

Infrared and Visible Image Fusion Based on Nonlinear Enhancement and NSST Decomposition

Xiaoxue XING (✉ xiaoxue8184@126.com)

Changchun University

Cheng LIU

Changchun University

Cong LUO

Changchun University

Tingfa XU

Beijing Institute Of Technology

Research

Keywords: Image Fusion, Non-sampled Shearlet Transform, Infrared and Visible Image, Nonlinear Enhancement, Fusion Weight

Posted Date: May 14th, 2020

DOI: <https://doi.org/10.21203/rs.2.24788/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Version of Record: A version of this preprint was published on August 24th, 2020. See the published version at <https://doi.org/10.1186/s13638-020-01774-6>.

Infrared and Visible Image Fusion Based on Nonlinear Enhancement and NSST Decomposition

Xiaoxue XING^{a*}, Cheng LIU^a, Cong LUO^a, Tingfa XU^b

(a College of Electronic Information Engineering, Changchun University, Changchun 130012, Jilin, China

b Beijing institute of technology, Beijing 100081, China

* corresponding author. E-mail address: xiaoxue8184@126.com)

Abstract-- In Multi-scale Geometric Analysis (MGA)-based fusion methods for infrared and visible images, adopting the same representation for the two types of the images will result in the non-obvious thermal radiation target in the fused image, which can hardly be distinguished from the background. To solve the problem, a novel fusion algorithm based on nonlinear enhancement and Non-Subsampled Shearlet Transform (NSST) decomposition is proposed. Firstly, NSST is used to decompose the two source images into low- and high-frequency sub-bands. Then, the Wavelet Transform (WT) is used to decompose high-frequency sub-bands into obtain approximate sub-bands and directional detail sub-bands. The “average” fusion rule is performed for fusion for approximate sub-bands. And the “max-absolute” fusion rule is performed for fusion for directional detail sub-bands. The inverse WT is used to reconstruct the high-frequency sub-bands. To highlight the thermal radiation target, we construct a non-linear transform function to determine the fusion weight of low-frequency sub-bands, and whose parameters can be further adjusted to meet different fusion requirements. Finally, the inverse NSST is used to reconstruct the fused image. The experimental results show that the proposed method can simultaneously enhance the thermal target in infrared images and preserve the texture details in visible images, and which is competitive with or even superior to the state-of-the-art fusion methods in terms of both visual and quantitative evaluations.

Keywords: Image Fusion, Non-sampled Shearlet Transform, Infrared and Visible Image, Nonlinear Enhancement, Fusion Weight

1.Introduction

Image fusion technology, which aims to combine images obtained from different sensors to create a single and rich fused image^[1], has been widely used in medical

imaging^{[2][3]}, remote sensing^{[4][5][6]}, object recognition^{[7][8]} and detection^[9]. Among the combination of different types of images, infrared and visible images fusion has attracted increasing attention^[11]. The infrared images record the thermal radiation of the scene, thus, the target in the infrared images is prominent and obvious. However, the infrared images have less detail information, low contrast, poor visual effects and poor imaging performance. In contrast, the visible images can provide abundant detail information, while the target will be inconspicuous and easily influenced by smoke, bad weather conditions and other factors. Therefore, fusion of the two types of the images can compensate for the insufficient imaging competence of infrared and visible sensors^[12]. The final fused image can possess clearer scene information as well as better target characteristics^[13].

There are seven main fusion methods: Multi-scale Geometric Analysis (MGA)-based, sparse representation-based^{[14][15][16]}, neural network-based^{[17][18]}, subspace-based^[19], saliency-based methods^[20], hybrid models^[21], and other methods. Among them, MGA-based methods are the most popular. MGA-based methods assume that the images can be represented by different coefficients in different scale. These methods decompose the source images into low- and high-component bands, combine the corresponding bands with specific fusion rules, and reconstruct the fused image with the inverse MGA transform^[22]. The key to MGA-based methods is the MGA transform, which decides the amount of the useful information that can be extracted from source images and integrated in the fused image. Popular transforms used for decomposition and reconstruction include Wavelet Transform^[23] (WT), Wedgelet Transform^[24], Curvelet Transform^{[25][26]}, Contourlet Transform^[27], NSCT^{[28][29]}, Shearlet Transform^[30] (ST), Non-Subsampled Shearlet Transform^[31] (NSST) and so on. Due to the characteristics of shift-invariant, high sensitivity, strong directivity, fast operation speed, and multi-directional processing, NSST has been widely used in the image fusion^[32]. Many researches have shown that NSST is more consistent with human visual characteristics than other MGA transforms, and the performance can make the fused images have better visual effects^[33]. However, it may be inappropriate for the infrared and visible image fusion. In infrared images, the target information is significant and easy to detect and recognize. While in visible images, the detailed information is mainly provided by gradients. Therefore, adopting the same representation for the two types of the images will cause the thermal

radiation target inconspicuous, which can hardly be distinguished from the background. In MGA-based fusion methods, it is difficult to keep the thermal radiation in infrared images and appearance information in visible images simultaneously.

To overcome the problem, we proposed a new fusion algorithm based on nonlinear enhancement and NSST decomposition for the infrared and visible images. Firstly, the NSST is used to decompose the two source images into low- and high-frequency sub-bands. Then, the high-frequency sub-bands are fused with WT-based method. To highlight the target, we construct a non-linear transform function to determine the fusion weight of low-frequency sub-bands, and whose parameters can be further adjusted to meet different fusion requirements. Finally, the inverse NSST is used to reconstruct the fused image. The experiments demonstrate that the proposed method can not only enhance the thermal target in infrared images, but also preserve the texture details in visible images. The presented method is competitive with or even superior to other methods in terms of both visual and quantitative evaluations.

The rest of this paper is organized as follows. The principle theoretical base and implementation steps of NSST are reviewed in Sec.2. The details of the proposed image fusion method are proposed in Sec.3. Experimental results and comparisons are presented in Sec.4. The main conclusion of this paper is drawn in Sec.5.

2 Related works

NSST is one of the most suitable multi-scale geometric analysis tools for fusion applications. The NSST provides an elegant sparse image representation with edges and much detail information. It does not introduce artifacts or noise when the inverse NSST is performed. In addition, the shearlet coefficients are well-localized in tight frames ranging at various locations, scales with anisotropic orientation. This achieves a successful fusion process and produces higher image quality and more clearness of image details and edges ^[34].

2.1 Basic principle of NSST

The shearlet construction is based on the non-sampled pyramid filter banks that provide the multi-scale decomposition and directional filtering generated using shear matrix that provides multi directional localization. When the dimension $n=2$, the affine system with synthetic expansion is $A_{AB}(\psi)^{[11]}$.

$$A_{AB}(\psi) = \left\{ \psi_{j,l,k}(x) = |\det A|^{\frac{j}{2}} \psi(B^l A^j x - k) : j, l \in \mathbb{Z}, k \in \mathbb{Z}^2 \right\} \quad (1)$$

Where $\psi \in L^2(\mathbb{R}^2)$, A and B are 2×2 invertible matrices and $|\det B| = 1$. If $A_{AB}(\psi)$ forms a Parseval tight framework for $L^2(\mathbb{R}^2)$, the elements of the system are called composite wavelets. For any $f \in L^2(\mathbb{R}^2)$, there is

$$\sum_{j,k,l} |\langle f, \psi_{j,l,k} \rangle|^2 = \|f\|^2 \quad (2)$$

Among them, matrix A^j and B^l are respectively associated with scale and geometric transformations, such as rotation and shear operations.

Where $A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$, $B_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$, the system can be shown as follows:

$$\left\{ \psi_{\text{ast}}(x) = a^{-\frac{3}{4}} \psi(A_a^{-1} b_s^{-1} x - t), a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2 \right\} \quad (3)$$

Equation (3) is a shearlet system, and $\psi_{\text{ast}}(x)$ is a shearlet.

Fig. 1 shows the tiling of the frequency plane induced by the shearlets and frequency supports of shearlet elements. It can be seen from Fig. 1 that each element $\hat{\psi}_{j,l,k}(x)$ is supported on a pair of trapezoidal pairs with the size of about $2^j \times 2^{2j}$, and the direction is along a straight line with a slope of $l2^{-j}$.

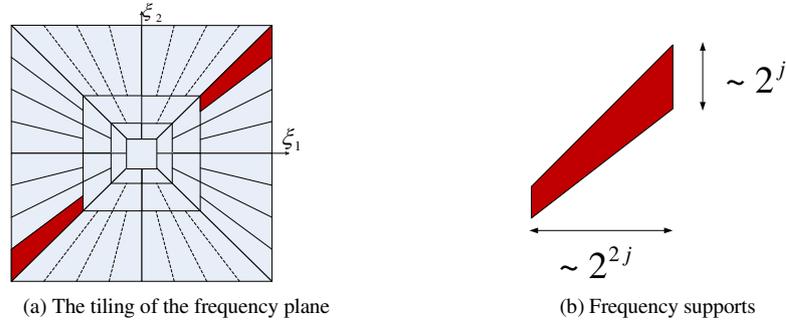


Fig 1. The tiling of the frequency plane induced by the shearlet and frequency support of shearlet elements

2.2 Implementation steps

The NSST can be realized through two steps:

(1) multi-scale decomposition. The nonsampled pyramid (NSP) filter bank decomposes each source image into a set of high- and low-frequency sub-images to attain multi-resolution decomposition. Firstly, the source image is decomposed into the low- and high-frequency coefficients with NSP. Then, the NSP decomposition of each layer will iterate on the low-frequency components obtained by the upper layer decomposition to get the singular points. Without the down-sampling operation, the

sub-band images will have the same size as the source image. Finally, for j level decomposition, we can obtain a low-pass image and j band-pass images.

