



Infrared and visual image fusion using LNSST and an adaptive dual-channel PCNN with triple-linking strength

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

Article history:

Received 2 November 2017

Revised 27 April 2018

Accepted 7 May 2018

Available online 22 May 2018

Communicated by Jungong Han

Keywords:

LNSST

ATD-PCNN

Image fusion

Singular value decomposition

Auxiliary linking strength

Triple-linking strength

ABSTRACT

This paper presents a novel fusion framework for infrared and visual image that uses an adaptive dual-channel pulse-coupled neural network (PCNN) with triple-linking strength (ATD-PCNN) in a local non-subsampled shearlet transform (LNSST) domain. First, the LNSST, which is an upgrade based on the NSST, is used as a multi-scale analysis tool to decompose the source images into low-pass and high-pass sub-images. Second, an improved sum-modified Laplacian (ISML) that reflects the energy characteristics of low-pass sub-images in the LNSST domain is used to stimulate the ATD-PCNN model, and an improved average gradient (IAGV) that reflects the textural details of high-pass sub-images is used to stimulate the ATD-PCNN neurons. Furthermore, to solve the problem of the spectral difference between infrared and visible light, three novel operators are used as the adaptive linking strength, and an optimization model based on l_2 -norm is used for merging the output coefficients. A series of images from diverse scenes are used for the fusion experiments, and the fusion results are evaluated subjectively and objectively. The results of the subjective and objective evaluations show that our algorithm exhibits superior fusion performance and is more effective than typical fusion techniques.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The concept of image fusion was proposed in the early 1970s, and it primarily involves the study of the intrinsic relationship between multi-source images via the use of appropriate algorithms to convert two or more pieces of redundant, diverse and complementary information into one fused image with richer information [1]. In general, the source images of the fusion are classified into two types: images acquired by the same type of sensor or images acquired by different types of sensors. Different types of sensors have different properties in principle, and the images obtained by different types of sensors have a greater difference and complementarity. Therefore, research on heterogeneous image fusion technology is important [2]. Infrared (IR) and visible light (VI) image fusion is the most widely used heterogeneous image fusion technique. An effective combination of such techniques can combine the advantages of the two types of images to form a fusion image with high contrast, background enhancement and target projection, which is more convenient for accurate detection of the target [3].

Researchers have proposed a large variety of fusion algorithms, and multi-scale geometric analysis (MGA) is currently the most widely used [4,5]. In this method, the source image is decomposed into a series of multi-direction and multi-scale sub-images and then appropriate fusion rules are applied to transform the corresponding sub-band images into fused images. Certain traditional fusion rules generally use the weighted-average method, which ensures the complementarity of the brightness of the IR and VI images and eventually leads to image contrast reduction and unnatural image reconstruction [6]. A pulse-coupled neural network (PCNN) [7] is obtained by studying neuron stimulation in the mammalian visual system, and it conforms to the visual information system of the human brain. So as a fusion rule, combining MGA with it, the spatial 2D information integrity of the input image is retained after the fusing process. Using the perfect combination of the PCNN model and visual characteristics, a better fusion effect is obtained. It is worth mentioning that, NSST [8] is the most advanced method used today among these numerous MGA methods, and it is constructed from synthetic wavelets that can obtain a multi-directional tight support structure. In addition, a sub-sampled operation is not included in the decomposition process; thus, the image can be finely decomposed. Because the large-sized shearlet filter will cause a spectral aliasing phenomenon in the directional localization stage of the NSST, a local non-subsampled

* Corresponding author at: Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China.

E-mail address: chengboyang15@mails.ucas.edu.cn (B. Cheng).

shearlet transform (LNSST) [9] is proposed as a multi-scale decomposition tool of the image based on the NSST. This method uses the local small-size shearlet filter to avoid the block effect and weaken the Gibbs-ringing phenomenon, which improves the time domain convolution calculation efficiency. Therefore, it is very valuable to study the fusion algorithm based on LNSST with PCNN model.

