reflects how much salient information



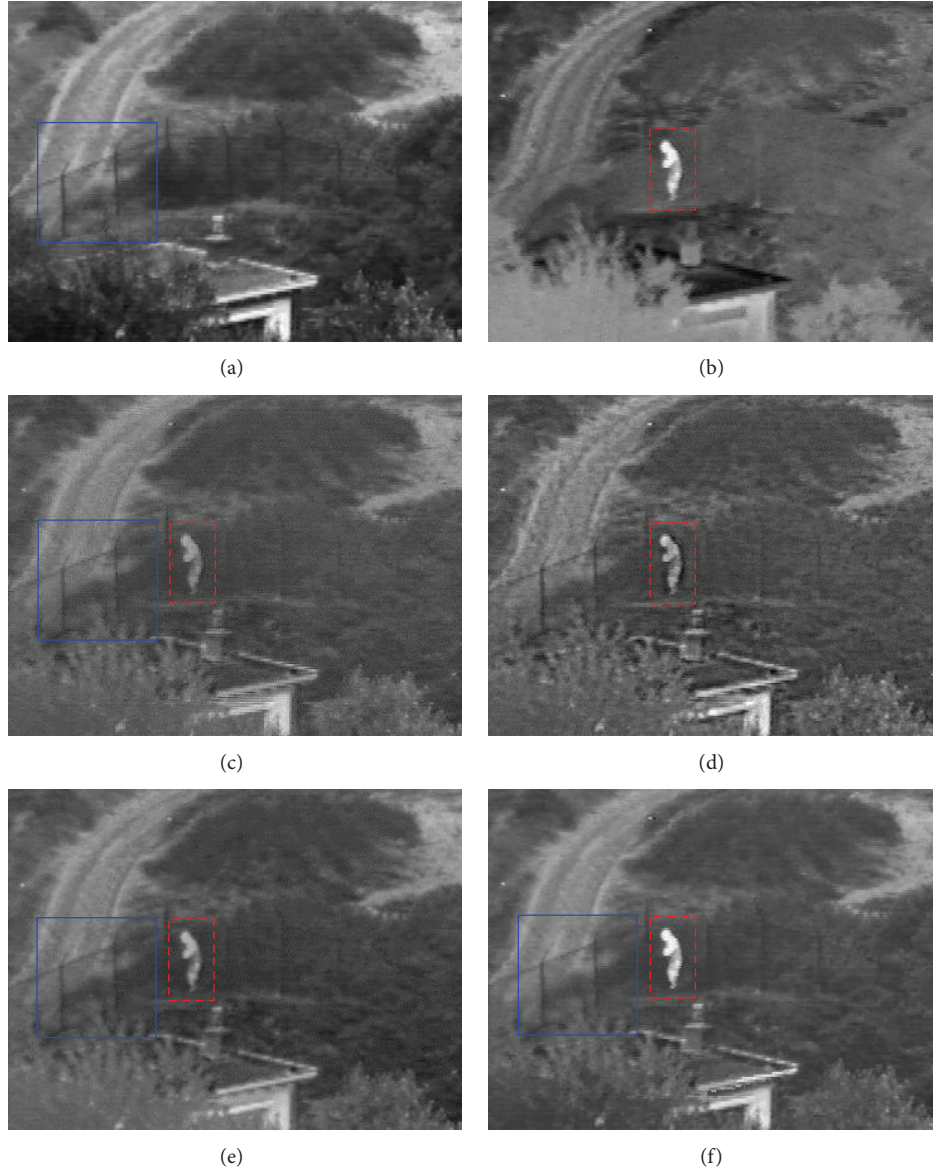


FIGURE 8: The fused results of different fusion methods on “UN Camp” images. (a) Visible image. (b) Infrared image. (c) DWT method. (d) MDBF method. (e) VWM method. (f) The proposed method.

in the input images is transferred into the fused image.  $Q_w$  is defined as

$$\begin{aligned}
 Q_w(I^A, I^B | I^F) &= \sum_{w \in W} m(w) (\lambda(w) Q_0(I^A, I^F | w) \\
 &\quad + (1 - \lambda(w)) Q_0(I^B, I^F | w)), \quad (13)
 \end{aligned}$$

where  $m(w) = \max(s(I^A | w), s(I^B | w))$ ;  $s(I | w)$  reflects the local entropy of image  $I$  within window  $w$ . The local weight

