Infrared Physics and Technology 92 (2018) 68–77

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Regular article

General fusion method for infrared and visual images via latent low-rank representation and local non-subsampled shearlet transform

Boyang Cheng[a,b,*], Longxu Jin[a], Guoning Li[a]

[a] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:
Latent low-rank representation
Local non-subsampled shearlet transform
Salient features
Image fusion

ABSTRACT

This study establishes the general fusion method for infrared and visual images via latent low-rank representation (LatLRR) and local non-sampled shearlet transform (LNSST) to effectively combine the salient information of both images and solve problems on low-contrasting heterogeneous image fusion. First, LNSST is used as a multi-scale analysis tool to decompose the source images into low-pass and high-pass sub-images. Second, the LatLRR, which is an effective method for exploring multiple subspace structural data, is used to extract the salient information of image sources. Thus, the LatLRR can be adopted to guide the adaptive weighted fusion of low-pass sub-images. Simultaneously, the average gradient, which can reflect image edge details, is regarded as the fusion rule for high-pass sub-images. A series of images from diverse scenes are used for the fusion experiments, and the results are evaluated subjectively and objectively. The subjective and objective evaluations show that our algorithm exhibited superior visual performance, and the values of the objective evaluation parameters increase by about 5–10% compared with other typical fusion methods.

1. Introduction

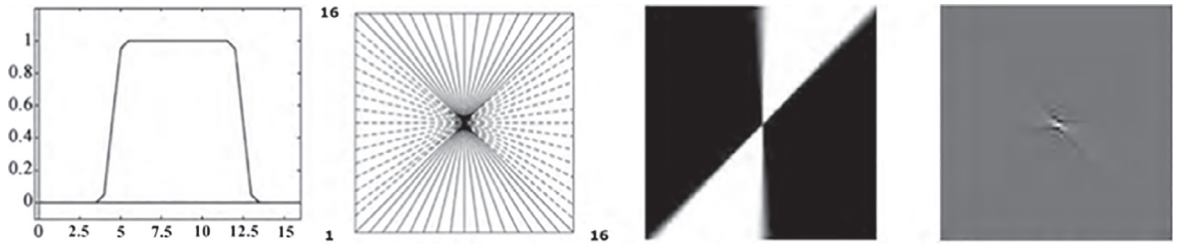
The development of infrared (IR) and visible image fusion technology is largely aimed at developing modern military detection technology. A visible image (VI) is a reflection image with several high-frequency components, and VI images can reflect scene details under certain illumination conditions. However, when illumination is not good, the resultant contrast of the VI image is relatively low. Meanwhile, an IR image is a radiation image. The gray level of IR images is determined by the temperature difference between the target and background, but resultant images cannot reflect real scenes [1]. Image fusion technology for IR and VI images can effectively synthesize and explore the combined characteristic information of two complementary images with the same resolution, enhance the understanding of a scene, and highlight image targets; thus, image fusion technology can find objects quickly and accurately despite confusing situations [2].

Several fusion approaches have been recently proposed, especially for pixel-level-based VI and IR image fusion [3]. A number of multi-scale analysis tools, such as contourlet transform (CT) [4], non-subsampled contourlet transform (NSCT) [5], and local non-subsampled shearlet transform (LNSST) [6], have been successfully used in the field of image fusion. LNSST is regarded the fastest MGA tool with the most disaggregation effect. Furthermore, LNSST can exhibit good local characteristics in space and frequency domains, avoid blocking effects, weaken the Gibbs-ringing phenomenon by using local small-sized shearlet filters, and improve the calculation efficiency of time domain convolutions. Thus, many researchers favor LNSST over other techniques. Lei et al. [7] proposed an adaptive fusion method based on the LNSST and non-negative matrix factorization to construct an algorithm that could guide the fusion of low-frequency coefficients, but the final fused images were dim and lost considerable textural details. Zhang et al. [8] presented a fusion algorithm based on saliency analysis and LNSST. This method utilized saliency detection to integrate IR target information into the VI image, but the fusion effect of background information required improvements. Wu et al. [9] combined LNSST and deep Boltzmann machine programming to solve fusion problems, but deep learning technology is not yet mature for fusion applications. Wang et al. [10] proposed a fusion algorithm for RDU-PCNN and ICA bases in the LNSST domain. Although the PCNN has a bionic mechanism, the final fused images introduce artifacts and have hazy image edges. Kong et al. [11] forwarded a technique for gray-scale VI and IR image fusion based on the LNSST. This method makes use of regional averaged energy and local directional contrast, but the fused image loses some important IR saliency information.

Latent low-rank representation (LatLRR) [12], an upgraded version

* Corresponding author at: Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China.
E-mail address: boyangwudi@163.com (B. Cheng).

https://doi.org/10.1016/j.infrared.2018.05.006
Received 23 December 2017; Received in revised form 23 April 2018; Accepted 7 May 2018
Available online 16 May 2018
1350-4495/



(a) Meyer wavelet window **(b)** Pseudo-polarized coordinates **(c)** Frequency domain shearlet filter **(d)** Time domain shearlet filter

Fig. 1. Shearlet filter formation procedure when $L = 15$.

of the low-rank representation (LRR), is an effective method for exploring the multiple subspaces of data structures. LatLRR can robustly extract salient features from images because the method utilizes an unsupervised feature extraction algorithm. Moreover, LatLRR is robust to noise. When an image matrix is decomposed by LatLRR, the image can be represented as a superposition of principal features, salient features, and sparse noise. Salient features show the spatial distribution of salient information of images, and the weighted-mean is usually treated as the fusion rule for exploring the low contrast and unnatural reconstruction of heterogeneous fused images. The LatLRR algorithm can precisely identify salient objects and regions in images to form a saliency map. The saliency map contains the weight information of the spatial distribution of a grayscale image and thus integrates a weighting function. The fusion rule can be changed from the weighted-mean to the weighted-adaptive approach to effectively merge the salient information into the fused image and improve the fusion effect.

Based on the above review, this study proposes a general fusion method for IR and visual images via LatLRR and LNSST. To the best of our knowledge, this is the first time that the LatLRR has been used in the field of heterologous image fusion, in which the heterologous source images have the same resolution [13]. In this study, the LNSST is first used as a multi-scale analysis tool to decompose image sources into a low-pass sub-image and a series of high-pass sub-images. Second, the saliency information of the image is extracted by LatLRR to guide the adaptive weighted fusion of low-pass sub-image and high-pass sub-images. Finally, each sub-image is modeled and the corresponding fusion coefficients are produced. An algorithm is adopted to effectively express image characteristics and obtain a good fusion effect by using IR and visible light images in the fusion experiments.

The remainder of this paper is organized as follows: Section 2 introduces the theory relevant to LNSST and LatLRR. Section 3 elaborates the algorithm based on the new fusion rule. Section 4 presents five experimental results and intuitively compares the proposed method with other methods. Section 6 provides a summary of the findings.

2. Relevant theory

2.1. LNSST [14,15]

When the dimension is $n = 2$, the shearlet system function with discrete parameters is as follows:

$$S_{AB}(\varphi) = \{\varphi_{j,l,k} = |\det A|^{j/2}\varphi(B^l A^j x - k);\ j, l \in Z,\ k \in Z^2\}. \tag{1}$$

where $\varphi \in L^2(R^2)$, A, and B are 2×2 reversible matrices; $|\det B| = 1$; j is the scale parameter; l is the direction parameter; and k is the spatial position.