Xiang et al. [10] proposed an adaptive dual-channel PCNN model based on the traditional PCNN. This new model can reflect the information from two source images at the same time, and simplify a large number of peripheral parameters. Moreover, the "linking strength", which is the most important parameter of PCNN model is adaptively set. These changes significantly increase the accuracy of the fusion, so they provide some ideas for the study of this paper [11]. Although a number of results have been achieved, problems are still encountered, such as target halo and black pseudo-noise, and the essential reason is that the selection of the linking strength in the model is inappropriate. Aiming at the fusion of images with large difference spectra between IR and VI, a single operator such as average gradient [12], spatial frequency [13] and edge gradient [14], which is treated as the linking strength, will lead to the decrease of the stability of the value of linking strength, so it eventually leads to discontinuity and edge blur in the fused images. The linking strength reflects the strength of the human visual system response to different characteristic regions in the image, thus, a better operator that can reflect the image characteristics is required as the external linking strength [15]. In addition, low-pass sub-images are different from high pass sub-images, and the same operator should not be used as the linking strength for both types. The low-pass sub-image often neglects the edge detail information, and the high-pass sub-image often ignores the main energy characteristics. Thus, a type of auxiliary linking strength is required to connect the information in the low-pass sub-image and high-pass sub-image and compensate for their deficiencies in linking strength. In addition, the low-frequency and high-frequency components of the image produced after LNSST decomposition are different in nature, and the same operator should not be used as the linking strength for both types. The low-pass sub-image often neglects the edge detail information, and the high-pass sub-image often ignores the main energy characteristics. Thus, a type of auxiliary linking strength is required to connect the information in the low-pass sub-image and high-pass sub-image and compensate for their deficiencies in linking strength. Based on the above analysis, the existing adaptive dual-channel PCNN model needs to find three new operators that can better reflect the intrinsic eigenvalue of the image to solve the IR and VI image fusion with large spectral differences.

The singular value of an image [16] contains the structural information of the image and thus concentrates the energy information and the regional characteristics of the image. Therefore, a local structure information factor (LSI) that is constructed of the local region singular value of the image is regarded as the main linking strength of the low-pass sub-image, and a sum of directional gradients (SDG) represents the image information changed in the horizontal, vertical and diagonal directions, which can be used as the main linking strength of the high-pass sub-image to characterize the texture details of the image. A Laplacian edge energy (LEE) operator can reflect the edge energy information of low-pass sub-images and highlight changes in the edges of details of high-pass sub-images; therefore, it can be used as a bridge to connect the information in the low-frequency and high-frequency domains and compensate for deficiencies in the original linking strength. Through the interaction of the above three operators, the feature information of the image can be more fully input into the PCNN model, and the stability of the value of linking strength can be consolidated so that the final fusion image is naturally transitioned. To

our best knowledge, this is a new adaptive PCNN model, which is called ATD-PCNN.

Based on the preceding review, due to the fact that the current adaptive linking strength operator cannot solve the spectral difference between IR and VI, and single linking strength cannot express the characteristics of the low-pass sub-images and high-pass sub-images simultaneously, our method presents a novel fusion framework of VI and IR images by using an adaptive dual-channel PCNN model with triple-linking strength in the LNSST domain. First, the LNSST is used as a multi-scale analysis tool to decompose the source images into a low-pass sub-image and high-pass sub-image. Second, the concrete selection principles of the low-pass sub-image and high-pass sub-image are separately discussed in detail in this paper. Furthermore, three new operators are proposed based on the original adaptive dual-channel PCNN model as the external linking strength to solve the problem of the spectral difference between IR and VI, and an optimization model based on l_2 -norm is used for merging the output coefficients. Finally, each sub-image is modeled, and the corresponding fusion coefficients are produced. The algorithm can effectively express the characteristics of the image and obtain a better fusion effect by using IR and VI images in the fusion experiments.

The remainder of this paper is organized as follows. Chapter 2 details the principle of the LNSST. Chapter 3 introduces the singular value decomposition of the image and its characteristics. Chapter 4 introduces the theoretical framework of the adaptive dual-channel PCNN. Chapter 5 elaborates the algorithm based on the new fusion rule. Chapter 6 presents six experimental results and intuitively compares the proposed method with other methods. Chapter 7 provides a summary of the findings.

2. Local non-subsampled shearlet transform [17,18]

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

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

where $\varphi \in L^2(\mathbb{R}^2)$; 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 $j \geq 0$, $-2^j \leq l \leq 2^j - 1$, $k \in \mathbb{Z}^2$, $d = 0, 1$, the Fourier transform of the shearlet can be expressed according to the tight support frame:

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

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 function $f \in L^2(\mathbb{R}^2)$ can be calculated by Eq. (3):

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

Eq. (3) shows that the shearlet transform is divided into two steps: the first step is a multi-scale decomposition, i.e., $\hat{f}(\xi) V(2^{-2j}\xi)$; and the second step is the direction of localization, i.e., $\hat{f}(\xi) V(2^{-2j}\xi) W_{j,l}^{(d)}(\xi)$.