(2) directional localization. The shearlet filter bank decomposes these high-frequency sub-images to attain multi-direction decomposition. Firstly, the pseudo polarization coordinates are mapped to Cartesian coordinates. Then, the ‘‘Meyer’’ wavelet is used to construct window function and generate shearlet filters. Finally, the sub-band image is convoluted with ‘‘Meyer’’ window function to obtain the direction sub-band images.

The two-level decomposition structure is shown in Fig. 2. The NSP decomposes the source image f into a low-pass filtered image f_a^1 and a high pass filtered image f_d^1 . In each iteration, the NSP decomposes the low-pass filtered image from the upper layer until the specified number of decomposition layers is reached. Finally, a low-pass low-frequency image and a series of high-frequency images are obtained.

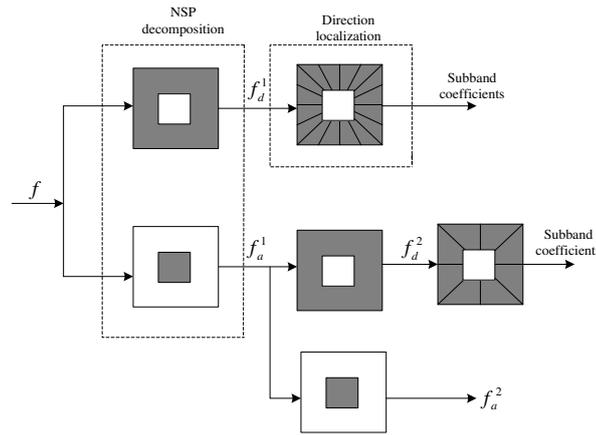


Fig.2 Two level decomposition diagram of NSST

3 Proposed method

In this section, we introduce the process of the proposed method and discuss the setting of parameters. The low- and high-frequency components obtained from the NSST decomposition represent different features information. For example, the low-frequency components carry the approximate features of the source image, and the high-frequency components carry the detailed features. The approximate parts of images provide more visually significant information and contrast information. The detailed parts of images provide more contour and edges information. Therefore, we should use different fusion rules to fuse the low-and high-frequency components. According to the stage of image data to be fused and the degree of information extraction in the fusion system, image fusion is divided into three levels: pixel level,

feature level and decision level. The proposed method focuses on the pixel level. The specific fusion scheme is shown in Fig.3. The steps of proposed method are as follows:

Step 1: Decompose the infrared and visible images with NSST into low- and high-frequency coefficients.

Step 2: Fuse low-frequency coefficients based on nonlinear enhancement algorithm.

Step 3: Fuse high-frequency coefficients based on WT-based method.

Step 4: Apply inverse NSST to obtain the fused image.

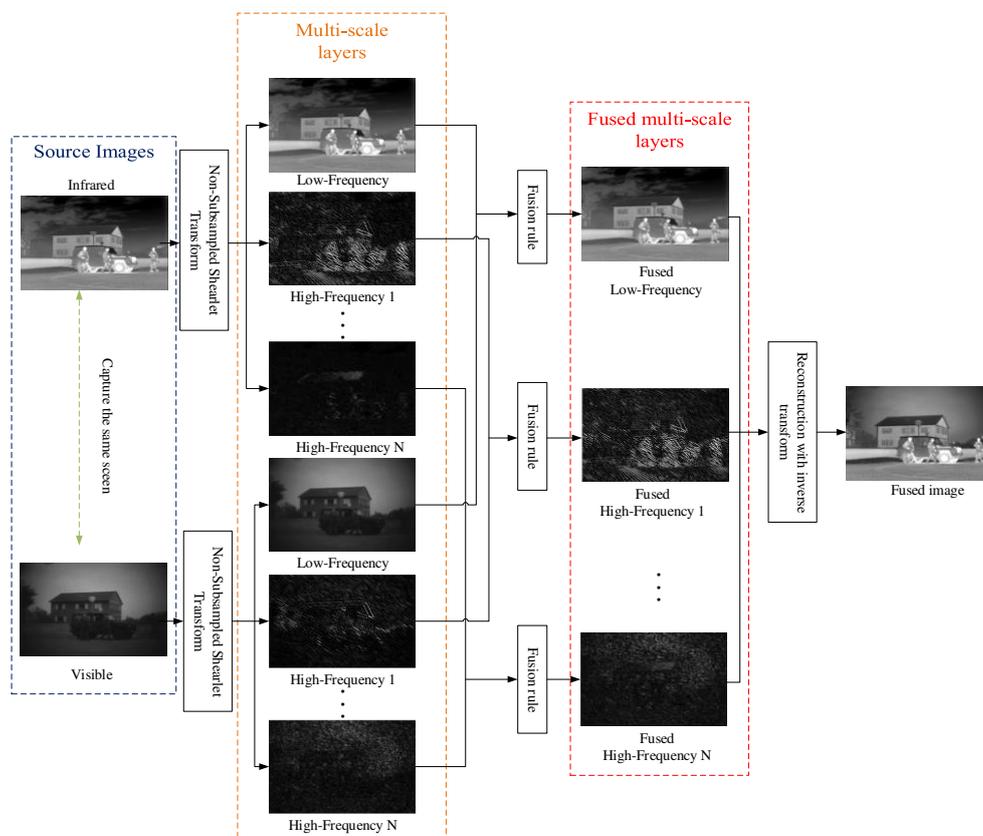


Fig.3 The diagram of fusion scheme

3.1 Low-frequency sub-band fusion

The low-frequency components reflect the contour information of the image, which contain a lot of energy information of the original image^[35]. Weighted average method is commonly used to fuse low-frequency sub-bands, however, unreasonable fusion weight will cause loss of source image information or poor image performance. We introduce a fusion strategy that construct a nonlinear transform function to determine the fusion weight of the low-frequency sub-bands to address the problems.

In infrared images, the target information is significant. Due to the large gray

values, the target is easy to detect and recognize. In order to highlight the target in the fused image, we extract the coefficients in the low-frequency component of the infrared image to determine the low-frequency fusion weight.

Each coefficient of the low-frequency components takes the absolute value as follows:

$$R = |LFC_{IR}| \quad (4)$$

Where LFC_{IR} represents the low-frequency sub-band of the infrared image after decomposition, R represents the significant infrared characteristic distribution. R_{mean} means the average of the LFC_{IR} . When R is larger than R_{mean} , it can be considered as a bright point; when R is smaller than R_{mean} , it can be considered as a dark point. The bright points are regarded as the target, while the dark points are regarded as backgrounds. In order to highlight the target, a nonlinear transform function is introduced to control the degree of the enhancement. The nonlinear transform function is as follows:

$$S(\lambda) = 1 - \frac{1}{1 + \left(\frac{R}{R_{mean}}\right)^\lambda} \quad (5)$$

Where, the parameter λ is belong to $(0, \infty)$.

The low-frequency information fusion weight can be expressed as:

$$C_{IR} = S(\lambda) \quad (6)$$

$$C_{VIS} = 1 - C_{IR} \quad (7)$$

Where, C_{IR} is the fusion weight of the infrared image, C_{VIS} is the weight of visible image, and they are both belong to $[0, 1]$.

As shown in equation 5-7, the parameter λ directly affects the fusion weight of the infrared image. Therefore, we can adjust λ to control the proportion of the infrared features of the fused image. Particularly, the larger the value of C_{IR} , the more obvious the target is. To strengthen the thermal radiation target, the value of C_{IR} should be relatively large.

The final low-frequency sub-band fusion result can be obtained as follows:

$$LFC_F = C_{IR} * LFC_{IR} + C_{VIS} * LFC_{VIS} \quad (8)$$

Where, LFC_F represents the low-frequency component of the fused image. LFC_{VIS} represents the low-frequency component decomposed by visible images.

3.2 High-frequency sub-band fusion

High-frequency components reflect details information, such as edges and

contours of the source image. To obtain more details information, we use the WT-based method to fuse the high frequency sub-bands of the infrared and visible images. Firstly, the WT is used to decompose high-frequency sub-bands to obtain approximate sub-bands (LFC_{IR} and LFC_{VIS}) and directional detail sub-bands (HFC_{IR} and HFC_{VIS}). Here, Haar wavelet is selected as the WT basis, and the decomposition layers is set to 1. Then, the “average” fusion rule is performed for fusion for approximate sub-bands. The approximate sub-bands fusion rule is defined as follows:

$$LFC_F = \frac{|LFC_{IR}| + |LFC_{VIS}|}{2} \quad (9)$$

And the “max-absolute” fusion rule is performed for fusion for directional detail sub-bands. The directional detail sub-bands fusion rule can be expressed as follows:

$$HFC_F = \begin{cases} HFC_{IR} & , |HFC_{IR}| > |HFC_{VIS}| \\ HFC_{VIS} & , \text{otherwise} \end{cases} \quad (10)$$

Where, LFC_F and HFC_F represent the approximate and directional detail sub-bands of high-frequency sub-bands images.

Finally, the inverse WT is implemented on LFC_F and HFC_F to get the high-frequency sub-bands of the fused image.

3.3 Analysis of parameter

In the nonlinear enhancement method, there is a main parameter which influence the enhancement performance, namely, the parameter λ . In this section, we draw the curve of the enhancement weight C_{IR} under different parameter λ shown in Fig. 4. The intensity of the target pixel in the fused image is determined by the value of C_{IR} . The larger the value of C_{IR} , the more evident the target is.

As shown in Fig.4, the C_{IR} curve with the the abscissa R (the gray level of the pixel) is “S” type, which shows that the target pixels can obtain larger enhancement than that of the background pixels. Moreover, the shape of C_{IR} become steep when the parameter λ increases. Therefore, it is convenient to adjust λ to get different fusion result.