$\lambda(w)$  indicates the relative importance of image  $I^A$  compared to image  $I^B$ .  $\lambda(w)$  is defined as

$$\lambda(w) = \frac{s(I^A | w)}{s(I^A | w) + s(I^B | w)}. \quad (14)$$

$Q_0$  is a local quality index which is calculated by

$$Q_0(x, y) = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\bar{x}^2 + \bar{y}^2)(\sigma_x^2 + \sigma_y^2)}, \quad (15)$$

where  $\bar{x}$  and  $\bar{y}$  denote the means of two real-valued sequences  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$ , respectively. In this study, values  $x_i$  and  $y_i$  are gray levels.  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_{xy}$

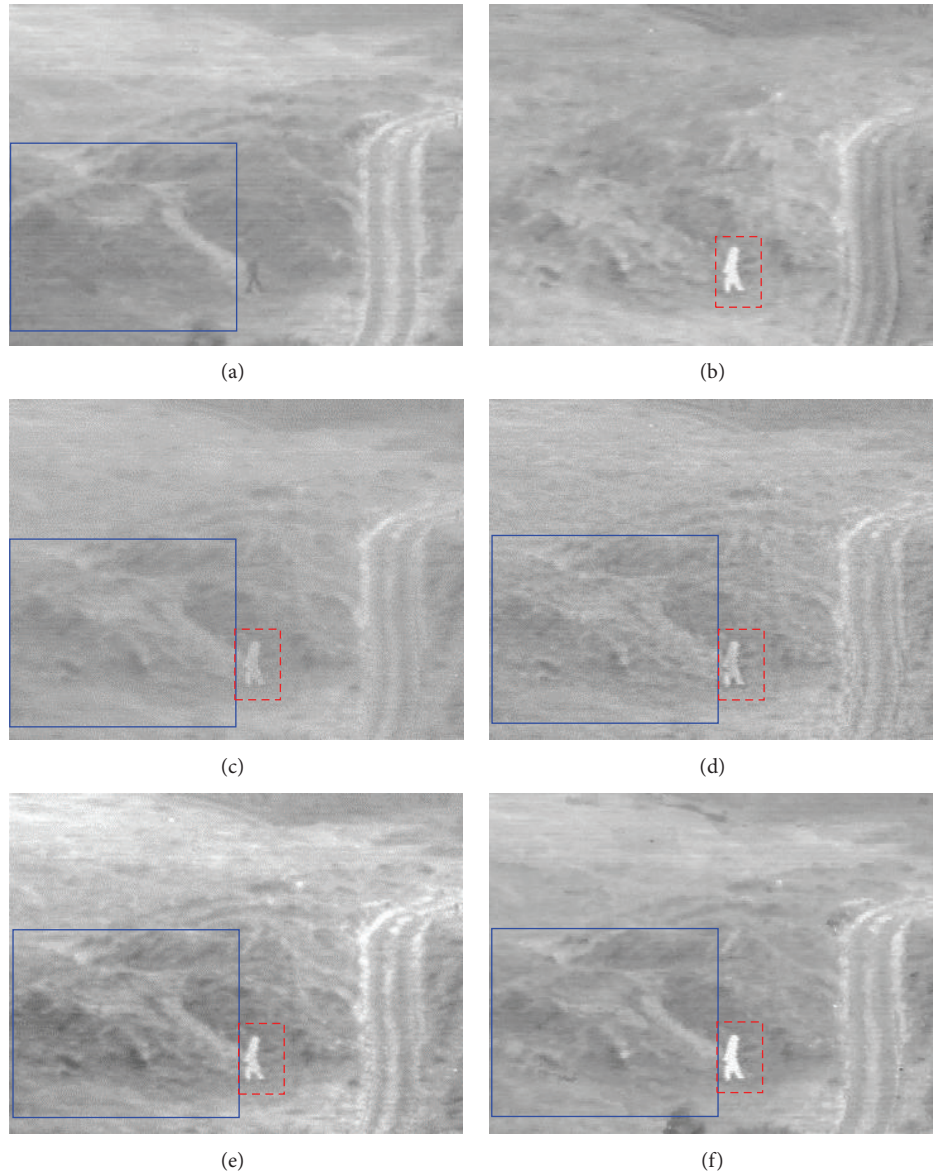


FIGURE 9: The fused results of different fusion methods on “dune” images. (a) Visible image. (b) Infrared image. (c) DWT method. (d) MDBF method. (e) VWM method. (f) The proposed method.

are variance of  $x$ , variance of  $y$ , and covariance of  $x$  and  $y$ , respectively.