Multi-scale decomposition: the image is subjected to non-subsampled pyramid decomposition (NSP) using a non-subsampled 2D 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 (SFs) and high-pass sub-images convolution calculations. The local window is $L \times L$, where $L = n \cdot (2^{j-1} + 1)$, j is the scale parameter, and n is any positive integer. Because the local small-size shearlet filter can avoid the block effect, weaken the

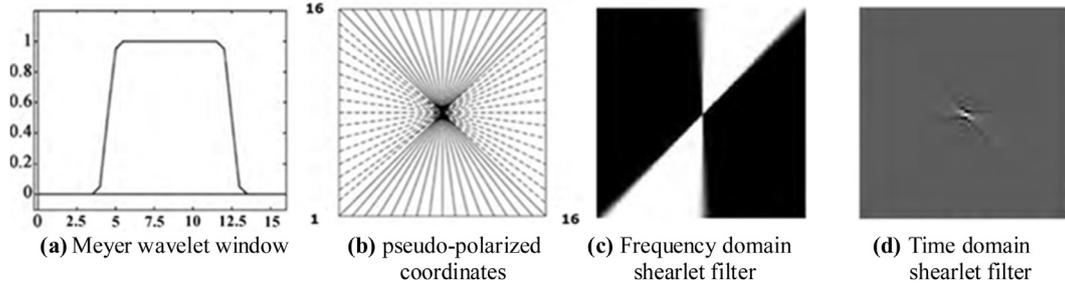


Fig. 1. Shearlet filter formation procedure when $L = 15$.

Gibbs-ringing phenomenon, and improve the time domain convolution calculation efficiency, j is usually 2 or 3 and the local window is usually 15×15 . The Meyer window function $g(\theta)$ is generated in the window using the Meyer wavelet function $v(x)$.

The Meyer wavelet function is defined as follows:

$$v(x) = \begin{cases} 0 & x < 0 \\ 35x^4 - 84x^5 + 70x^6 - 20x^7 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases} \quad (4)$$

The Meyer window function is as follows:

$$g(\theta) = \begin{cases} 1 & |x - 1/2| < 1/3 \\ \cos^2 \frac{\pi}{2} (\nu(3|x - 1/2| - 1)) & 1/3 \leq |x - 1/2| \leq 2/3 \\ 0 & \text{other} \end{cases} \quad (5)$$

where $g(\theta)$ is subjected to discrete resampling in the pseudo-polarized network and is then changed from pseudo-polarized coordinates to the Cartesian coordinate system, $\hat{W}_{j,l}^s$ forms a frequency domain implemented shearlet filter with a size of $L \times L$, and $\sum_{l=-2^j}^{2^j-1} \hat{W}_{j,l}^s(\xi_1, \xi_2) = 1$. $\hat{W}_{j,l}^s$ is represented as the time domain-implemented shearlet filter. For any image f , we have the following:

$$\sum_{l=-2^j}^{2^j-1} f \times W_{j,l}^s = f, \quad j \geq 0. \quad (6)$$

When $L = 15$, the formation process of the shearlet filter is expressed as shown in Fig. 1.

The above shearlet transformation is called the local non-subsampled shearlet transform. The local non-subsampled shearlet removes the sampling operation in the decomposition stage, which has translation invariance, because the local small-size shearlet filter can avoid the aliasing of the spectrum and make the decomposition and reconstruction of the image better.

The process of local non-subsampled shearlet reconstruction is divided into two steps: (1) the coefficients of high-pass sub-images at the same scale but in different directions are accumulated after decomposition and then the reconstructed coefficients are obtained; and (2) an inverse transformation of the decomposed low-pass sub-image coefficients and the reconstructed high-pass sub-images coefficients is performed to obtain the reconstructed image. A local non-subsampled shearlet has good local properties, excellent direction selectivity, and parabolic edge characteristics. The image f is decomposed by the m -layer LNSST to obtain $\Sigma_m 2^{dm}$ high-pass directional sub-images and one 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 Trui 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 8 (the number of stages is 3), and the size of the shearlet filter is 15×15 .

3. Singular value decomposition of image

Singular value decomposition provides a new method of extracting algebraic features of images, and it has been diffusely adapted for data compression, signal processing, and pattern recognition [19,20]. The singular value decomposition theorem and its properties are described in detail as follows.

If the matrix $A \in R^{m \times n}$, then the two orthogonal matrices $U = [u_1, u_2, \dots, u_m] \in R^{m \times m}$ and $V = [v_1, v_2, \dots, v_n] \in R^{n \times n}$ occur and one diagonal group $S = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_p]$ occurs. For $U^T U = I$, $V^T V = I$, $p = \min(m, n)$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p > 0$, and the following is obtained:

$$A = USV^T = \sum_{i=1}^p \sigma_i u_i v_i^T. \quad (7)$$

Eq. (7) is considered the singular value decomposition of matrix A , where the matrix S is regarded as the singular value matrix. Here, σ_i ($i = 1, 2, \dots, p$) indicates the singular value of matrix A . The singular value of the matrix represents the essential nature of the image rather than the visual characteristics and reflects the intrinsic properties of the image. The energy E of matrix A can be expressed as $E = \|A\|_F^2$, which can be proven in Eq. (8):

$$\begin{aligned} E &= \|A\|_F^2 = \|S\|_F^2 = \|USV^T\|_F^2 = \text{tr}[(USV^T) \cdot (USV^T)^T] \\ &= \text{tr}(S \cdot S^T) = \|S\|_F^2 = \sum_{i=1}^p \sigma_i^2. \end{aligned} \quad (8)$$

As shown in Eq. (8), after image A is decomposed by a singular value, the energy information is concentrated in the singular value matrix S . Specifically, the singular value of the image represents the energy characteristic of the image. When Fig. 3(a) is singularly decomposed and reconstructed, the singular value vector is replaced with the unit diagonal matrix of the same size, and Fig. 3(b) is obtained.