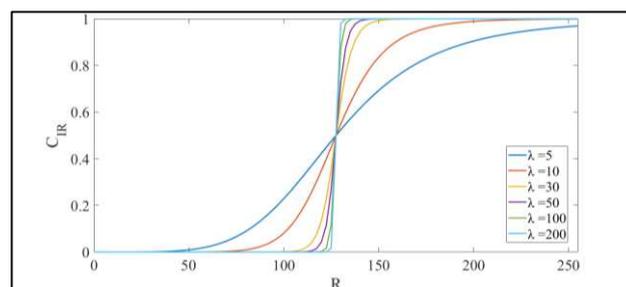


Fig 4 The curve under different parameter λ

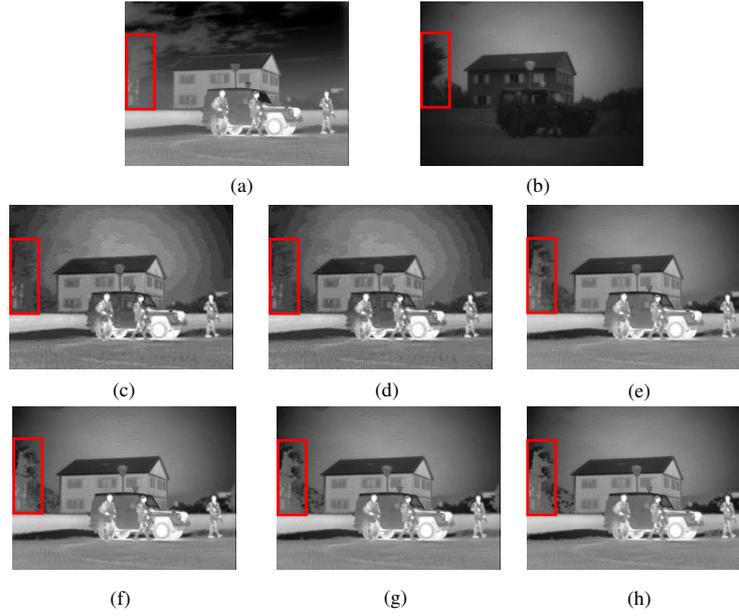


Fig. 5 Fused images under different parameter λ . (a) infrared image; (b) Visible image; (c)-(h) fused images at $\lambda = 5$, $\lambda = 10$, $\lambda = 30$, $\lambda = 50$, $\lambda = 100$, $\lambda = 200$, respectively. The differences are highlighted in red rectangles.

Fig. 5 shows the fused images under the parameter λ of 5, 10, 30, 50, 100 and 200. As seen in Fig.5, the pixel intensity distribution of infrared images is strengthened with the increase of λ . However, when λ reaches a certain degree, the distortion in the fused image will occur. The parameter λ should be appropriately large to meet different fusion requirements. In this paper, the value of λ is 10. **The proposed algorithm is summarized as Table I.**

Table I Algorithmic module

The proposed algorithmic: Based on Nonlinear Enhancement and NSST Decomposition

Input: Infrared image IR, visible images VIS, parameter λ

Output: Fused image F

1. IR and VIS are decomposed into high- and low-frequency components by NSST.
 2. The parameter λ set to 10; compute fusion weight C_{IR} and C_{VIS} by nonlinear function $S(\lambda)$ (equation 5-8).
 3. Obtain fused low-frequency component LFC_F using $LFC_F = C_{IR} * LFC_{IR} + C_{VIS} * LFC_{VIS}$.
 4. Fuse high-frequency components using WT-based method with “mean” and “maximum absolute” rules to obtain fused high-frequency component HFC_F .
 5. Repeat 4 until all corresponding components are fused.
 6. The final fused images F us obtained by inverse NSST.
-

4 Experimental results and discussion

4.1 Experimental scheme

To evaluate the performance of the proposed algorithm, two groups of simulation experiments have been carried out. Firstly, we compare the proposed method with six MGA-based methods. Then, we compare our method with other five advanced methods. Finally, qualitative and quantitative analysis of experimental results is achieved. The infrared and visible images to be fused are collected from TNO Image Fusion Dataset. Our experiments are performed using MATLAB Code on a computer with 2.6Hz Intel Core CPU and 4GB memory.

4.2 Fusion quality evaluation

4.2.1 Subjective evaluation

The subjective evaluation methods assess the quality of the fused image according to the evaluator's own experience and feeling. To some extent, it is a relatively simple, direct, fast and convenient method. However, the lower efficiency and poorer real-time performance limit its practical applications. Table.2 shows the common used subjective evaluation criteria.

Table.2 Subjective evaluation criteria

Score	Quality level	Obstruction level
5	Perfect	Do not see the deterioration of image quality
4	Good	Could see the deterioration of image quality, do not affect viewing
3	Ordinary	See clearly the deterioration of image quality, affect viewing slightly
2	Bad	Affect viewing
1	Worst	Affect viewing seriously

4.2.2 Objective evaluation

According to the different subjects, the objective evaluation indicators of image fusion quality can be divided into three categories: the characteristics of the fusion image itself, the relationship between the fusion image and the standard reference image, and the relationship between the fusion image and the source images ^[11]. We use A, B and F to infrared, visible and fused image, respectively, and R to be the ideal reference image. Here are the five objective evaluation parameters we used.

(1) Entropy (E)

E can be directly used to measure the richness of image information. The larger the E value, the better the fusion effects are. The calculation formula is shown in Equation (11):

$$E = - \sum_{i=0}^{L-1} p_i \log_2 p_i \quad (11)$$

Where, L is the total number of gray levels of the image, and p_i is the probability with the gray value i in the image.

(2) Average Gradient (AG)

AG is used to reflect the micro-details contrast and texture variation in the image. The larger the AG value, the more gradient information the fused image contains. The calculation formula is shown in equation (12):

$$\Delta \bar{G} = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N \sqrt{\frac{\Delta F_x^2(m, n) + \Delta F_y^2(m, n)}{2}} \quad (12)$$

Where, ΔF_x is the difference in the x direction of the fused image F , and ΔF_y is the difference in the y direction.

(3) Standard Deviation (SD)

SD is used to reflect the distribution of pixel gray values and the contrast of the fused image. It is defined as follows:

$$SD = \sqrt{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (F(m, n) - \bar{F}(m, n))^2} \quad (13)$$

(4) Spatial Frequency (SF)

SF is used to reflect the overall activity of the image in the spatial domain. The solution of SF is defined in Equation (16). The larger the SF , the better the fusion effects are.

$$RF = \sqrt{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N [F(m, n) - F(m, n - 1)]^2} \quad (14)$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N [F(m, n) - F(m - 1, n)]^2} \quad (15)$$

$$SF = \sqrt{RF^2 + CF^2} \quad (16)$$

Where, RF and CF are the row and column frequency of image respectively.

(5) Edge information retention ($Q^{AB/F}$)

$Q^{AB/F}$ measures the amount of edge information that is transferred from the source image to the fused image. $Q^{AB/F}$ is defined as follows:

$$Q^{AB/F} = \frac{\sum_{\forall m,n} Q_{m,n}^{AF} w_{m,n}^A + A_{m,n}^{BF} w_{m,n}^B}{\sum_{\forall m,n} w_{m,n}^A + w_{m,n}^B} \quad (17)$$

w^A and w^B denotes the weight of the importance of infrared and visible images to the fused image. Q^{AF} and Q^{BF} are calculated from the edges. A large $Q^{AB/F}$ means that considerable edge information is transferred to the fused image. For a perfect fusion result, $Q^{AB/F}$ is 1.

4.3 Experiments and results

4.3.1 Comparison with MGA-based methods

In the first group of simulation tests, we used the presented method to fuse five typical infrared and visible images in the TNO datasets, namely, “Men in front of house”, “Bunker”, “Sandpath”, “Kaptein_1123” and “barbed_wire_2”. In addition, six MGA-based methods are selected for comparison experiments, including WT [24], TEMST [36], NSST with weighted average [37], NSST with WT [38], NSCT with WT [39] and CURV with WT [40].

The key of MGA-based fusion schemes is the selection of the transforms. WT- and CURV-based methods have block artifacts, reduce the contrast of the image and can't capture abundant directional information of images. NSCT-based method can capture the geometry of image edges well, while the number of the directions at every level are fixed. In NSST-based methods, the number of the directions can be set arbitrarily, and thus the more detailed information can be obtained. But the more directions, the longer running time is. We replaced LP with NSST in TEMST as a comparative experiment.

In the proposed method, the pyramid filter for NSST is set as “maxflat”, the decomposition level of NSST is set for 3, and the number of the directions is set for {4,4,4}. The high-frequency sub-bands are decomposed into 1 level by WT (with the basis of Harr). The results are shown in Fig.6. The first two rows in Fig. 6 show the infrared and visible images. The six remaining rows denote the fused images of our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT. The subjective and objective evaluation parameters introduced

earlier are used to analyze the fusion results.