For $j \geq 0$, $-2^j \leq l \leq 2^j - 1$, $k \in Z^2$, and $d = 0,1$, the Fourier transform of the shearlet can be expressed on the basis of the tight support frame.

$$\hat{\varphi}_{j,l,k}^{(d)} = 2^{3j/2} V(2^{-2j}\xi) W_{j,l}^{(d)}(\xi) e^{-2\pi i \xi A_d^{-j} B_d^{-l} k}. \tag{2}$$

where $V(2^{-2j}\xi)$ is the scale function; $W_{j,l}^{(d)}$ is the window function localized on the trapezoidal pair; A_d is the heterosexual expansion matrix; and B_d is the shear matrix. The shearlet transform of the $f \in L^2(R^2)$ function can be calculated by Eq. (3).

$$\langle f, \hat{\varphi}_{j,l,k}^{(d)} \rangle = 2^{3j/2} \int_{R^2} \hat{f}(\xi) \overline{V(2^{-2j}\xi) W_{j,l}^{(d)}(\xi)} e^{-2\pi i \xi A_d^{-j} B_d^{-l} k} d\xi. \tag{3}$$

As shown in Eq. (3), the shearlet transform is divided into two steps. The first step is a multi-scale decomposition [i.e., $\hat{f}(\xi)\overline{V(2^{-2j}\xi)}$] and the second step is the direction of the localization, i.e., $\hat{f}(\xi)\overline{V(2^{-2j}\xi) W_{j,l}^{(d)}(\xi)}$.

Multi-scale decomposition: The image is subjected to non-subsampled pyramid decomposition using a non-subsampled 2*D* filter bank of dual channels to generate a low-pass sub-image and multiple high-pass sub-images with perfect reconstruction.

Directional localization: Directional localization is achieved by small-scale shearlet filters and high-pass sub-images convolution calculations. The local window is $L \times L$, where $L = n \cdot (2^{j-1} + 1)$ with j as the scale parameter and n as any positive integer. The local small-size shearlet filter can avoid the blocking effect, weaken the Gibbs-ringing phenomenon, and improve the calculation efficiency of time domain convolution. Thus, j is usually 2 or 3 and the local window is usually 15×15.

The above shearlet transformation is called the LNSST, a technique that removes the sampling operation in the decomposition stage. LNSST involves translation invariance because the local small-size shearlet filter can avoid spectrum aliasing to improve image decomposition and reconstruction. The shearlet filter formation process for L = 15 is shown in Fig. 1.

The image f is decomposed by the m-layer LNSST to obtain $\Sigma_m 2^{dm}$ high-pass directional sub-images and a low-pass sub-image. Each sub-image is the same size as the original image, and d_m is the number of m-layers of directional localization. A two-layer LNSST decomposition of Linda is shown in Fig. 2. The number of high-pass sub-images in the first layer is 4 (the number of stages is 2), the number of high-pass sub-images in the second layer is 4 (the number of stages is 2), and the size of the shearlet filter is 15×15.

2.2. LatLRR

The image sources of the fusion usually contain a certain amount of noise, but LatLRR can automatically extract salient features from noisy images. LatLRR is robust to noise, and the saliency map obtained is more accurate than other saliency detection-based methods [16–18].

The core idea behind LatLRR is that an image matrix can be represented as a superposition of principal features, salient features, and sparse noise given the low rank and sparse optimization criteria. For an image matrix $X \in R^{M \times N}$, the idea may be interpreted as

$$X = XL + SX + E. \tag{4}$$

where L represents the low-rank matrix, $L \in R^{N \times N}$; S represents the sparse matrix, $S \in R^{M \times M}$; E represents the sparse noise, $E \in R^{M \times N}$; XL

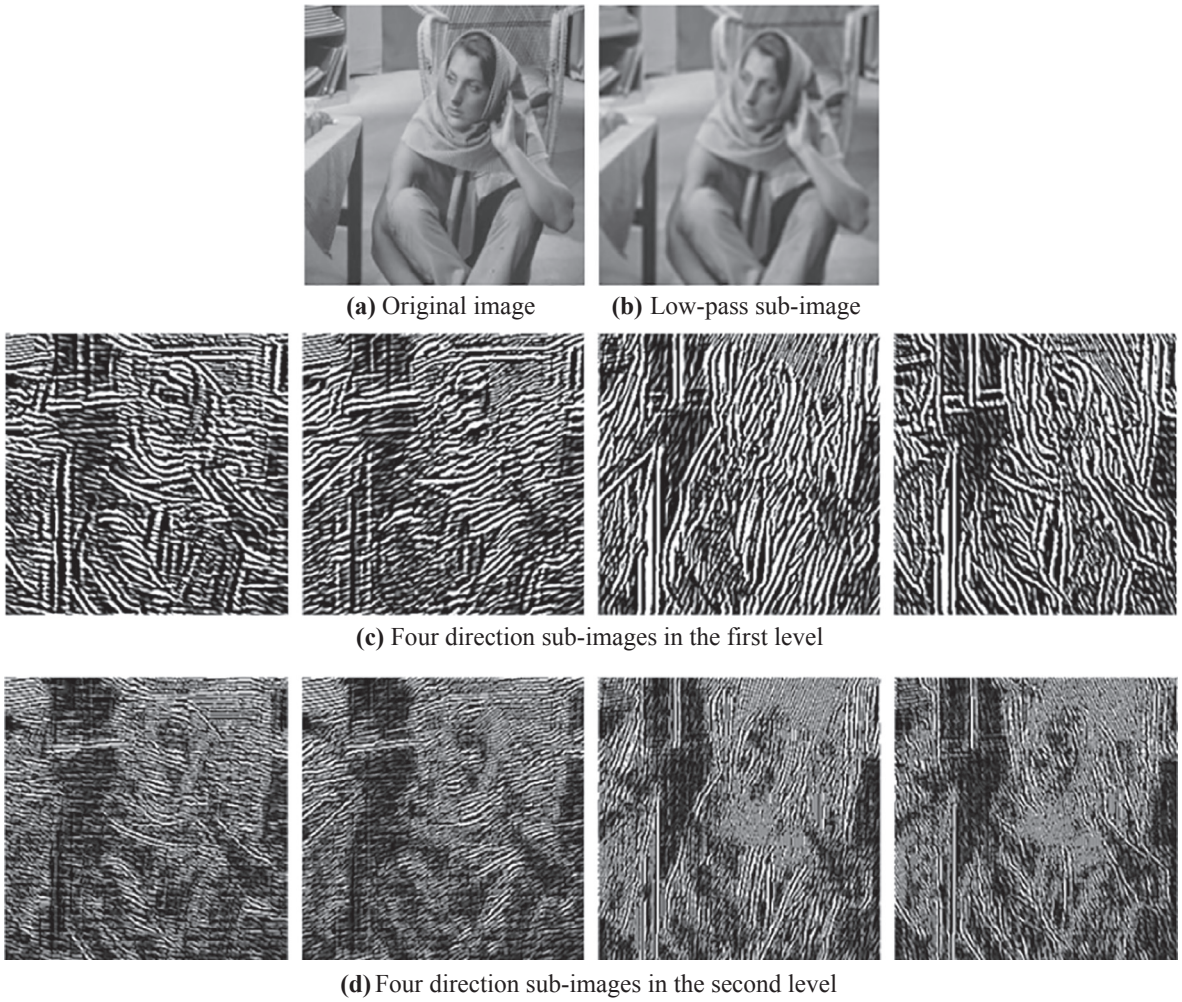


Fig. 2. Two-level LNSST of the image Linda.

represents an approximate part of the image; and SX represents the saliency information of the image.