As shown in Fig. 3(a), after the singular value vector is stripped, the energy information of the image is almost completely shielded. Furthermore, the singular value vector of the image contains most of the energy information of the image. This principle is also the theoretical basis of the proposed algorithm.

4. Adaptive dual-channel PCNN

The PCNN is obtained by studying the neuron stimulation in the mammalian visual system, and it conforms to the visual information system of the human brain. Therefore, the PCNN model with a bionic mechanism can retain the spatial 2D information integrity of the input image the image is fused. Using the perfect combination of the PCNN model and visual characteristics, a better fusion effect is obtained [21–23]. Traditional PCNN models have many parameters, and most of them cannot be ignored; thus, we use a simplified dual-channel PCNN model. The expression is as follows:

$$F_{ij}^1(n) = S_{ij}^1(n). \quad (9)$$

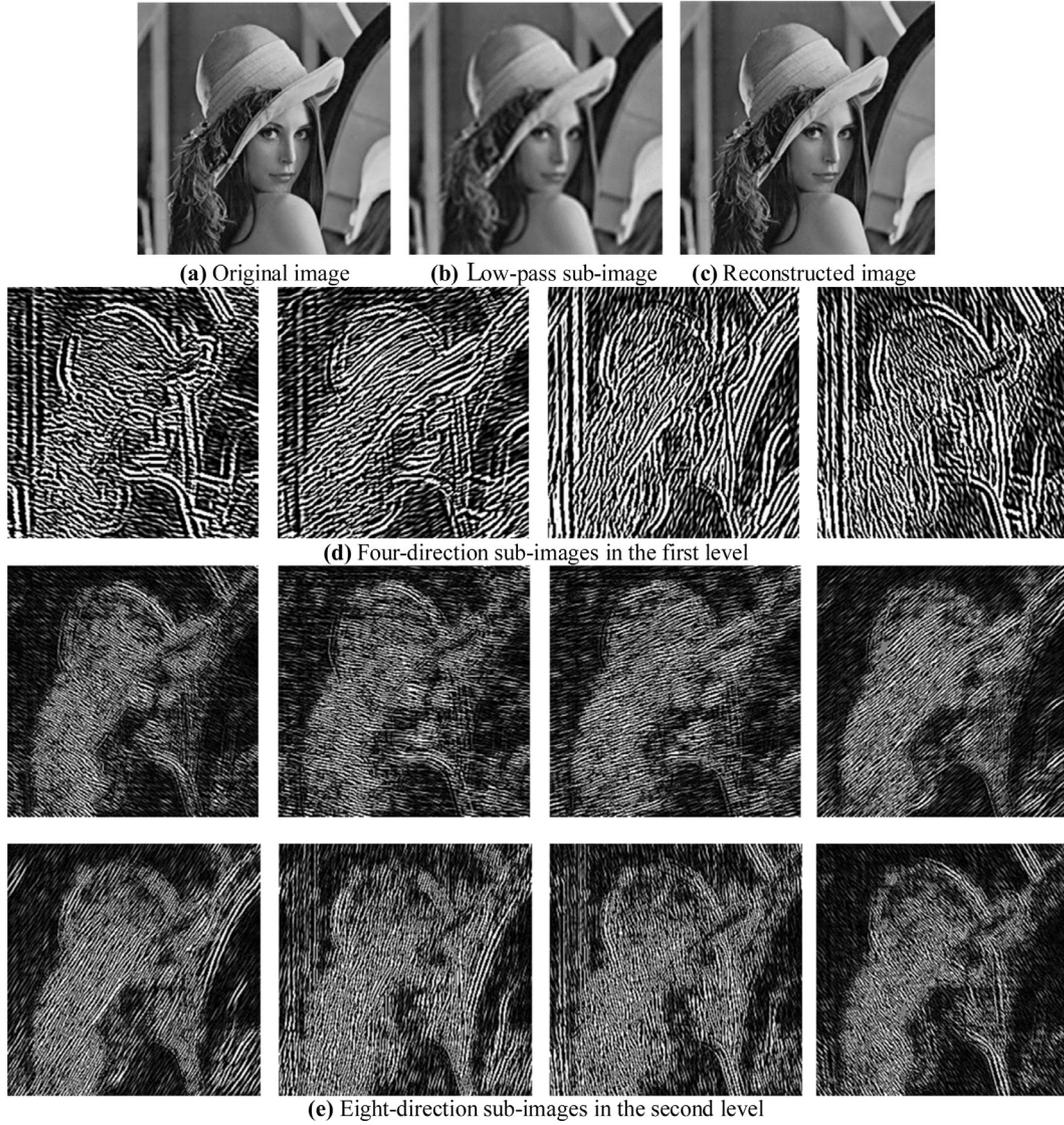


Fig. 2. Two-level LNSST of the image Linda.