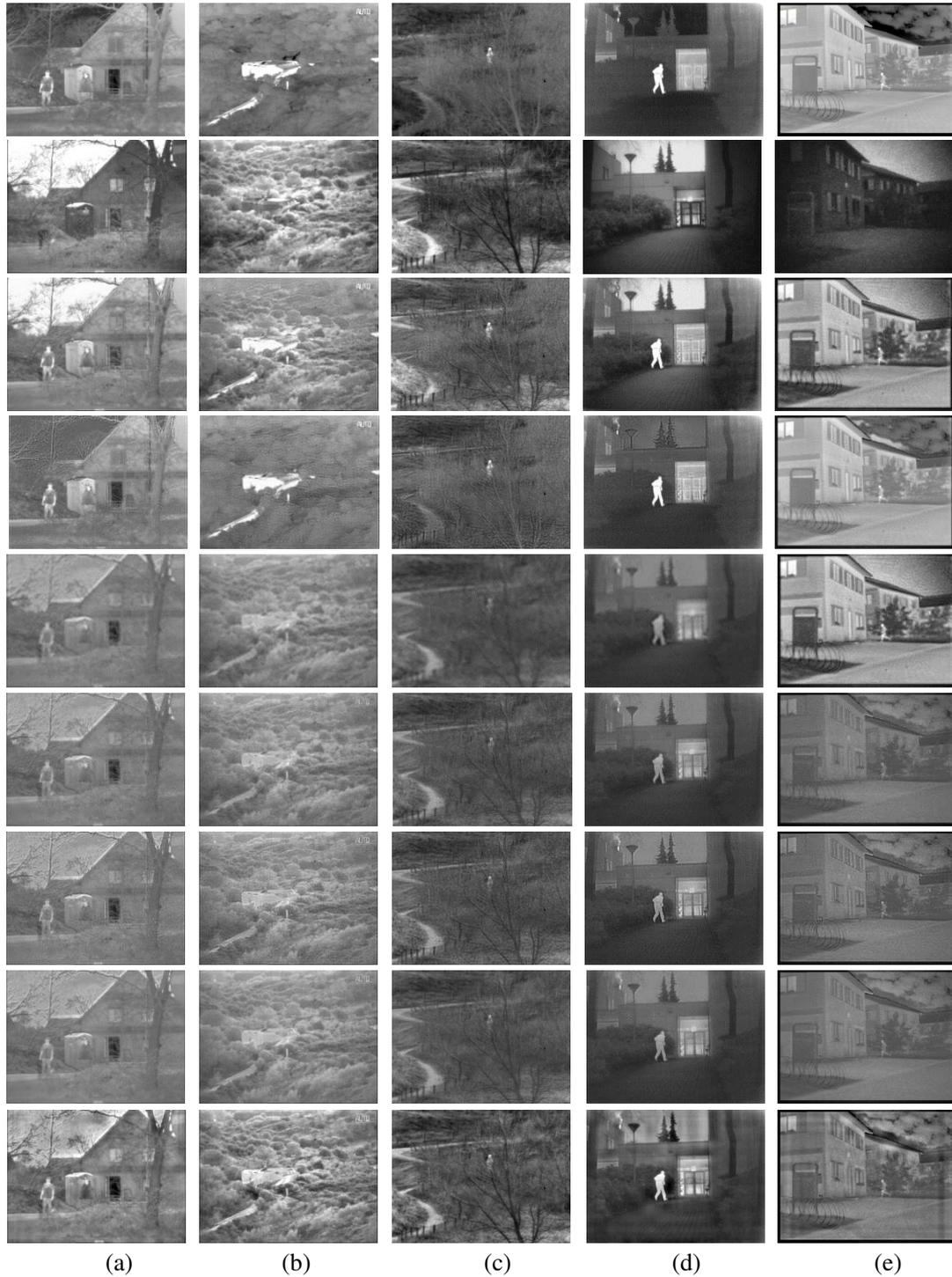


Fig.6 Fusion results on five typical infrared and visible images in the TNO datasets. (a) Men in front of house; (b) Bunker; (c) Sandpath; (d) Kaptein_1123; (e) barbed_wire_2. From top to bottom: infrared image, visible image, fused images of our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT.

The above five assessment indicators (i.e. E , AG , SD , SF and $Q^{AB/F}$) on the five typical infrared and visible images are shown in Fig.7. The larger their values, the better the fusion effects are.

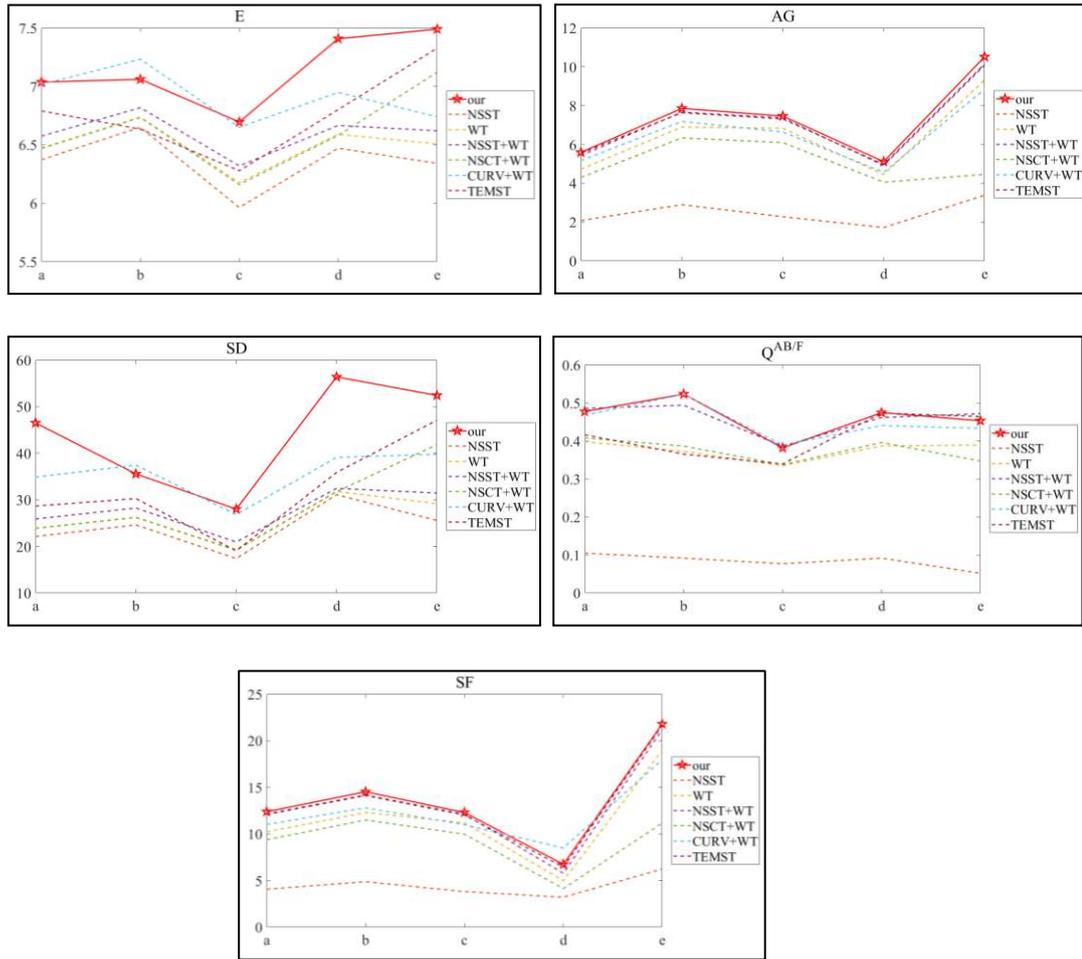


Fig.7 Comparison of five evaluation parameters, including E, AG, SD, $Q^{AB/F}$, SF. The seven methods are our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT for five pairs datasets.

4.3.2 Comparison with the state-of-the-art methods

In this part, seven typical infrared and visible images in the TNO datasets (i.e., Men in front of house, Bunker, Soldier_behind_smoke_1, Nato_camp_sequence, Kaptein_1123, Lake and Barbed_wire_1) are chosen to evaluate the effectiveness of the proposed method. We compare the proposed method with other 5 advanced methods, including: Guided filtering-based weighted average technique (GF) [41], Multi-resolution Singular Value Decomposition (MSVD)[42], Fourth Order Partial Differential Equations (FPDE) [43], Different Resolutions via Total Variation (DRTV) [44] and Visual Attention Saliency Guided Joint Sparse Representation (SGJSR)[45].

The fused images are shown in Fig. 8. The values of the five evaluations metrics on the seven infrared and visible images are shown in Fig.9.