The larger entropy value which reflects more information is contained in an image. A larger  $Q_w$  value reflects more salient information contained in each of the input images having been transferred into the fused image. The experimental results of entropy and  $Q_w$  values of the proposed method and the other three methods for the four test images are shown in Table 2. The proposed method gets the highest entropy and  $Q_w$  values for all test images. The experimental results reveal that the proposed method not only contains the most quantitative information, but also preserves the most saliency of the input images in the fused image.

## 5. Discussions and Conclusions

In this paper, we proposed a region-based method for fusing spatially registered visible and infrared images. The advantages of the proposed method are improving the contrast of the fused images and keeping the saliency of objects in the fused images. The contributions are described as follows:

- (1) segmenting the base layers of the input images instead of the original images such that the over-segmentation problem is then reduced,
- (2) using the decision map to indicate the importance of every region according to the gray-level differences

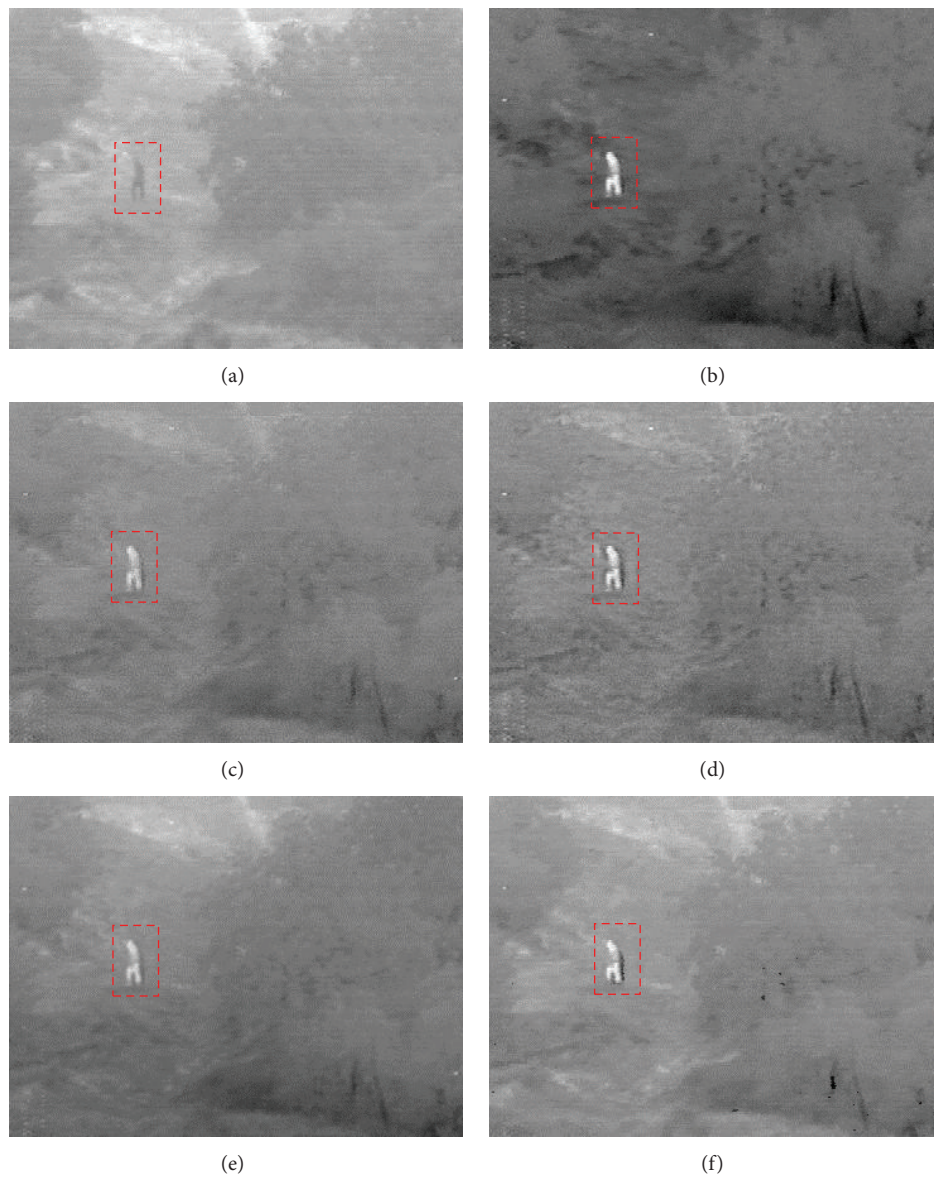


FIGURE 10: The fused results of different fusion methods on “Trees” images. (a) Visible image. (b) Infrared image. (c) DWT method. (d) MDBF method. (e) VWM method. (f) The proposed method.