We can use the convex optimization function to solve the problem of Eq. (4) (i.e., the norm is minimized). The expression is as follows:

$$\min_{L,S,E} \|L\|_* + \|S\|_* + \lambda \|E\|_1, \ s.t. \ X = XL + SX + E. \tag{5}$$

where $\boldsymbol{\lambda} > 0$, $||\ ||_*$ denotes the nuclear norm of a matrix (i.e., sum of singular values of the matrix) and $||\ ||_1$ represents the l_1-norm (i.e., sum of absolute values of all matrix elements). An example of the decomposition of IR and VI images for the same scene by LatLRR is shown in Figs. 3 and 4. Figs. 3(b) and 4(b) show the main features of the image, i.e., the approximate part; Figs. 3(c) and 4(c) present the salient features of the image; and Figs. 3(d) and 4(d) represent sparse noise.

The IR image can be easily determined by the temperature difference between the target and the background, and highlighting the significant features in the interior can be obtained through LatLRR decomposition. The VI image also contains rich scene information and textural details. By using the same technique, all salient information in the VI image can be obtained through LatLRR.

3. Fusion method

3.1. Fusion framework

The fusion algorithm used in this study aims at fusing heterogeneous images. The specific fusion framework is shown in Fig. 5, and the proposed fusion steps can be summarized as follows:

(1) LNSST is used for the multi-scale decomposition of VI and IR images, in which the low-pass sub-image coefficients and high-pass

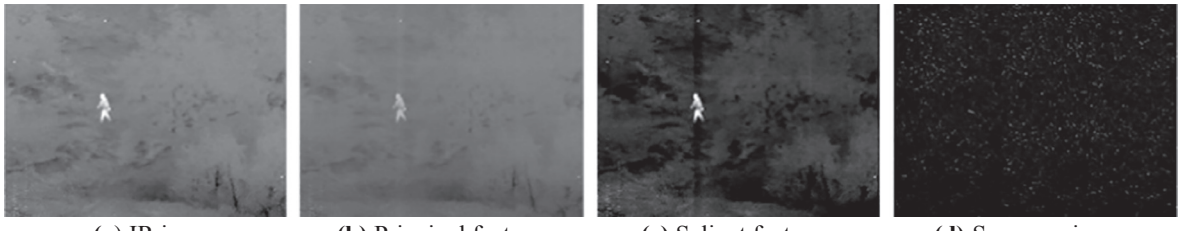


Fig. 3. LatLRR mechanism of IR image decomposition.

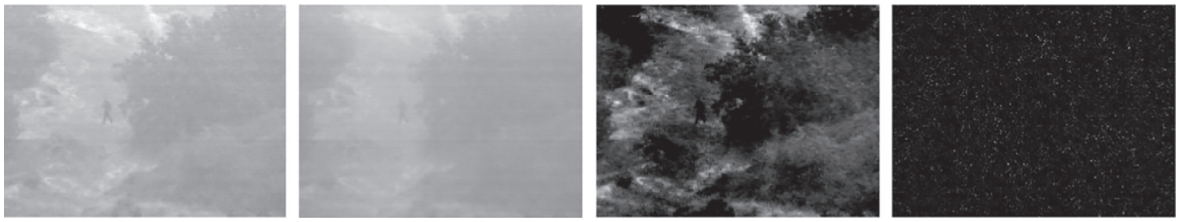


Fig. 4. LatLRR mechanism of VI image decomposition.

sub-images coefficients are $[C_{l,k}^{IR}(i,j), C_{l,k}^{VI}(i,j)]$ and $[D_{l,k}^{IR}(i,j), D_{l,k}^{VI}(i,j)]$; l is the number of decomposition layers; k represents the direction of decomposition of each layer; and (i, j) represents the position of the pixel element.

(2) IR and VI images are decomposed by LatLRR simultaneously, then their salient features are obtained to form saliency maps. Saliency maps contain the spatial information distribution of salient targets. Thus, saliency maps can be used to guide the adaptive-weighted fusion of low-pass sub-image and high-pass sub-images.

(3) The coefficients of high-pass sub-images contain the edge energy information of the image. Thus, the average gradient (AVG) is calculated and compared as a part of the fusion rule.

(4) The fused image is reintegrated based on the inverse LNSST.

3.2. Fusion rule of low-pass sub-images

The low-pass component of the image represents its main energy, which is an approximate part of the image. Thus, the fusion rules of low-pass components determine the final fusion effect. LatLRR algorithm is used to guide the fusion of low-pass sub-image and avoid the disadvantage of incompatible special characteristics of heterologous images. The LatLRR algorithm can also decompose the salient features of the image, which provide the basis for the rules of adaptive-weighted fusion. The fusion rules of the low-pass sub-image are provided below.

Step 1: The LatLRR algorithm is used to obtain the saliency features of IR and VI images. The corresponding saliency maps of S_{IR} and S_{VI} are generated. The gray values of the saliency maps are normalized to form the weighting coefficient matrices of S_1 and S_2.

$$S_1(i,j) = \frac{S_{IR}(i,j) - \min S_{IR}(i,j)}{\max S_{IR}(i,j) - \min S_{IR}(i,j)}. \tag{6}$$

$$S_2(i,j) = \frac{S_{VI}(i,j) - \min S_{VI}(i,j)}{\max S_{VI}(i,j) - \min S_{VI}(i,j)}. \tag{7}$$

Step 2: S_1 and S_2 are used to guide adaptive-weighted fusion based on the saliency extraction of IR and VI images, respectively. The specific expression is as follows:

$$C_{l,k}(i,j) = \frac{\begin{array}{l} S_1(i,j)\cdot C_{l,k}^{IR}(i,j) + (1-S_1(i,j))\cdot C_{l,k}^{VI}(i,j) + (1-S_2(i,j))\cdot C_{l,k}^{IR}(i,j) \\ + S_2(i,j)\cdot C_{l,k}^{VI}(i,j) \end{array}}{2}. \tag{8}$$

where $C_{l,k}(i,j)$ represents the final low-pass fusion coefficients. The saliency map contains the weight information of the spatial distribution of the grayscale value of the image, so this fusion method adaptively integrates the salient information of an IR image into the VI image with textural details. Since the spectral difference of the two source images can be compensated by the adaptive-weighted of the saliency map, the problem of low-contrasting fused image will be solved. Similarly, the saliency information of the two source images combined appropriately,