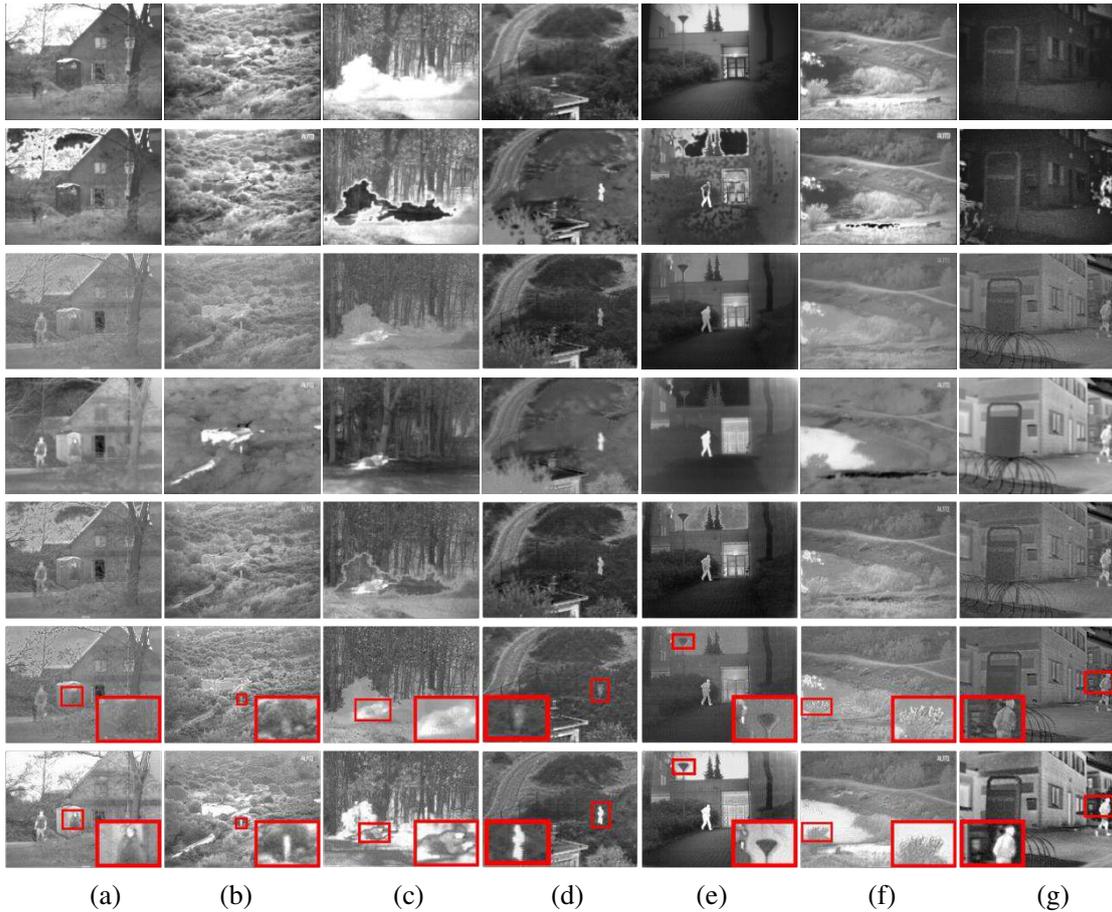
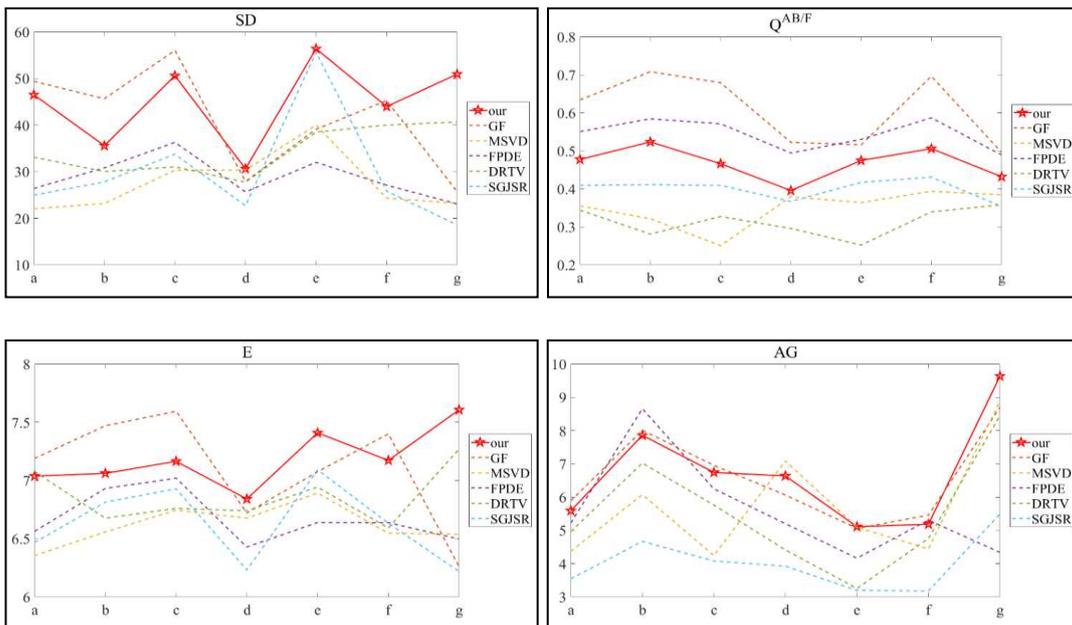


Fig.8 Qualitative fused images on seven sets of typical infrared and visible images in the TNO_datasets. (a) Men in front of house; (b) Bunker; (c) Soldier_behind_smoke_1; (d) Nato_camp_sequence; (e) Kaptein_1123; (f) Lake; (g) Barbed_wire_1. From top to bottom: infrared image, visible image, the results of GF, MSVD, DRTV, SGJSR, FPDE and our method. We pick out a small area (i.e. the red rectangle) in each fusion result, enlarge and put it on the bottom.



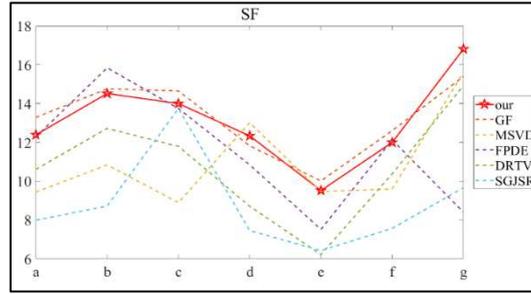


Fig.9 Comparison of five evaluation parameters, including: E, AG, SD, $Q^{AB/F}$ and SF. The six methods are GF, MSVD, DRTV, SGJSR, FPDE and our method.

4.3.3 Results and discussion

As seen in Fig.6-9, the above 12 methods can implement the effective fusion of infrared and visible images. In the other MGA-based methods, the fused image is dark and the target is not prominent, which can be clearly seen from the sky in images “Men in front of house” and “Kaptein_1123” in Fig.6. It can be seen that the proposed method can achieve apparently and easily identifiable target information. In terms of objective evaluation parameters, our proposed method is generally higher than other methods as seen in Fig.7. In short, the presented method in this paper is superior to other MGA-based methods.

Compared with five advanced methods, the presented method can achieve the best visual quality as shown in Fig.8. However, analyzing the objective evaluation parameters (i.e., E , AG, SD, SF and $Q^{AB/F}$) as seen in Fig.9, there is a fluctuation. Our method can't always get the highest values, but it can get the more stable image quality. In all, our method is competitive with the five advanced fusion methods.

5 Conclusions

In this study, we propose a new fusion algorithm for infrared and visible images based on nonlinear enhancement and NSST decomposition. It can be demonstrated that this algorithm can not only retain the texture details of the visible image, but also highlight the targets in the infrared image. Compared with other MGA-based and advanced algorithms, it is competitive or even superior in terms of qualitative and quantitative evaluation. And the fusion performance is beneficial for target detection and tracking in complex environments.

Abbreviation

MGA: Multi-scale Geometric Analysis; NSST: Non-Subsampled Shearlet Transform; WT: wavelet transform; CURV: Curvelet Transform; NSCT: Non-Subsampled Shearlet Transform; TEMST: target-enhanced multiscale transform

decomposition; GF: Guided filtering-based weighted average technique; MSVD: Multi-resolution Singular Value Decomposition; FPDE: Fourth Order Partial Differential Equations; DRTV: Different Resolutions via Total Variation; SGJSR: Visual Attention Saliency Guided Joint Sparse Representation.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work is supported in part by the Natural Science Foundation of China under NO: 61805021 and in part by the Department of Science and Technology Plan Projects of Jilin Province under NO: JJKH20191196KJ.

Authors' contributions

XX is the main writer of this paper. She proposed the main idea and construct the nonlinear function. LC (the second author) and LC (the third author) completed simulation experiment, and compared with other algorithms. XT gives some important suggestions for the simulation. All authors read and approve the final manuscript.

Acknowledgments

The research is supported by the National Natural Science Foundation of China under NO: 61805021 and the Department of Science and Technology of Jilin Province under NO: JJKH20191196KJ.

References

- [1] Liu W, Wang Z. A novel multi-focus image fusion method using multiscale shearing non-local guided averaging filter[J]. *Signal Processing*, (2020). doi:10.1016/j.sigpro.2019.107252.
- [2] Darwish S M. Multi-level fuzzy contourlet-based image fusion for medical applications [J]. *IET Image Processing*. 7(7):694-700 (2013).
- [3] Venkatrao P H, Damodar S S. HWFusion: Holoentropy and SP-Whale optimisation-based fusion model for magnetic resonance imaging multimodal image fusion[J]. *IET Image Processing*. 12(4):572-581 (2018).
- [4] Wei X. Adaptive remote sensing image fusion under the framework of data assimilation[J]. *Optical Engineering*. 50(6):067006 (2011).
- [5] Simone G, Farina A, Morabito F C, et al. Image fusion techniques for remote sensing applications[J]. *Information Fusion*. 3(1):3-15 (2002).
- [6] Li W, Hu X, Du J, et al. Adaptive remote-sensing image fusion based on dynamic gradient

- sparse and average gradient difference[J]. *International Journal of Remote Sensing*. 38(23):7316-7332 (2017).
- [7] Raghavendra R, Busch C. Novel image fusion scheme based on dependency measure for robust multispectral palmprint recognition[J]. *Pattern Recognition*. 47(6):2205-2221 (2014).
- [8] Singh R, Vatsa M, Noore A. Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition[J]. *Pattern Recognition*. 41(3):880-893 (2008).
- [9] Han J, Bhanu B. Fusion of color and infrared video for moving human detection[J]. *Pattern Recognition*. 40(6):1771-1784 (2007).
- [10] Liu Y, Chen X, Wang Z, et al. Deep learning for pixel-level image fusion: Recent advances and future prospects [J]. *Information Fusion*. 42:158-173 (2018).
- [11] Zhou Z, Dong M, Xie X, et al. Fusion of infrared and visible images for night-vision context enhancement[J]. *Applied Optics*. 55(23):6480 (2016).
- [12] Ding M, Wei L, Wang B. Research on fusion method for infrared and visible images via compressive sensing[J]. *Infrared Physics & Technology*. 57:56-67 (2013).
- [13] Gao S, Jin W, Wang L. Objective color harmony assessment for visible and infrared color fusion images of typical scenes[J]. *Optical Engineering*. 51(11):117004 (2012).
- [14] Zhang Q, Fu Y, Li H, et al. Dictionary learning method for joint sparse representation-based image fusion[J]. *Optical Engineering*. 52(5):057006 (2013).
- [15] Wang M, Mi Z, Shang J, et al. Image fusion-based video deraining using sparse representation[J]. *Electronics Letters*. 52(18):1528-1529 (2016).
- [16] Fengtao X, Zhang J, Pan L, et al. Robust image fusion with block sparse representation and online dictionary learning[J]. *IET Image Processing*. 12(3):345-353 (2018).
- [17] Kong W, Liu J. Technique for image fusion based on nonsubsampling shearlet transform and improved pulse-coupled neural network[J]. *Optical Engineering*. 52(1):017001-1-017001-12 (2013).
- [18] Wang G, Tang H, Xiao B, et al. Pixel convolutional neural network for multi-focus image fusion[J]. *Information Sciences: An International Journal*. 433/434:125-141 (2018).
- [19] Li S, Yao Z, Yi W. Frame fundamental high-resolution image fusion from inhomogeneous measurements[J]. *IEEE Transactions on Image Processing*. 21(9):4002-4015 (2012).
- [20] Bavirisetti D P, Dhuli R. Two-scale image fusion of visible and infrared images using saliency detection[J]. *Infrared Physics & Technology*. 76:52-64 (2016).
- [21] Petrusca L, Cattin P, De Luca V, et al. Hybrid ultrasound/magnetic resonance simultaneous acquisition and image fusion for motion monitoring in the upper abdomen[J]. *Investigative Radiology*. 48(5):333-340 (2013).
- [22] Kong W. Technique for gray-scale visual light and infrared image fusion based on non-subsampling shearlet transform[J]. *Infrared Physics & Technology*. 63:110-118 (2014).
- [23] Zhou Z, Tan M. Infrared image and visible image fusion based on wavelet transform[J].