TABLE 2: The performance of four methods on the four famous images.

Input images	Evaluation measurement	DWT	MDBF	VWM	The proposed
UN Camp	Entropy	6.3410	6.4597	6.4015	<b>6.6652</b>
	$Q_w$	0.6218	0.6717	0.6667	<b>0.7045</b>
Dune	Entropy	5.7882	5.8246	5.9515	<b>5.9816</b>
	$Q_w$	0.6783	0.6954	0.5119	<b>0.7437</b>
Trees	Entropy	5.9692	5.9984	6.1500	<b>6.1740</b>
	$Q_w$	0.6662	0.6943	0.6746	<b>0.7636</b>
OCTEC	Entropy	6.3901	6.4649	6.5500	<b>6.5824</b>
	$Q_w$	0.7523	0.8489	0.8370	<b>0.8999</b>

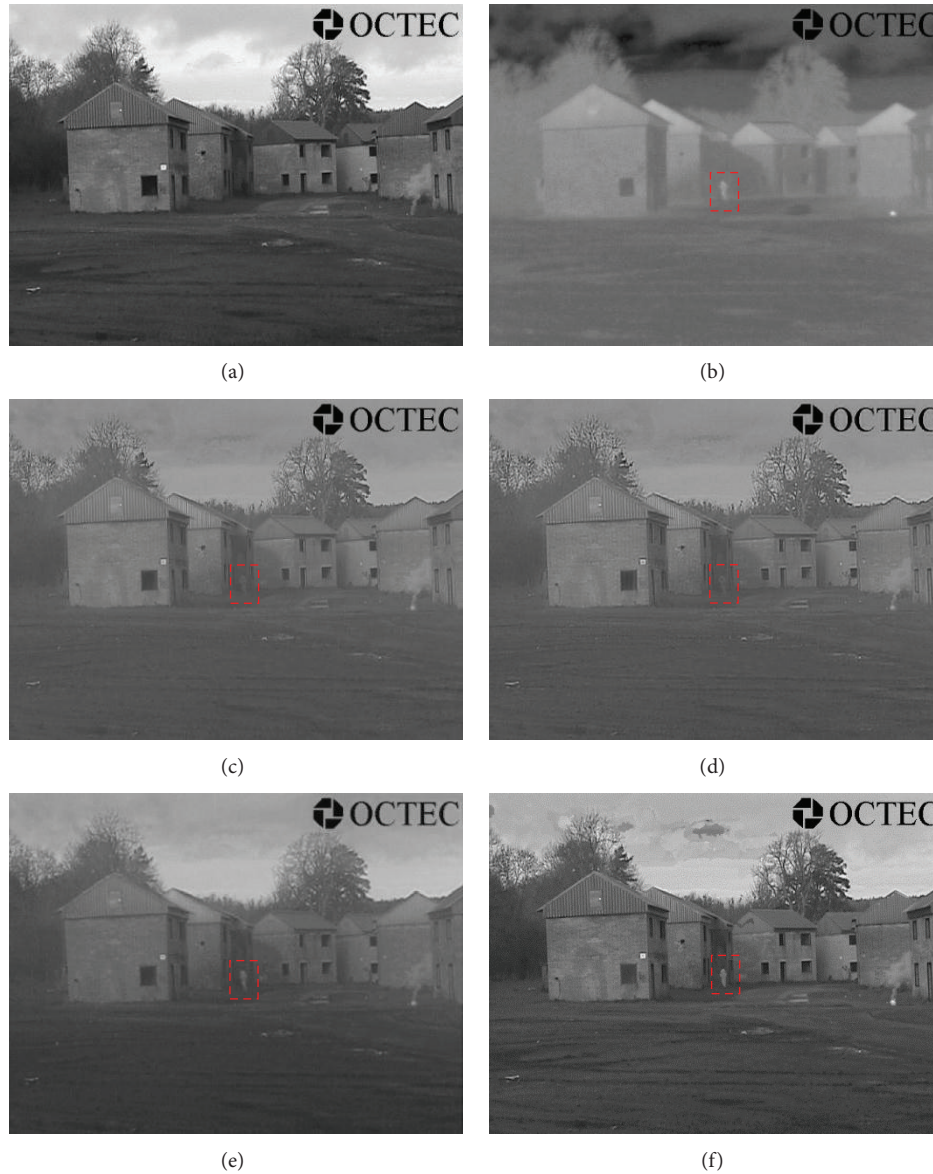


FIGURE 11: The fused results of different fusion methods on “OCTEC” images. (a) Visible image. (b) Infrared image. (c) DWT method. (d) MDBF method. (e) VWM method. (f) The proposed method.