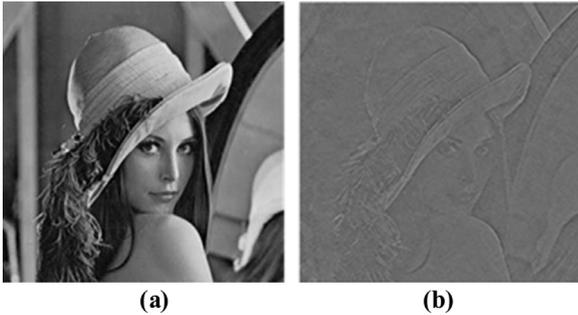


Fig. 3. Residual image after the singular value is stripped.

$$F_{ij}^2(n) = S_{ij}^2(n), \quad (10)$$

$$L_{ij}(n) = \begin{cases} 1 & \text{if } \sum_{k,l \in N(i,j)} Y_{kl}(n-1) > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (11)$$

$$U_{ij}(n) = \max \{ F_{ij}^1(n)(1 + \beta_{ij}^1 L_{ij}(n)), F_{ij}^2(n)(1 + \beta_{ij}^2 L_{ij}(n)) \}, \quad (12)$$

$$Y_{ij}(n) = \begin{cases} 1 & \text{if } U_{ij}(n) \geq \theta_{ij}(n-1) \\ 0 & \text{otherwise} \end{cases}, \quad (13)$$

$$\theta_{ij}(n) = \theta_{ij}(n-1) - \Delta + V_{\theta} Y_{ij}(n), \quad (14)$$

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

This model also includes the following parts: receiving domain, information fusion domain, and pulse generator, where S_{ij}^1 and S_{ij}^2 represent the gray-scale values normalized at the (i, j) position of the two source images. Their values are regarded as the external stimulus for the model. L_{ij} is the linking input, and F_{ij}^1 and F_{ij}^2 represent two symmetrical feedback inputs. Furthermore, β_{ij}^1 and β_{ij}^2 represent the linking strength; Y_{ij} denotes the external output of the neurons after internal processing; U_{ij} is treated as the internal activity item of the neurons; Δ is the declining extent of the

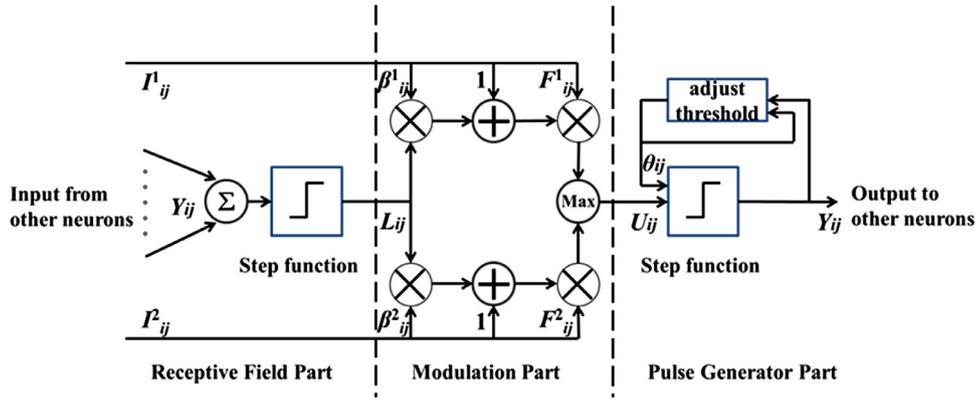


Fig. 4. Adaptive dual-channel PCNN model.