- Advanced Materials Research. 756-759(2):2850-2856 (2013).
- [24] Donoho, David L. Wedgelets: nearly minimax estimation of edges[J]. *The Annals of Statistics*. 27(3):859-897 (1999).
- [25] Ali F E, El-Dokany I M, Saad A A, et al. A curvelet transform approach for the fusion of MR and CT images[J]. *Journal of Modern Optics*. 57(4):273-286 (2010).
- [26] Guo L, Dai M, Zhu M. Multifocus color image fusion based on quaternion curvelet transform[J]. *Optics Express*. 20(17):18846 (2012).
- [27] Do M N, Member, IEEE, et al. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation[J]. *IEEE Transactions on Image Processing*. 14(12):2091-2106 (2006).
- [28] Bhatnagar G, Wu Q, Liu Z. Directive contrast based multimodal medical image fusion in NSCT domain[J]. *IEEE Transactions on Multimedia*. 15(5):1014-1024 (2013).
- [29] Li Y, Sun Y, Huang X, et al. An image fusion method based on sparse representation and sum Modified-Laplacian in NSCT domain[J]. *Entropy*. 20(7):522 (2018).
- [30] Fan Z, Bi D, Gao S, et al. Adaptive enhancement for infrared image using shearlet frame[J]. *Journal of Optics*. 18(8):085706 (2016).
- [31] Padma Ganasala, Vinod Kumar. Multimodality medical image fusion based on new features in NSST domain[J]. *Biomedical Engineering Letters*. 4(4):414-424 (2015).
- [32] Kong W, Wang B, Lei Y. Technique for infrared and visible image fusion based on non-subsampled shearlet transform and spiking cortical model[J]. *Infrared Physics & Technology*. 71:87-98 (2015).
- [33] Xu L, Gao G, Feng D. Multi-focus image fusion based on non-subsampled shearlet transform[J]. *IET Image Processing*. 7(6):633-639 (2013).
- [34] Miao Q, Shi C, Xu P, et al. A novel algorithm of image fusion using shearlets[J]. *Optics Communications*. 284(6):1540-1547 (2011).
- [35] Zhang Y, Zhang L, Bai X, et al. Infrared and visual image fusion through infrared feature extraction and visual information preservation[J]. *Infrared Physics & Technology*. 83:227-237 (2017).
- [36] Chen J, Li X, Luo L, Mei G, Ma J. Infrared and visible image fusion based on target-enhanced multiscale transform decomposition[J]. *Information Sciences*. (2020). doi:10.1016/j.ins.2019.08.066.
- [37] Liu X, Mei W, Du H. Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion[J]. *Neurocomputing*. 235:131-139 (2017).
- [38] Qu Z, Xing Y, Song Y. An Image Enhancement Method Based on Non-Subsampled Shearlet Transform and Directional Information Measurement[J]. *Information*. 9(12):308 (2018).
- [39] Wu Y, Zhang H, Zhang F, et al. Fusion of Visible and Infrared Images Based on Non-Sampling Contourlet and Wavelet Transform[J]. *Applied Mechanics and Materials*. 3360(1200):1523-1526 (2014).

- [40] Bhutada G G, Anand R S, Saxena S C. Edge preserved image enhancement using adaptive fusion of images denoised by wavelet and curvelet transform[J]. Digital Signal Processing. 21(1):118-130 (2011).
- [41] Li S, Kang X, Hu J. Image fusion with Guided Filtering[J]. IEEE Transactions on Image Processing. 22(7):2864-2875 (2013).
- [42] V.P.S. Naidu. Image fusion technique using multi-resolution singular value decomposition[J]. Defence science journal. 61(5):479-484 (2011).
- [43] D. P. Bavirisetti, Xiao G, Liu G, "Multi-sensor image fusion based on fourth order partial differential equations," 2017 20th International Conference on Information Fusion (Fusion), Xi'an. pp. 1-9 (2017). doi:10.23919/ICIF.2017.8009719
- [44] Du Q, Han X, et al. Fusing infrared and visible images of different resolutions via total variation model[J]. Sensors (Basel, Switzerland). (2018). doi:10.3390/s18113827
- [45] Yang B, Li S. Visual attention guided image fusion with sparse representation[J]. Optik - International Journal for Light and Electron Optics. 125(17):4881-4888 (2014).

Figure Title and Legend Section

Fig 1. Tiling of the frequency plane induced by the shearlet and frequency support of shearlet elements

Fig.2 Two level decomposition diagram of NSST

Fig.3 The diagram of fusion scheme

Fig 4 The curve under different parameter λ

Fig. 5 Fused images under different parameter λ . (a) infrared image; (b) Visible image; (c)-(h) fused images at $\lambda = 5$, $\lambda = 10$, $\lambda = 30$, $\lambda = 50$, $\lambda = 100$, $\lambda = 200$, respectively. The differences are highlighted in red rectangles.

Fig.6 Fusion results on five typical infrared and visible images in the TNO_datasets. (a) Men in front of house; (b) Bunker; (c) Sandpath; (d) Kaptein_1123; (e) barbed_wire_2. From top to bottom: infrared image, visible image, fused images of our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT

Fig.7 Comparison of five evaluation parameters, including E, AG, SD, $Q^{AB/F}$ and SF. The seven methods are our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT for five pairs datasets.

Fig.8 Qualitative fused images on seven sets of typical infrared and visible images in the TNO_datasets. (a) Men in front of house; (b) Bunker; (c) Soldier_behind_smoke_1; (d) Nato_camp_sequence; (e) Kaptein_1123; (f) Lake; (g) Barbed_wire_1. From top to bottom: infrared image, visible image, the results of GF,

MSVD, DRTV, SGJSR, FPDE and our method. We pick out a small area (i.e. the red rectangle) in each fusion result, enlarge and put it on the bottom.

Fig.9 Comparison of five evaluation parameters, including: E, AG, SD, $Q^{AB/F}$ and SF. The six methods are GF, MSVD, DRTV, SGJSR, FPDE and our method.

Figures

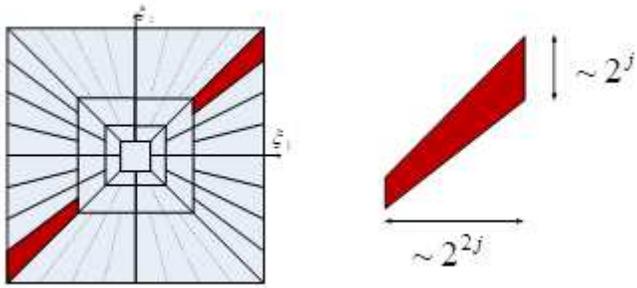


Figure 1

The tiling of the frequency plane induced by the shearlet and frequency support of shearlet elements. (a) The tiling of the frequency plane. (b) Frequency supports.