between a region and its neighboring regions such that the contrast of regions is remained in the final fused images,

- (3) accompanying with the decision map, the proposed fusing rules preserve the significant features and saliency properties of input images in the fused image.

Experimental results show that the proposed method achieves superior results over the other three fusion methods in improving contrast of the fused images and preserving saliency of objects in the fused images.

There is still a drawback in the proposed fusion method. In the proposed method, an object may be segmented into several regions. If an object has partial small gray levels in one input image and partial high gray levels in another

input image, the object will be segmented into several regions and the regions may be fused by different fusion rules. Consequentially, the fused image may show an unnatural appearance. As one example shown in Figure 12, the sky is composed of several different light regions in the visible image and is composed of several uneven dark regions in the infrared image. The sky was first segmented into several regions due to the uneven gray regions in the infrared image. Then, the sky shows unnatural appearance in the fused image as shown in the dashed block. In conclusion, if an object is oversegmented, the object may show an unnatural appearance in the fused image, named as the inconsistency problem. The proposed approach can only alleviate this phenomenon; it cannot completely eliminate the problem. In the future, we will design more robust fusion rules and

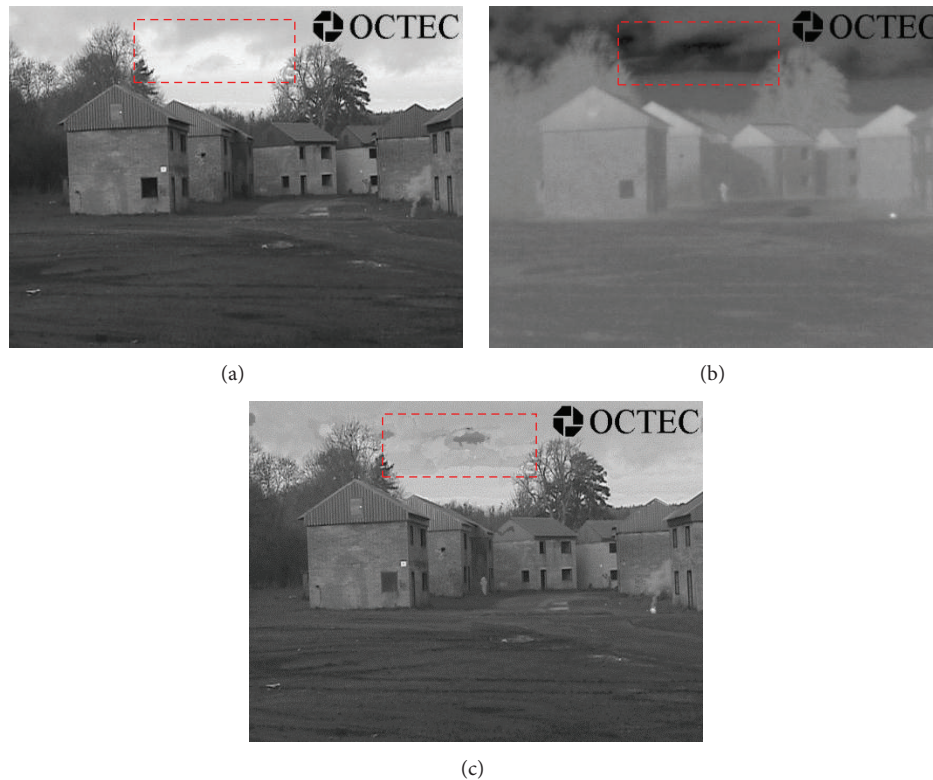


FIGURE 12: The fused result of the proposed method on “OCTEC” images. (a) Visible image. (b) Infrared image. (c) Fused image.

better image segmentation method to decrease the unnatural artifact in the fused images.

### Conflict of Interests

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

### Acknowledgment

The authors would like to thank the *TNO* Human Factors Research Institute for their kind support of supplying the original infrared and visible images of “UN Camp” and “OCTEC.” All images are available on <http://www.imagefusion.org>.