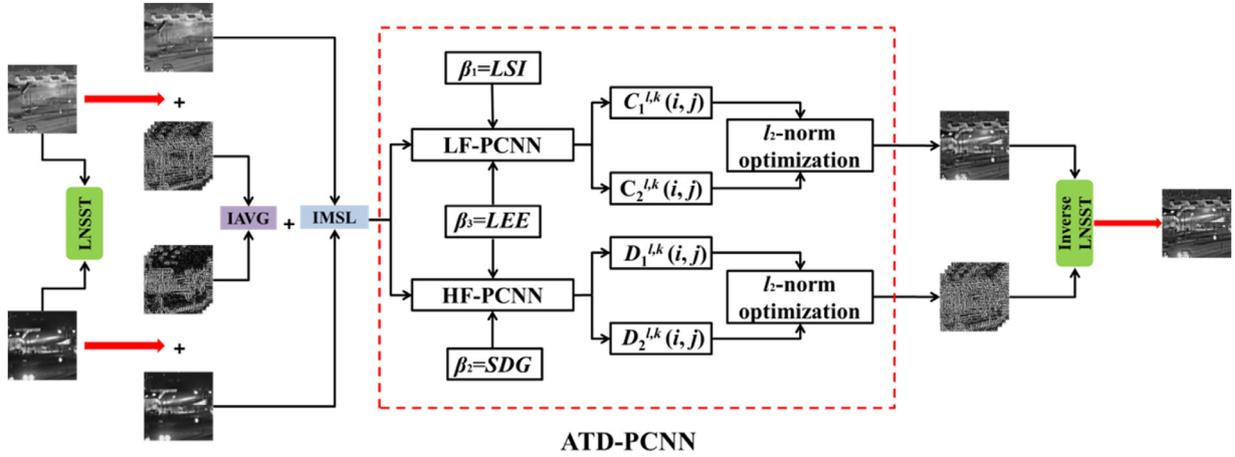


Fig. 5. Schematic of our proposed algorithm.

dynamic threshold, which can be regarded as 0.01 to ensure that the decay rate of the dynamic threshold is mild; θ_{ij} is the threshold function; and V_θ determines the threshold of the fired neuron and thus should be set to a relatively large value, e.g., 10. The time matrix T_{ij} [24] is used to adaptively determine the number of reasonable iterations. The simplified dual-channel PCNN model is illustrated in Fig. 4.

5. Fusion rule

5.1. The fusion framework

The algorithm used in this paper aims at solving the fusion of heterogeneous images. The specific fusion framework is shown in Fig. 5. The proposed fusion steps can be summarized as follows:

- (1) The LNSST is used for multi-scale decomposition of the IR and VI images.
- (2) Because of the essential difference between the high-frequency and low-frequency components, two adaptive dual-channel PCNN models with different linking strength and external stimuli are used to process them separately. For the low-frequency components of PCNN (LF-PCNN), a local structure information factor is used as the main linking strength, and an improved sum-modified Laplacian (ISML) is used as the external stimulus. For the high-frequency components of PCNN (HF-PCNN), a sum of the direction gradients is treated as the main linking strength, and an improved average gradient (IAVG) is treated as the external stimulus.

- (3) This paper proposes a Laplacian edge energy operator as the auxiliary linking strength, which is also regarded as the linking strength of the LF-PCNN and HF-PCNN, and is simultaneously input into the respective dual-channel PCNN model. Because the dual-channel PCNN model can directly output the coefficients of the fused image, the two sets of fusion coefficients will be outputted when the model has two linking strengths. For this reason, an optimization model [23] based on l_2 -norm is used to combine the two sets of fusion coefficients into one group.

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

5.2. Fusion rule of low-pass sub-images

The low-pass component of the image represents the main energy of the image, which is an approximate part of the image, so the fusion rules of low-frequency components determine the final fusion effect. The local structure information factor is an excellent mathematical feature that concentrates the energy of the image and contains most of the structural information of the image. The value can reflect the regional features and the low-frequency characteristics of the image. An improved sum-modified Laplacian (ISML) algorithm that indicates the energy information of the low-pass sub-images is utilized as an external stimulus for the LF-PCNN.

Step 1: In this paper, the ISML is an upgrade based on the novel sum-modified Laplacian (NSML). The NMSL represents the horizontal and vertical directions of the Laplacian energy and lacks energy information in the diagonal direction, which means that important details could be missing from the fusion result. Therefore, the

specific expression is improved as follows:

$$\begin{aligned} \nabla^2 ML_1^k(i, j) &= \nabla^2 ML_1^{l,k}(i, j) + \nabla^2 ML_2^{l,k}(i, j) \\ &\quad + \nabla^2 ML_3^{l,k}(i, j) + \nabla^2 ML_4^{l,k}(i, j), \end{aligned} \quad (16)$$

$$\nabla^2 ML_1^{l,k}(i, j) = |2C_1^k(i, j) - C_1^k(i-1, j) - C_1^k(i+1, j)|, \quad (17)$$

$$\nabla^2 ML_2^{l,k}(i, j) = |2C_1^k(i, j) - C_1^k(i, j-1) - C_1^k(i, j+1)|, \quad (18)$$

$$\begin{aligned} \nabla^2 ML_3^{l,k}(i, j) \\ = |2C_1^k(i, j) - C_1^k(i-1, j-1) - C_1^k(i+1, j+1)|/\sqrt{2}, \end{aligned} \quad (19)$$

$$\begin{aligned} \nabla^2 ML_4^{l,k}(i, j) \\ = |2C_1^k(i, j) - C_1^k(i-1, j+1) - C_1^k(i+1, j-1)|/\sqrt{2}, \end{aligned} \quad (20)$$

$$ISML_1^k(i, j) = \sum_m \sum_n W(m, n) (\nabla^2 ML_1^k(i+m, j+n))^2, \quad (21)$$

$$W(m, n) = \frac{1}{15} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \quad (22)$$

where $C_l^k(i, j)$ represents the low-pass multi-scale decomposition coefficients of the pixels, l represents the decomposition number, k represents the direction of decomposition of each layer, and $W(m, n)$ represents the weighted templates. In this study, 3×3 windows are used, $ISML_1^k(i, j)$ is the improved sum-modified Laplacian operator, and the operator contains the row energy ML_1 , column energy ML_2 , and diagonal energy ML_3 and ML_4 .