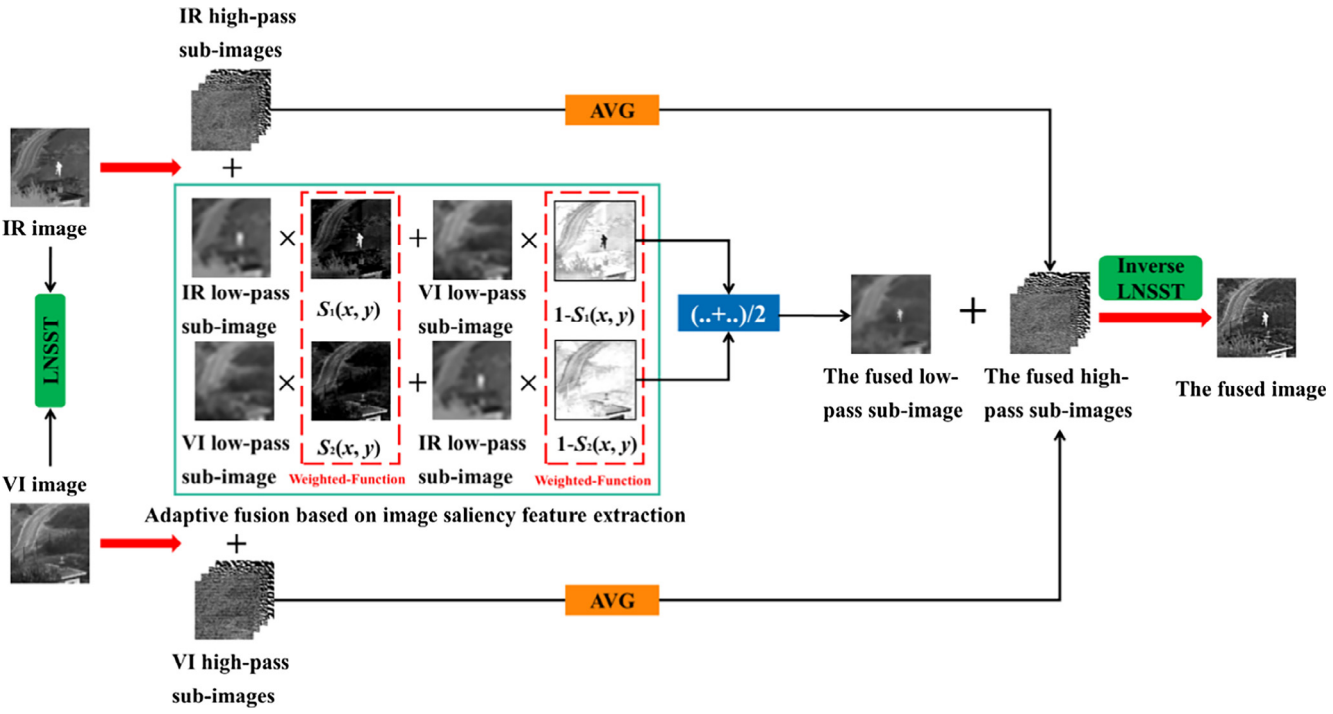


Fig. 5. Schematic of our proposed algorithm.

(a) VI image **(b)** The AVG of (a) **(c)** IR image **(d)** The AVG of (c)

Fig. 6. Images calculated by AVG.

which greatly improves the final fusion effect.

3.3. Fusion rule of high-pass sub-images

High-pass sub-images reflect the edge features and textural details of the image. The traditional "maximum absolute value" [19] is used as fusion rule, but results in significant loss of detailed information. By contrast the AVG [20] can reflect image edge details and represent changes in the gray image. Thus, the final fusion coefficients of high-pass sub-images can be replaced by AVG calculation.

Salient maps of image sources can present salient scene information. Large salient maps suggest relatively more salient information of an image source. Fusing high-pass sub-images coefficients is guided by the combination of AVG and LatLRR. The specific integration steps are listed below.

Step 1: The AVG of high-pass sub-images coefficients is calculated and expressed as follows:

$$AVG(i,j) = \frac{1}{M\times N}\sum_{i=(1-M)/2}^{(M-1)/2}\sum_{j=(1-N)/2}^{(N-1)/2}\sqrt{\frac{\nabla D_h^2(i,j)+\nabla D_v^2(i,j)+\nabla D_{md}^2(i,j)+\nabla D_{vd}^2(i,j)}{2}}, \tag{9}$$

$$\nabla D_h(i,j) = D_{l,k}(i,j) - D_{l,k}(i+1,j), \tag{10}$$

$$\nabla D_v(i,j) = D_{l,k}(i,j) - D_{l,k}(i,j+1), \tag{11}$$

$$\nabla D_{md}(i,j) = [D_{l,k}(i,j) - D_{l,k}(i+1,j+1)]/\sqrt{2}, \tag{12}$$

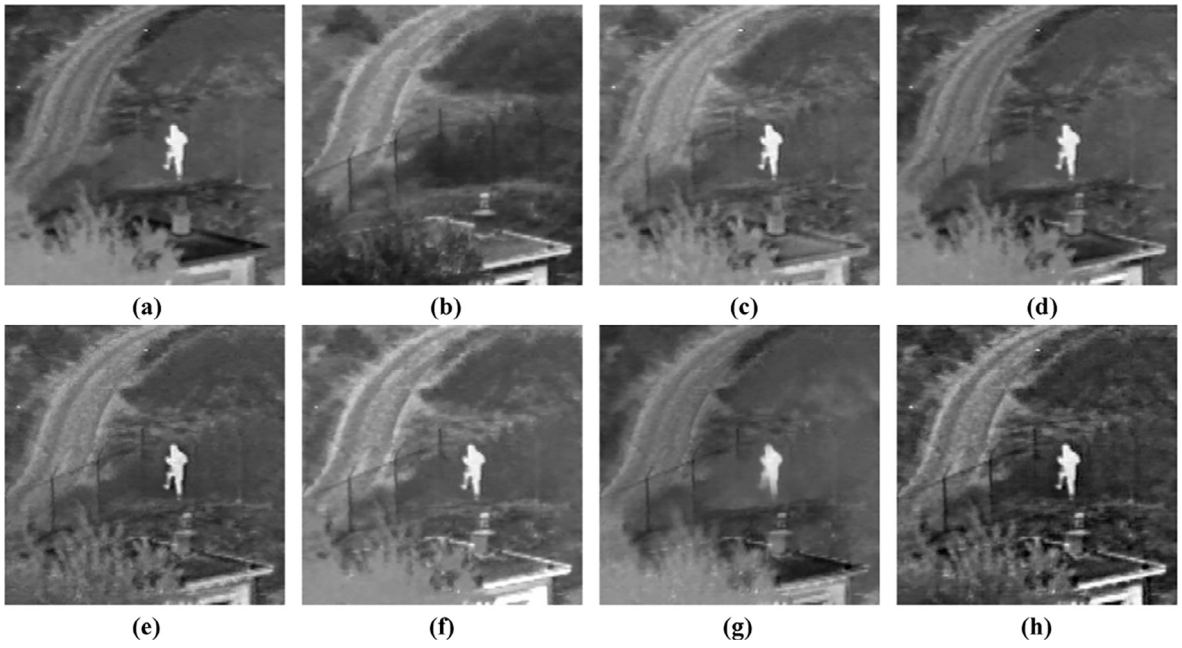
$$\nabla D_{vd}(i,j) = [D_{l,k}(i+1,j) - D_{l,k}(i,j+1)]/\sqrt{2}. \tag{13}$$

where AVG is represented by the (i, j) position; $M \times N$ is designated as 3×3; and $\nabla D_h(i, j)$, $\nabla D_v(i, j)$, $\nabla D_{md}(i, j)$, and $\nabla D_{vd}(i, j)$ represent the gradient changes in the horizontal, vertical, main diagonal, and oblique diagonal directions, respectively.

Step 2: The sizes of S_1 and S_2 and the AVG values between the coefficients of high-pass sub-images are compared separately, and then the final fusion coefficients are determined. The fusion rule is defined as

$$D_{l,k}(i,j) = \begin{cases} D_{l,k}^{IR}(i,j) & AVG_{l,k}^{IR}(i,j) > AVG_{l,k}^{VI}(i,j) \text{ and } S_1(i,j) > S_2(i,j) \\ D_{l,k}^{VI}(i,j) & others \end{cases}. \tag{14}$$

The AVG operator can reflect image edge details and represent changes in the gray image. The weighting coefficient matrices of S_1 and S_2 contain salient scene information. The information of high-pass details can be fully integrated because of the effective combination of the above matrices. The AVGs of the IR and VI image effect maps of the same scene are shown in Fig. 6. As shown in Fig. 6(b) and (d), the image edge information is extracted and enlarged by the AVG operator, thus making it more intuitive and accurate than the "absolute maximum principle," which is used to fuse the texture information between two source images.