### References

- [1] A. A. Goshtasby and S. Nikolov, “Image fusion: advances in the state of the art,” *Information Fusion*, vol. 8, no. 2, pp. 114–118, 2007.
- [2] C. Pohl and J. L. Van Genderen, “Multisensor image fusion in remote sensing: concepts, methods and applications,” *International Journal of Remote Sensing*, vol. 19, no. 5, pp. 823–854, 1998.
- [3] G. Pajares and J. M. de la Cruz, “A wavelet-based image fusion tutorial,” *Pattern Recognition*, vol. 37, no. 9, pp. 1855–1872, 2004.
- [4] V. S. Petrović and C. S. Xydeas, “Gradient-based multiresolution image fusion,” *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 228–237, 2004.
- [5] M. Ding, L. Wei, and B. Wang, “Research on fusion method for infrared and visible images via compressive sensing,” *Infrared Physics and Technology*, vol. 57, pp. 56–67, 2013.
- [6] S. Li and B. Yang, “Multifocus image fusion using region segmentation and spatial frequency,” *Image and Vision Computing*, vol. 26, no. 7, pp. 971–979, 2008.
- [7] X. Luo, J. Zhang, and Q. Dai, “A regional image fusion based on similarity characteristics,” *Signal Processing*, vol. 92, no. 5, pp. 1268–1280, 2012.
- [8] Z. Zhang and R. S. Blum, “A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application,” *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315–1326, 1999.
- [9] J. J. Lewis, R. J. O’Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, “Pixel- and region-based image fusion with complex wavelets,” *Information Fusion*, vol. 8, no. 2, pp. 119–130, 2007.
- [10] J. Hu and S. Li, “The multiscale directional bilateral filter and its application to multisensor image fusion,” *Information Fusion*, vol. 13, no. 3, pp. 196–206, 2012.
- [11] L. Guo, M. Dai, and M. Zhu, “Multifocus color image fusion based on quaternion curvelet transform,” *Optics Express*, vol. 20, no. 17, pp. 18846–18860, 2012.
- [12] P. Wang, H. Tian, and W. Zheng, “A novel image fusion method based on FRFT-NSCT,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 408232, 9 pages, 2013.
- [13] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, “Edge-preserving decompositions for multi-scale tone and detail manipulation,” *ACM Transactions on Graphics*, vol. 27, no. 3, article 67, 2008.

- [14] 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.
- [15] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the 6th IEEE International Conference on Computer Vision*, pp. 839–846, Bombay, India, January 1998.
- [16] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, "Bilateral filtering: theory and applications," *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1–73, 2009.
- [17] J. Zhao, Q. Zhou, Y. Chen, H. Feng, Z. Xu, and Q. Li, "Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition," *Infrared Physics and Technology*, vol. 56, pp. 93–99, 2013.
- [18] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information Fusion*, vol. 4, no. 4, pp. 259–280, 2003.
- [19] T. Wan, N. Canagarajah, and A. Achim, "Segmentation-driven image fusion based on alpha-stable modeling of wavelet coefficients," *IEEE Transactions on Multimedia*, vol. 11, no. 4, pp. 624–633, 2009.
- [20] J. Saeedi and K. Faez, "Infrared and visible image fusion using fuzzy logic and population-based optimization," *Applied Soft Computing Journal*, vol. 12, no. 3, pp. 1041–1054, 2012.
- [21] J. W. Han, J. H. Kim, S. H. Cheon, J. O. Kim, and S. J. Ko, "A novel image interpolation method using the bilateral filter," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 1, pp. 175–181, 2010.
- [22] P. Zhou, W. Ye, Y. Xia, and Q. Wang, "An improved Canny algorithm for edge detection," *Journal of Computational Information Systems*, vol. 7, no. 5, pp. 1516–1523, 2011.
- [23] J. Zhao, H. Feng, Z. Xu, Q. Li, and T. Liu, "Detail enhanced multi-source fusion using visual weight map extraction based on multi scale edge preserving decomposition," *Optics Communications*, vol. 287, pp. 45–52, 2013.
- [24] U. Manjusha, "Image fusion and image quality assessment of fused images," *International Journal of Image Processing*, vol. 4, no. 5, pp. 484–508, 2010.
- [25] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *proceedings of the 10th International Conference on Image Processing*, pp. 173–176, Barcelona, Spain, September 2003.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