Step 2: The low-pass sub-image is separated into blocks by a sliding window. The size of the block is 3×3 , the center point is (i, j) , and then calculate the singular value of each sub-block image.

Step 3: The singular value of each sub-block image is calculated based on Eq. (23), and then the value of the element is equivalent to the corresponding main linking strength.

$$\beta_1(i, j) = LSI(i, j) = \|\sigma(i, j)\|_F^2, \quad (23)$$

where $\sigma(i, j)$ is the singular value in the local region, and $\|\cdot\|_F$ denotes the F -norm of the matrix. The value of $LSI(i, j)$ reflects the basic structure of the image area and the richness of the detail of the local area information, which can describe the changes in the image features.

Step 4: To fully utilize the energy information of the image and effectively extract the details of the image, this paper proposes a LEE operator as the auxiliary linking strength, and its purpose is to consolidate the stability of the main linking strength. When LSI is used as the main linking strength, let LEE input into the PCNN model at the same time, and its expression is as follows:

$$\begin{aligned} \omega_1 &= \begin{bmatrix} -1 & -1 & -1 \\ -1 & A+8 & -1 \\ -1 & 1 & -1 \end{bmatrix}, \quad \omega_2 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & A+4 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \\ w(m, n) &= \frac{1}{15} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \end{aligned} \quad (24)$$

$$E_1^{l,k}(i, j) = \left[\sum_{m=-1}^1 \sum_{n=-1}^1 \omega_1(m+2, n+2) C_1^k(i+m, j+n) \right]^2, \quad (25)$$

$$E_2^{l,k}(i, j) = \left[\sum_{m=-1}^1 \sum_{n=-1}^1 \omega_2(m+2, n+2) C_1^k(i+m, j+n) \right]^2, \quad (26)$$

$$EOE_1^k(i, j) = E_1^{l,k}(i, j) + E_2^{l,k}(i, j), \quad (27)$$

$$\beta'(i, j) = LEE_1^k(i, j) = \sum_m \sum_n w(m, n) \cdot EOE(i, j). \quad (28)$$

where ω_1 and ω_2 are the Laplacian operators, $w(m, n)$ represents the weighted templates, $C_l^k(i, j)$ represents the multi-scale decomposition coefficients of low-pass sub-images, and $LEE_1^k(i, j)$ denotes the Laplacian edge energy. This operator can simultaneously reflect the energy of the edge information and the change in the edge information; therefore, it can assist the local structure information factor as the low-frequency linking strength to clarify the edge energy of the low-pass sub-image so that the low-pass energy information can be fully inputted to the dual-channel PCNN model.

Step 5: Iterate the Eqs. (9)–(15) until all the neurons are ignited and then calculate $U_{ij}(n)$, $L_{ij}(n)$, $\theta_{ij}(n)$, $T_{ij}(n)$, and $Y_{ij}(n)$. The fused coefficients can be chosen as follows:

$$C_l^k(i, j) = \begin{cases} C_L^{l,k}(i, j) & U_{ij}(n) = U_{ij}^L(n) \\ C_V^{l,k}(i, j) & U_{ij}(n) = U_{ij}^V(n) \end{cases}, \quad (29)$$

$$U_{ij}^L(n) = F_{ij}^L(n)(1 + \beta_{ij}^L L_{ij}(n)), \quad (30)$$

$$U_{ij}^V(n) = F_{ij}^V(n)(1 + \beta_{ij}^V L_{ij}(n)). \quad (31)$$

where the low-pass multi-scale decomposition coefficients of the IR image, VI image, and fusion image are represented by $C_L^{l,k}(i, j)$, $C_V^{l,k}(i, j)$ and $C_l^k(i, j)$, respectively. Here, n denotes the total ignition time.

Step 6: An optimization model based on l_2 -norm [25] is used to combine the two sets of fusion coefficients into one group, and the process is as follows:

$$\arg \min \left\{ \|C_1^k(i, j) - C_1^{l,k}(i, j)\|_2^2 + \lambda \|C_1^k(i, j) - C_2^{l,k}(i, j)\|_2^2 \right\}, \quad (32)$$

where $C_l^k(i, j)$ represents the final fusion coefficients of each sub-images, $C_1^{l,k}(i, j)$ and $C_2^{l,k}(i, j)$ represent the fusion coefficients generated by the two linking strength of LF-PCNN model, respectively. This model constrain $C_l^k(i, j)$ close to the final fusion coefficients, and the parameter λ is used to control the balance of $C_1^{l,k}(i, j)$, $C_2^{l,k}(i, j)$ and $C_l^k(i, j)$.