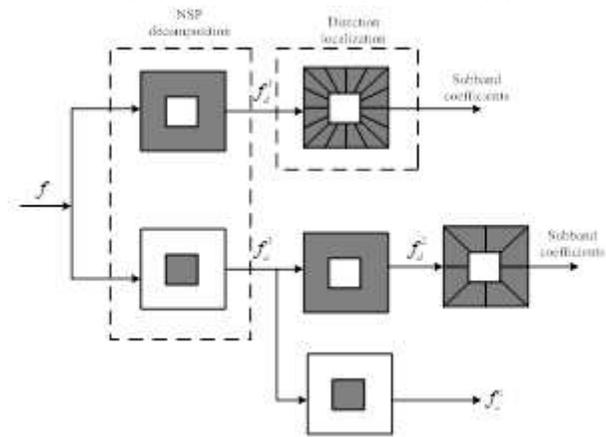


Figure 2

Two level decomposition diagram of NSST

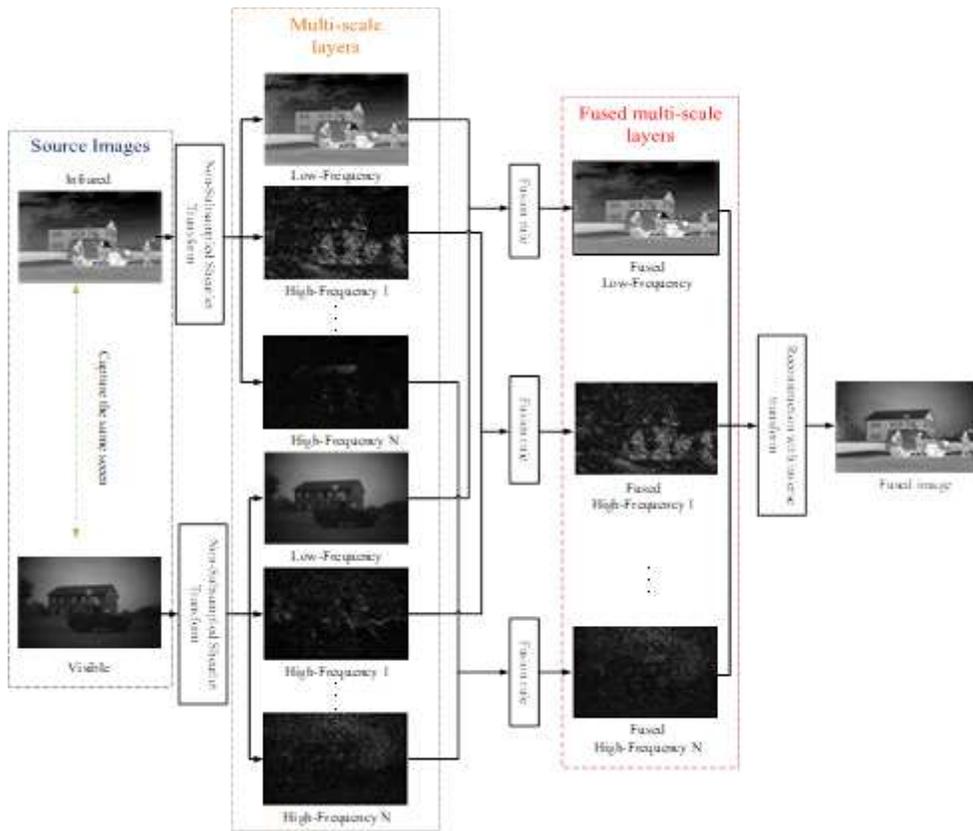


Figure 3

The diagram of fusion scheme

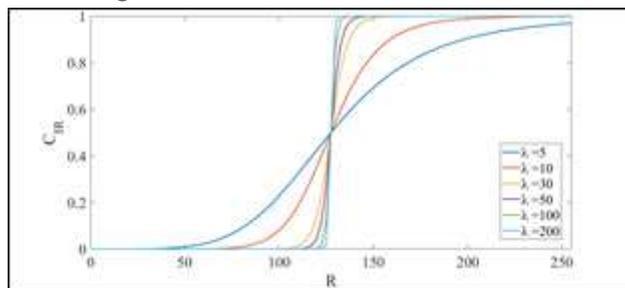


Figure 4

The curve under different parameter λ

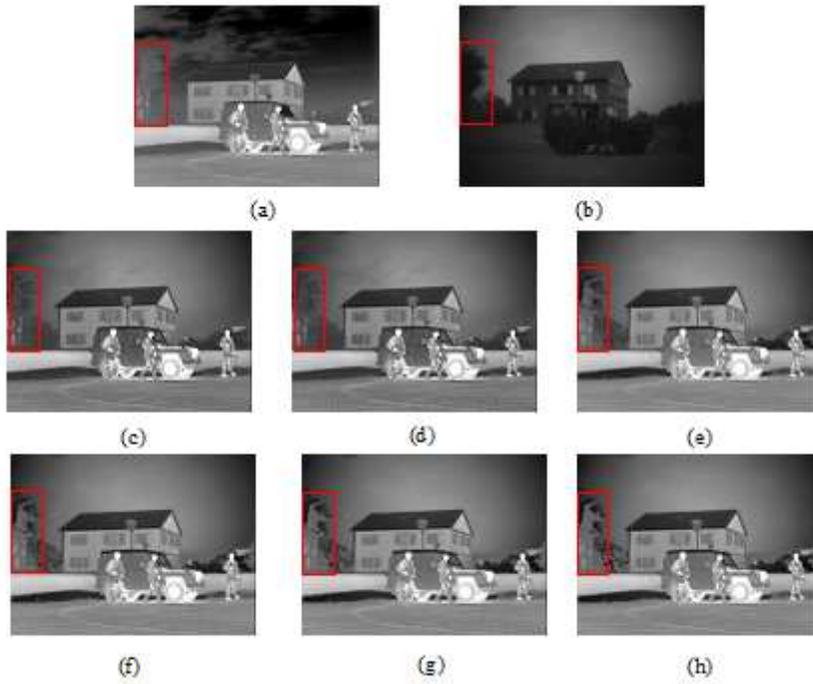


Figure 5

Fused images under different parameter λ . (a) infrared image; (b) Visible image; (c)-(h) fused images at $\lambda = 5, \lambda = 10, \lambda = 30, \lambda = 50, \lambda = 100, \lambda = 200$, respectively. The differences are highlighted in red rectangles.

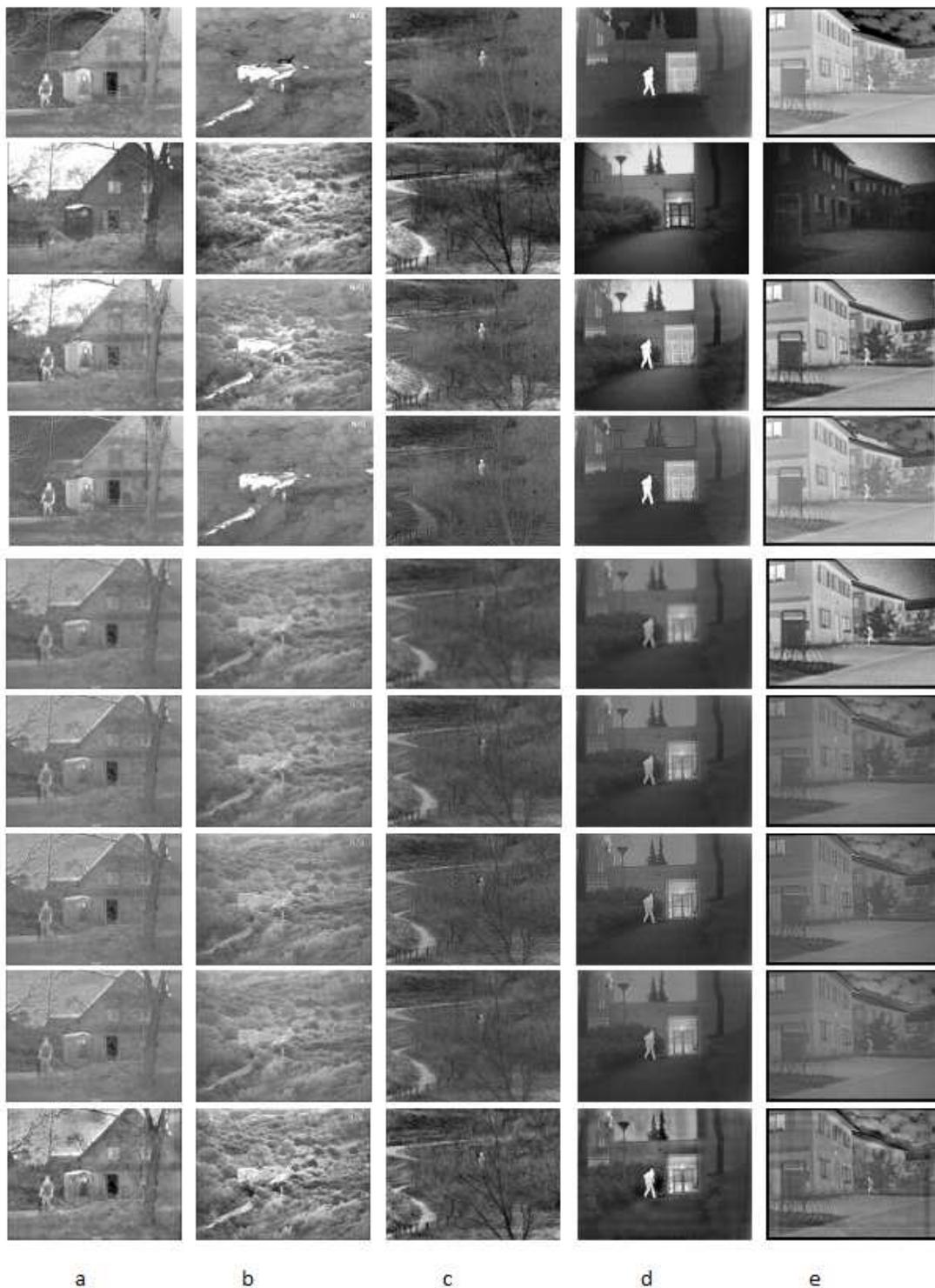


Figure 6

Fusion results on five typical infrared and visible images in the TNO datasets. (a) Men in front of house; (b) Bunker; (c) Sandpath; (d) Kaptein_1123; (e) barbed_wire_2. From top to bottom: infrared image, visible image, fused images of our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT.