(a) **(b)** **(c)** **(d)**

(e) **(f)** **(g)** **(h)**

Fig. 7. First group of fusion experiments.

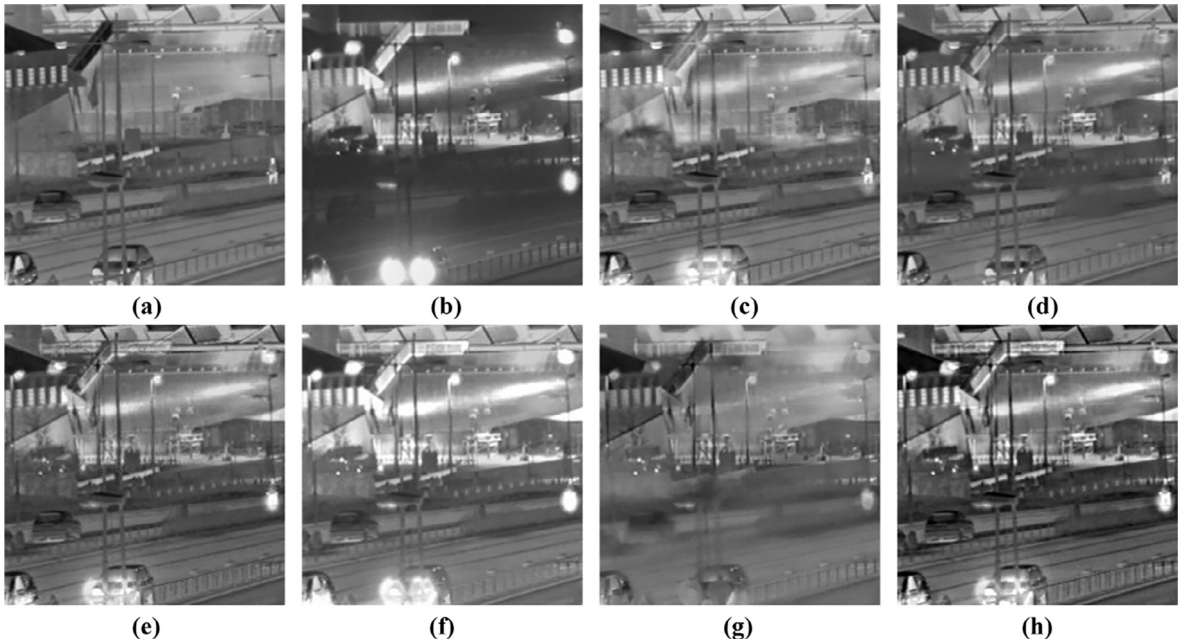


Fig. 8. Second group of fusion experiments.

4. Experimental results and analysis

Our method is compared with NSST-SF-PCNN [21], GFF [22], VSM [23], IFE [24], and GTF [25] to verify the superiority of our proposed method. Five different environments are selected to show the scope of the algorithm. The first group shows a man hidden in the forest, the second group depicts an airport at night, the third group shows a port, the fourth group shows a man walking in the forest, the fifth group depicts a ship on the sea, and the sixth group shows the coast. The original image size of the first three groups is 256 × 256, the size of the fourth group is 360 × 270, and the last two groups is 505 × 510. The parameters of the comparison algorithms are set as follows:

In the NSST-SF-PCNN method, the level of multi-scale decompositions is set to be 3, and the number of direction from coarser to finer scale is set to be 8, 16 and 16, respectively. The size of the shearing window is set to be 3. The size of the neighborhood is 3 × 3. W_{ijkl} is [0.1035, 0.1465, 0.1035; 0.1465, 0.0000, 0.1465; 0.1035, 0.1465, 0.1035]. Δ and V_θ are set to be 0.01 and 10, respectively.

In the GFF method, the default parameters are set as: $r_1 = 45$, $eps_1 = 0.3$, $r_2 = 7$, $eps_2 = 10^{-6}$.

In the VSM method, the number of decomposition levels is typically set as $N = 4$, which is good enough to produce satisfactory fusion results. The initial spatial weight is set as $\sigma^0{}_s = 2$. Generally, the values of λ in the range of [0.005, 0.02] would all obtain satisfactory results and we set λ = 0.01 in this paper.

In the IFE method, the default parameters are set as: QuadNormDim = 512; QuadMinDim = 32; GaussScale = 9; MaxRatio = 0.001; StdRatio = 0.8.

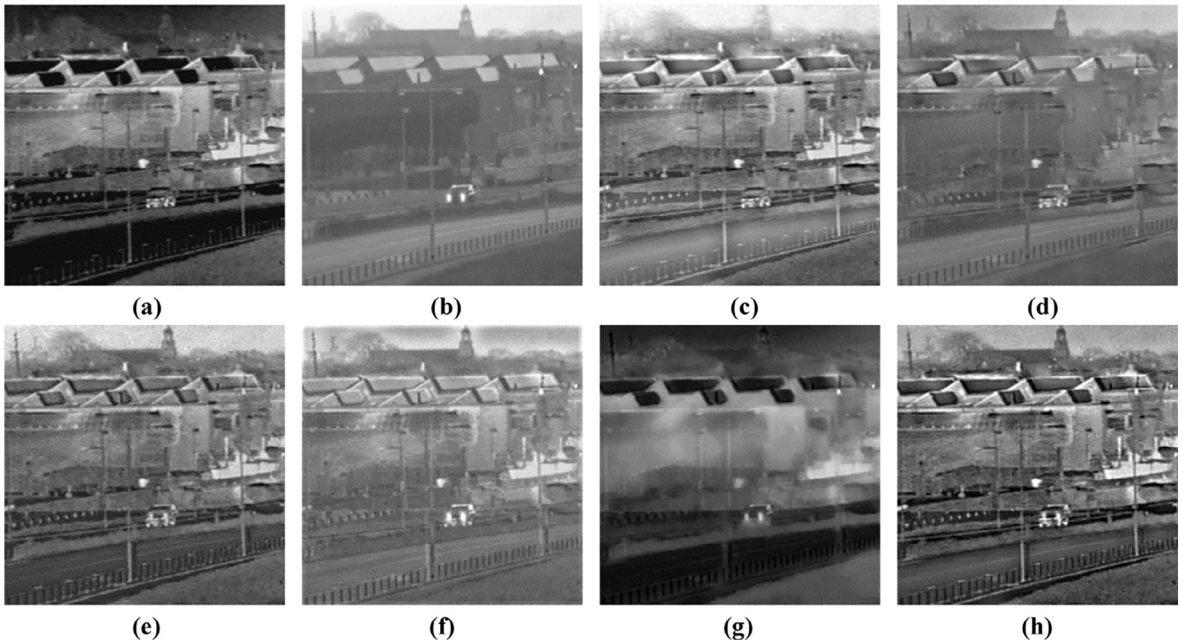


Fig. 9. Third group of fusion experiments.

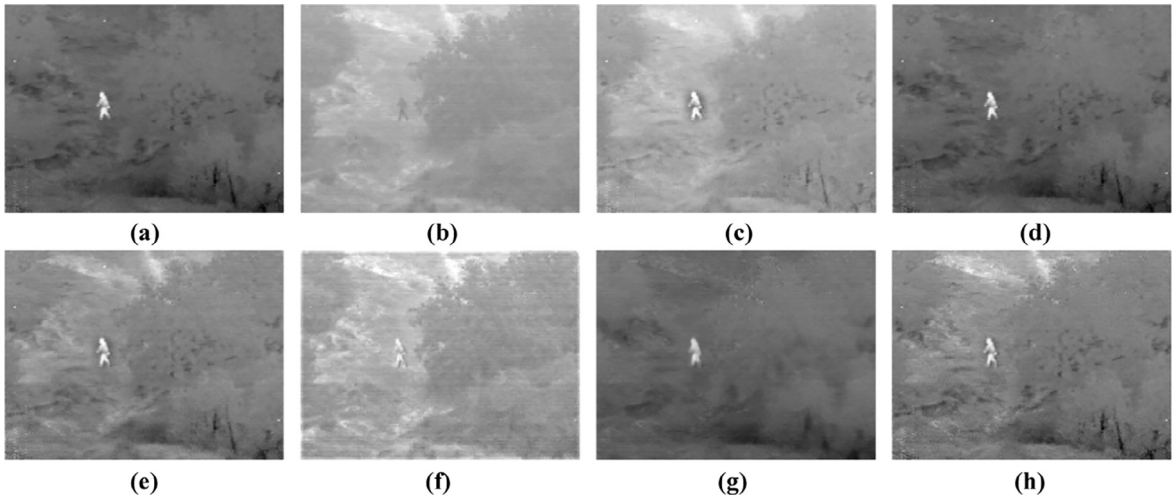


Fig. 10. Fourth group of fusion experiments.

In the GTE method, they fix λ to 4 as an empirical value, which can achieve good visual effects in most cases.

Our method takes "maxflat" and [2, 2, 3, 3] as the direction filter and the pyramid filter, respectively. For the fairness of the algorithm comparison, the parameters of each algorithm are all selected optimally, and all experiments are conducted with MATLAB 2012a programming using a PC with Intel Core i7/3.4 GHz/4G processor.

4.1. Subjective evaluation

The first to fifth group of experiments are shown from top to bottom in Figs. 7–12. The NSST-SF-PCNN, GFF, CNN, IFE-VIP, GTF and the proposed method are shown in succession from (a) to (f) in each group. In the first group of experiments, the depiction of Fig. 7(c), (d), and (g) are close to IR images, which also implies that the information of VI images is not well restored. The brightness of Fig. 7(f) is somewhat too high, image contrast is low, and textural details at the trees are not well restored. Fig. 7(e) is relatively better than the earlier images; however, compared with the algorithm in this study, the image is not clear and the reconstruction of the high-pass details is not as good as Fig. 7(f). Thus, the fusion image based on our algorithm has the best view.

In the second group of experiments, Fig. 8(c) and (g) are ambiguous with some images details lost. Fig. 8(f) still has the disadvantage of excessively high brightness, which eventually leads to unnatural image reconstruction. Fig. 8(d) does not restore the headlight information of the VI image. Therefore, information loss in the fused image is apparent. The contrast is low in Fig. 8(e), and the restoration of background details is not as good as Fig. 8(h).

In the third group of experiments, except for our algorithm, the fused images of other contrast algorithms are ambiguous, in which Fig. 9(g) is the most serious while Fig. 9(d) and (e) are dark. The background information of Fig. 9(c) is blurred, and the overall brightness of Fig. 9(f) is too high to cause important details to be missed. By contrast, Fig. 9(f) has the best contrast, sharpness, and detail reduction.

In the fourth group of experiments, Fig. 10(d) and (g) are close to IR images, and Fig. 10(c) and (f) are close to VI images. Therefore, none of the four fusion algorithms fully combines the characteristics of the two image sources. The depiction of Fig. 10(e) is close to the proposed algorithm, and textural details are as good as Fig. 10(f).

In the fifth group of experiments, the backgrounds of Fig. 11(c) and

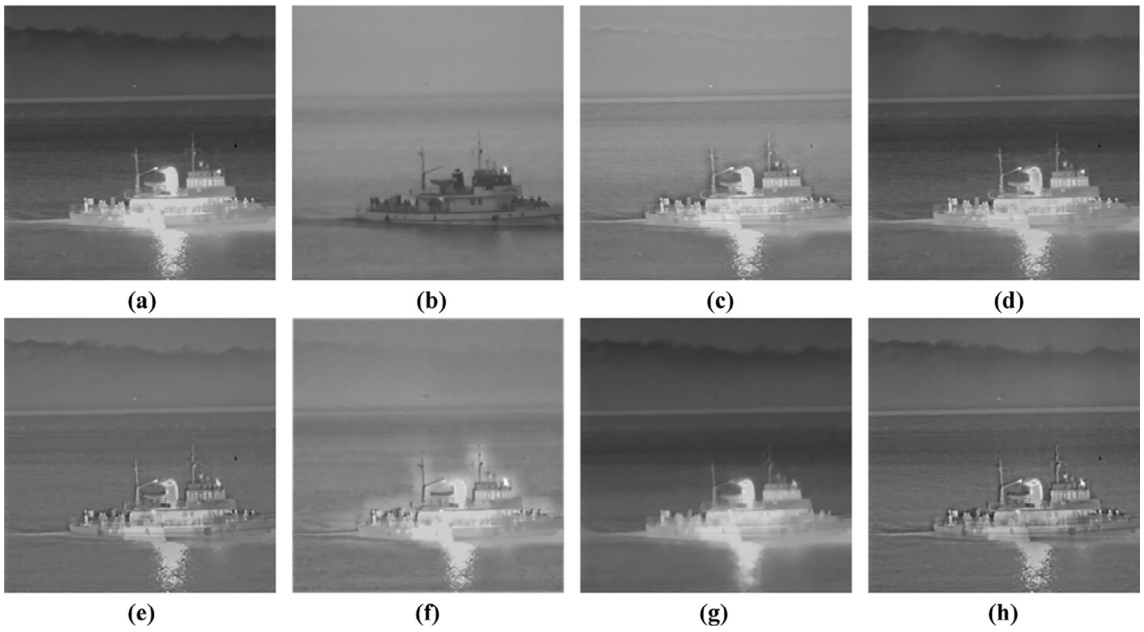


Fig. 11. Fifth group of fusion experiments.

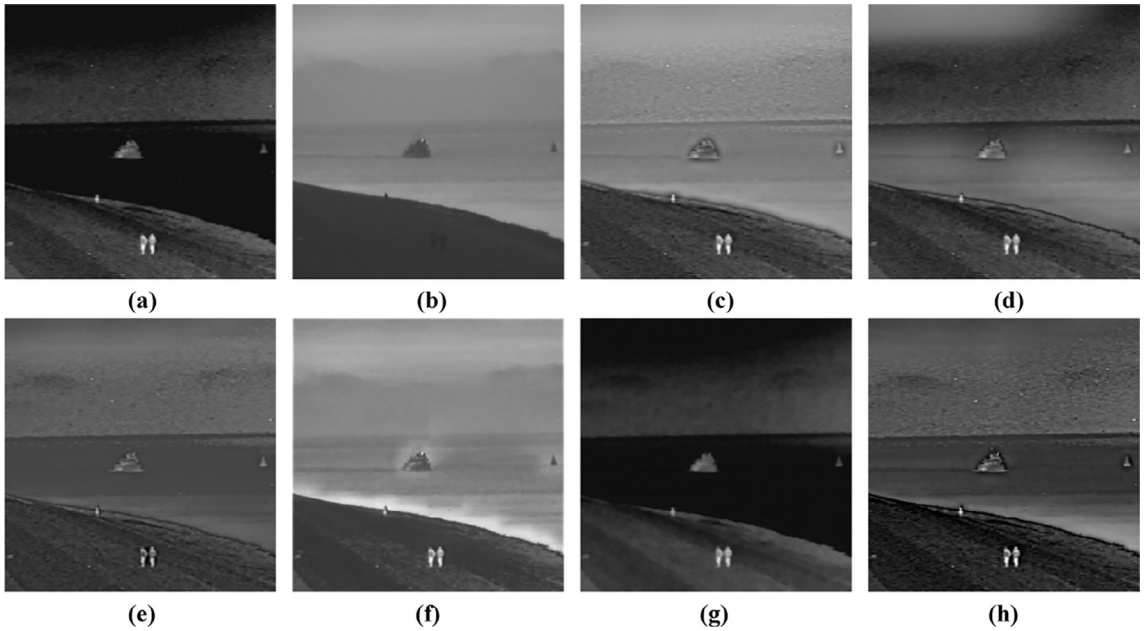


Fig. 12. Sixth group of fusion experiments.