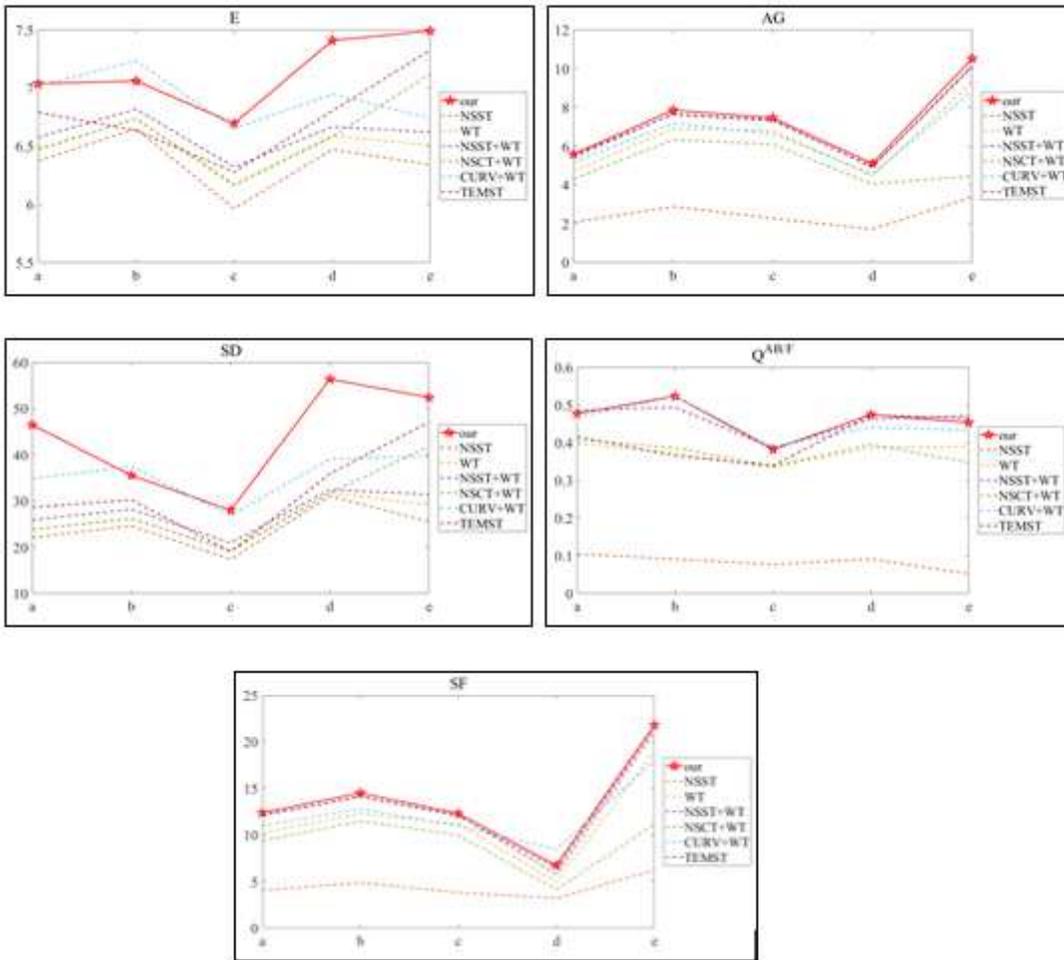


Figure 7

Comparison of five evaluation parameters, including E, AG, SD, QAB/F, SF. The seven methods are our method, TEMST, NSST with weighted average, WT, NSST with WT, NSCT with WT, CURV with WT for five pairs datasets.

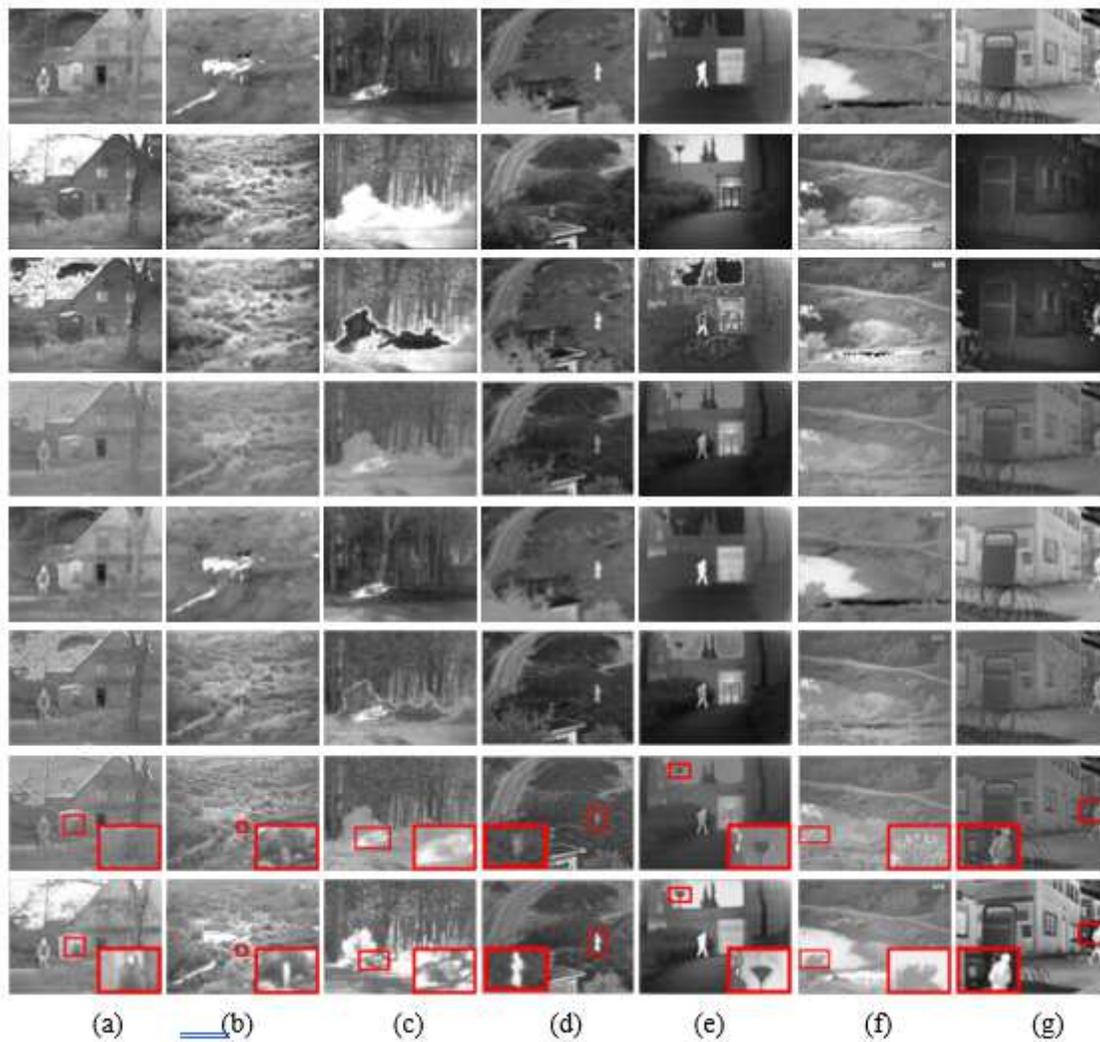


Figure 8

Qualitative fused images on seven sets of typical infrared and visible images in the TNO_datasets. (a) Men in front of house; (b) Bunker; (c) Soldier_behind_smoke_1; (d) Nato_camp_sequence; (e) Kaptein_1123; (f) Lake; (g) Barbed_wire_1. From top to bottom: infrared image, visible image, the results of GF, MSVD, DRTV, SGJSR, FPDE and our method. We pick out a small area (i.e. the red rectangle) in each fusion result, enlarge and put it on the bottom.

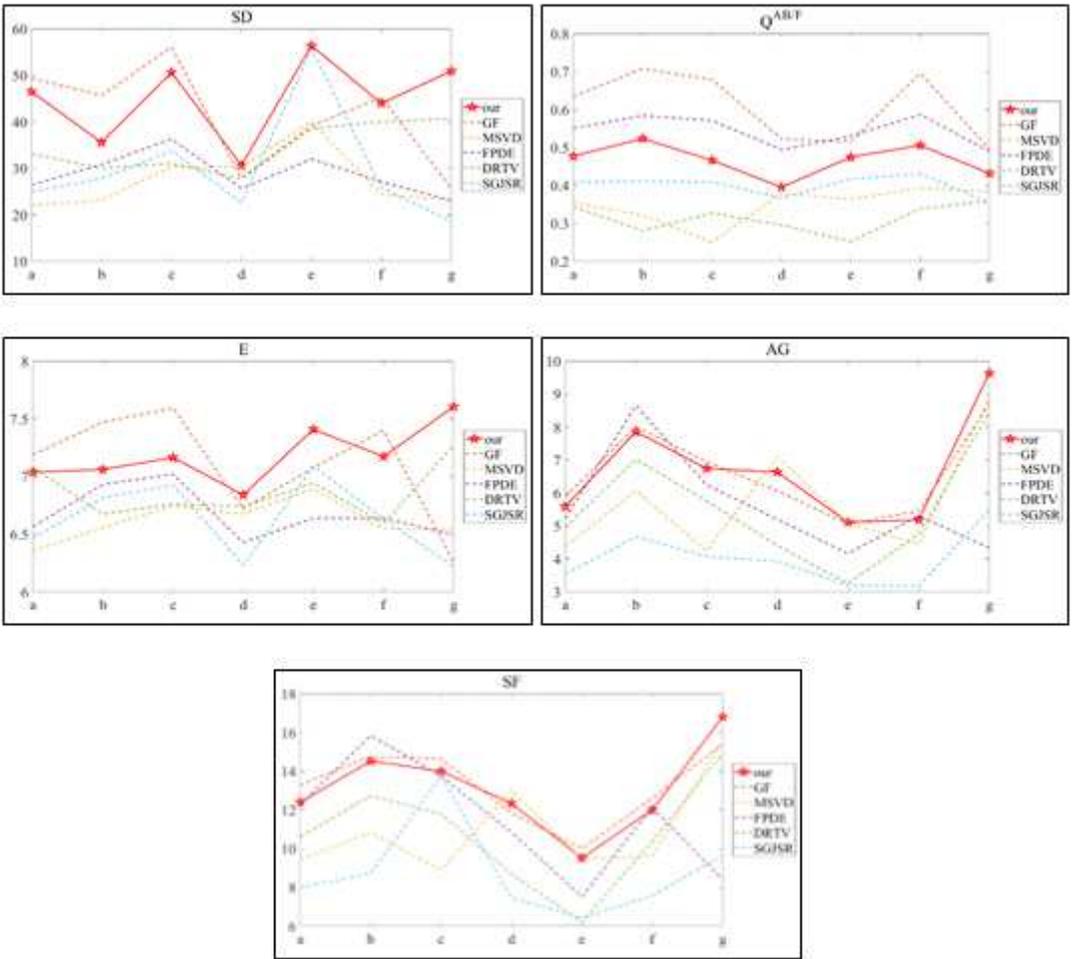


Figure 9

Comparison of five evaluation parameters, including: E, AG, SD, QAB/F and SF. The six methods are GF, MSVD, DRTV, SGJSR, FPDE and our method.