(g) do not fuse well the information of the mountain profile in the IR image. Moreover, the hull of Fig. 11(d) and (g) are close to IR images, thus losing the edge information of VI images. The contrast of Fig. 11(e) is not as good as the algorithm in this study. In addition, the water wave texture of the sea surface is not as good as Fig. 11(h).

In the sixth group of experiments, the depiction of Fig. 12(g) is almost similar to that of the IR image. Fig. 12(d) has the disadvantage of non-uniform gray value transition in the background. Fig. 12(e) is too dark and the contrast is low. Fig. 12(f) shows a white pseudo-noise coastline. This study has the best view compared with the other algorithms, and the texture transition of water waves on the sea surface is very natural.

4.2. Objective evaluation

Results of image fusion are usually evaluated subjectively and objectively, and they present limited differences in most circumstances. However, evaluating fusion results correctly and subjectively is a difficult task. Thus, the fusion effect is frequently evaluated based on objective quality evaluations.

The following five objective quality indexes are selected as the evaluation criteria: (1) AVG [26], (2) information entropy (IE) [27], (3) edge retentiveness ($Q^{AB/F}$) [28], (4) space frequency (SF) [29], and (5) standard deviation (SD) [30]. Meanwhile, AVG is used to reflect small details of contrast and texture changes in the image. IE can be used directly to measure the richness of image information. SF incarnates the overall activity of the image in spatial domains. SD is used to reflect the pixel distribution of gray values. $Q^{AB/F}$ computes the amount of edge information, which is shifted from image sources to fused images. For all five indexes, the greater the value is, the better the performance will be.

A detailed quantitative evaluation of the six groups of IR and VI images is shown in Tables 1–3. Values in boldface represent the best results given the same index in the aforementioned method.

In the fourth group of experiments, the value of AVG of the proposed algorithm is slightly lower than that of the VSM algorithm. The main reason is that the fused image of the AVG algorithm has a higher contrast ratio than the proposed algorithm. However, the other evaluation parameters of the VSM algorithm are not as good as the proposed algorithm of this paper, which shows that the texture features and edge details of the source images are well transferred to the proposed fused image. So the proposed algorithm has more ability to restore gradient information compared with other algorithms. Apart from the findings of the fourth experiment, the AVG values of the proposed algorithm are higher than those of other algorithms in each group of experiments. Similarly, except for the third group of experiments, the IE value of the proposed algorithm is also higher than those of the other algorithms, which indicates that the information of fused images using this algorithm is rich and combines the distinctive features of IR images with the textural details of VI images altogether. In the first, second, and sixth group of experiments, the SD value of the IFE algorithm is always the highest because image brightness values based on the IFE algorithm is too high, which leads to an abnormal increase in gray values; thus, the SD is unrealistic. In the first and fifth experiments, the $Q^{AB/F}$ values of the proposed algorithm are slightly lower than that of the GFF algorithm. On the basis of objective evaluation parameters, the result is the same as that of the subjective evaluation. The proposed algorithm is superior to other algorithms in terms of image gray value distribution, edge detail, and clarity.

The last column in the Table 1 represents the running time of each method. It can be found by comparison that the NSST-PCNN method has poor timeliness and its running time is much higher than other

Table 1
Objective evaluation results of the first two groups of fused images.

Group	Methods	Evaluation index					Time/s
		AVG	IE	$Q^{AB/F}$	SF	SD	
1	NSST-PCNN	6.1628	6.7568	0.4354	11.5985	30.2343	16.95
	GFF	5.8794	6.7518	**0.5547**	11.5417	30.1531	1.21
	VSM	6.2863	6.6799	0.4316	13.7775	29.2309	1.77
	IFE	6.0695	6.7760	0.4905	12.1391	**34.4136**	0.97
	GTF	5.7834	6.5689	0.4456	11.4568	28.4232	1.56
	Proposed	**6.2934**	**6.8043**	0.5441	**13.8054**	34.0326	**2.01**
2	NSST-PCNN	8.4300	7.2840	0.4932	16.0309	40.2327	15.99
	GFF	8.1631	7.1380	0.6027	16.3177	36.7515	1.32
	VSM	9.1998	7.2895	0.5436	18.8321	41.7515	1.70
	IFE	9.1782	7.3270	0.5436	18.3845	**49.2668**	0.95
	GTF	8.1545	7.2290	0.5265	17.2381	40.4321	1.62
	Proposed	**9.2547**	**7.3347**	**0.6176**	**19.0025**	48.0779	**2.05**

Table 2
Objective evaluation results of the two middle groups of fused images.

Group	Methods	Evaluation index					Time/s
		AVG	IE	$Q^{AB/F}$	SF	SD	
3	NSST-PCNN	12.3447	7.3316	0.4451	23.3175	46.2647	17.06
	GFF	11.4227	6.7608	0.4089	20.0220	49.8611	1.26
	VSM	12.4900	**7.3771**	0.4986	25.4927	46.2628	1.65
	IFE	12.3972	6.7644	0.4223	24.7946	47.2392	0.90
	GTF	11.3987	6.7465	0.4344	23.0087	45.9325	1.60
	Proposed	**13.1121**	7.2959	**0.5061**	**26.2108**	**50.1198**	**2.11**
4	NSST-PCNN	2.7707	5.9209	0.5654	7.1446	19.8795	23.41
	GFF	2.4447	5.4152	0.6568	6.4845	20.4683	1.73
	VSM	**2.9407**	5.3915	0.5724	6.2482	19.8898	1.80
	IFE	2.4104	5.7929	0.4641	6.2482	19.8898	1.23
	GTF	2.5032	5.6892	0.5634	6.0216	18.8255	2.20
	Proposed	2.8907	**6.0250**	**0.6768**	**7.2053**	**21.2323**	**2.90**

Table 3
Objective evaluation results of the last two groups of fused images.

Group	Methods	Evaluation index					Time/s
		AVG	IE	$Q^{AB/F}$	SF	SD	
5	NSST-PCNN	2.5746	6.5777	0.6793	5.4259	25.9003	59.82
	GFF	2.4487	6.5940	**0.7191**	5.2255	27.1206	2.57
	VSM	2.3942	6.1115	0.6622	5.0313	19.8495	3.67
	IFE	2.0091	6.6387	0.6081	5.3325	29.6111	2.50
	GTF	2.0647	5.4326	0.5889	5.2424	27.7524	4.12
	Proposed	**2.7062**	**6.6203**	0.7021	**5.7212**	**31.7661**	**5.74**
6	NSST-PCNN	5.0978	6.2414	0.4732	9.0530	20.0885	60.11
	GFF	4.8765	6.1547	0.4948	8.7761	19.3435	2.60
	VSM	5.9468	6.3167	0.4981	**10.5541**	21.2330	3.60
	IFE	5.2898	6.4205	0.4446	9.3023	**23.0881**	2.61
	GTF	5.2327	6.3752	0.4216	9.5228	21.8340	4.09
	Proposed	**6.0764**	**6.5396**	**0.5384**	10.2142	22.6234	**5.82**

algorithms. In addition, the running time of the proposed algorithm is the longest among the other comparison algorithms, which shows that the complexity of the proposed algorithm is relatively high, so its timeliness needs to be improved. However, the good visual effects of a fusion algorithm often sacrifice certain timeliness, so the running time of the proposed algorithm is still acceptable.

5. Conclusion

This study presented an adaptive fusion framework for IR and VI images based on the LatLRR in LNSST domain. The LNSST was used as a multi-scale decomposition tool for the image in our proposed method. For low-pass components, adaptive-weighted fusion based on salient feature extraction using the LatLRR algorithm was adopted. For high-pass components, the AVG operator was used to process fusion coefficients. Six different scenarios were considered to verify fusion performance. The outcome showed that our algorithm could effectively fuse high-contrasting visible light and IR images while retaining considerable textural and detailed information without any artifacts. The experimental results of subjective and objective evaluations showed that our algorithm offered better fusion performance than typical fusion methods.

6. Conflict of interest

No conflict of interest.

Acknowledgments

The authors would like to thank the anonymous reviewers and editors for their invaluable suggestions. The paper is jointly supported by National High-tech R&D Program of China (No. 863-2-5-1-13B).

References

[1] Chunhui Zhao, Yunting Guo, Yulei Wang, A fast fusion scheme for infrared and visible light images in NSCT domain, Infrared Phys. Technol. 72 (2015) 266–275.
[2] Xiangzhi Bai, Morphological center operator based infrared and visible image fusion through correlation coefficient, Infrared Phys. Technol. 76 (2016) 546–554.
[3] Jufeng Zhao, Guangmang Cui, Xiaoli Gong, Yue Zang, Shuyin Tao, Daodang Wang, Fusion of visible and infrared images using global entropy and gradient constrained regularization, Infrared Phys. Technol. 81 (2017) 201–209.
[4] M.N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.
[5] A.L. Da Cunha, Zhou Jianping, M.N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE Trans. Image Process. 15 (10) (2006) 3089–3101.
[6] G. Easley, D. Labate, W. Lim, Sparse directional image representations using the discrete shearlet transform, Appl. Comput. Harmon. Anal. 25 (1) (2008) 25–46.
[7] Weiwei Kong, Yang Lei, Huaixun Zhao, Adaptive fusion method of visible light and infrared images based on non-subsampled shearlet transform and fast non-negative matrix factorization, Infrared Phys. Technol. 67 (2014) 161–172.
[8] Lu. Zhang Baohua, Pei Haiquan Xiaoqi, Zhao Ying, A fusion algorithm for infrared and visible images based on saliency analysis and non-subsampled Shearlet transform, Infrared Phys. Technol. 73 (2015) 286–297.
[9] Wu. Wei, Zongming Qiu, Min Zhao, Qiuhong Huang, Yang Lei, Visible and infrared image fusion using NSST and deep Boltzmann machine, Optik 157 (2018) 334–342.
[10] Zhanwen Liu, Yan Feng, Yifan Zhang, Xu Li, A fusion algorithm for infrared and visible images based on RDU-PCNN and ICA-bases in NSST domain, Infrared Phys. Technol. 79 (2016) 183–190.
[11] Weiwei Kong, Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform, Infrared Phys. Technol. 63 (2014) 110–118.
[12] Guangcan Liu, Shuicheng Yan, Latent low-rank representation for subspace segmentation and feature extraction, in: IEEE International Conference on Computer Vision, ICCV, 2011, pp. 1615–1622.
[13] J. Li, Y. Zhang, S. Liu, Z.J. Wang, Self-calibration method based on surface micromaching of light transceiver focal plane for optical camera, Remote Sens. 8 (11) (2016) 893.
[14] Xingbin Liu, Wenbo Mei, Du. Huiqian, Structure tensor and non-subsampled shearlet transform based algorithm for CT and MRI image fusion, Neurocomputing 235 (2017) 131–139.
[15] Wei Wua, Zongming Qiu, Min Zhao, Qiuhong Huang, Yang Lei, Visible and infrared image fusion using NSST and deep Boltzmann machine, Optik 157 (2018) 334–342.
[16] R. Achanta, S. Hemami, F. Estrada, et al., Frequency-tuned salient region detection, computer vision and pattern recognition, in: Proceedings of the IEEE Conference on CVPR 2009, 2009, pp. 1597–1604.
[17] Durga Prasad Bavirisetti, Ravindra Dhuli, Two-scale image fusion of visible and infrared images using saliency detection, Infrared Phys. Technol. 76 (2016) 52–64.
[18] Jufeng Zhao, Qiang Zhou, Yueting Chen, Huajun Feng, Xu. Zhihai, Qi Li, Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition, Infrared Phys. Technol. 56 (2013) 93–99.
[19] Fu. Zhizhong, Xue Wang, Xu. Jin, Ning Zhou, Yufei Zhao, Infrared and visible images fusion based on RPCA and NSCT, Infrared Phys. Technol. 77 (2016) 114–123.
[20] Tianzhu Xiang, Li Yan, Rongrong Gao, A fusion algorithm for infrared and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain,

Infrared Phys. Technol. 69 (2015) 53–61.
[21] Weiwei Kong, Longjun Zhang, Yang Lei, Novel fusion method for visible light and infrared images based on NSST–SF–PCNN, Infrared Phys. Technol. 65 (2014) 103–112.
[22] S. Li, X. Kang, J. Hu, Image fusion with guided filtering, IEEE Trans. Image Process. 22 (2013) 2864–2875.
[23] Jinlei Ma, Zhiqiang Zhou, Bo Wang, Hu.a. Zong, Infrared and visible image fusion based on visual saliency map and weighted least square optimization, Infrared Phys. Technol. 82 (2017) 8–17.
[24] Yu Zhang, Lijia Zhang, Xiangzhi Bai, Li Zhang, Infrared and visual image fusion through infrared feature extraction and visual information preservation, Infrared Phys. Technol. 83 (2017) 227–237.
[25] Jiayi Ma, Chen Chen, Chang Li, Jun Huang, Infrared and visible image fusion via gradient transfer and total variation minimization, Inform. Fusion 31 (2016) 100–109.
[26] Xiangzhi Bai, Infrared and visual image fusion through feature extraction by morphological sequential toggle operator, Infrared Phys. Technol. 71 (2015) 77–86.
[27] Yong Ma, Jun Chen, Chen Chen, Fan Fan, Jiayi Ma, Infrared and visible image fusion using total variation model, Neurocomputing 202 (2016) 12–19.
[28] Weiwei Kong, Yang Lei, Minmin Ren, Fusion method for infrared and visible images based on improved quantum theory model, Neurocomputing 212 (2016) 12–21.
[29] Ming Yin, Puhong Duan, Wei Liu, Xiangyu Liang, A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation, Neurocomputing 226 (2017) 182–191.
[30] Yu. Xiangzhi Bai, Fugen Zhou Zhang, Bindang Xue, Quadtree-based multi-focus image fusion using a weighted focus-measure, Inform. Fusion 22 (2015) 105–118.