5.3. Fusion rule of high-pass sub-images

The high-pass sub-images always reflect the edge features and texture details of the image, so an IAVG that represents the details of the texture of the high-pass sub-images to stimulate the HF-PCNN is recommended. The SDG can also reflect the high-frequency information, so it is chosen as the main linking strength of HF-PCNN.

Step 1: An average gradient (AVG) is proposed in other papers as a gradient feature of the image, and it reflects the details of the edge of the image and represents the change in the gray value of the image. Similar to the IMSL principle, two diagonal gradient changes are added based on previously cited studies to more completely fuse the edge information of two images to prevent a loss of details. The IAVG is shown in the following equations:

$$\begin{aligned} IAVG_1^k(i, j) &= \frac{1}{M \times N} \sum_{i=(1-M)/2}^{(M-1)/2} \sum_{j=(1-N)/2}^{(N-1)/2} \\ &\quad \times \sqrt{\frac{\nabla D_h^2(i, j) + \nabla D_v^2(i, j) + \nabla D_{md}^2(i, j) + \nabla D_{vd}^2(i, j)}{2}}, \end{aligned} \quad (33)$$

$$\nabla D_h(i, j) = D_l^k(i, j) - D_l^k(i + 1, j), \quad (34)$$

$$\nabla D_v(i, j) = D_l^k(i, j) - D_l^k(i, j + 1), \quad (35)$$

$$\nabla D_{md}(i, j) = [D_l^k(i, j) - D_l^k(i + 1, j + 1)]/\sqrt{2}, \quad (36)$$

$$\nabla D_{vd}(i, j) = [D_l^k(i + 1, j) - D_l^k(i, j + 1)]/\sqrt{2}. \quad (37)$$

where $D_l^k(i, j)$ represents the multi-scale decomposition coefficients of high-pass sub-images, $M \times N$ is taken as 3×3 , and $\nabla D_h(i, j)$, $\nabla D_v(i, j)$, $\nabla D_{md}(i, j)$, $\nabla D_{vd}(i, j)$ represent the gradient changes in the horizontal, vertical, main diagonal, and oblique diagonal directions, respectively.

Step 2: The sum of the direction gradients can reflect the change in the edge details of the image and combine the information in the horizontal, vertical and diagonal direction at the pixel level, which is one of the key indexes for reflecting the image clarity. Thus, this index can be used as the linking strength of the high-pass sub-images. The mathematical expression is as follows:

$$SDG = \nabla D_h + \nabla D_v + \nabla D_{md} + \nabla D_{vd}, \quad (38)$$

$$\nabla C_h = \sqrt{(D_l^k(i, j) - D_l^k(i - 1, j))^2 + (D_l^k(i, j) - D_l^k(i + 1, j))^2}, \quad (39)$$

$$\nabla C_v = \sqrt{(D_l^k(i, j) - D_l^k(i, j - 1))^2 + (D_l^k(i, j) - D_l^k(i, j + 1))^2}, \quad (40)$$

$$\nabla C_{md} = \sqrt{((D_l^k(i, j) - D_l^k(i - 1, j - 1))^2 + (D_l^k(i, j) - D_l^k(i + 1, j + 1))^2)/\sqrt{2}}, \quad (41)$$

$$\nabla C_{vd} = \sqrt{((D_l^k(i, j) - D_l^k(i - 1, j + 1))^2 + (D_l^k(i, j) - D_l^k(i + 1, j - 1))^2)/\sqrt{2}}, \quad (42)$$

$$\beta_2(i, j) = SDG_l^k(i, j). \quad (43)$$

where $D_l^k(i, j)$ is the high-pass coefficient obtained by multi-scale analysis, l is the number of decomposed layers, k is the number of decomposition directions of each layer. SDG_l^k represents the sum of the direction gradients at the pixel element; ∇C_h and ∇C_v represent changes in the horizontal and vertical directions, respectively; and ∇C_{md} and ∇C_{vd} represent the changes in the diagonal directions. Greater direction gradient sums correspond to higher local area clarity and faster neuronal activation.

Step 3: Similar to the fourth step in Section 5.2, the Laplacian edge energy of the high-frequency sub-images is calculated by using Eqs. (24)–(28). For the high-pass sub-images, because this operator can also represent the transformation of the edge information, it can facilitate the sum of the direction gradients with the same effect, and the high-pass detail information is also more fully inputted into the fused image. The Eqs. (44) and (45) is changed as follows:

$$E_1^{l,k}(i, j) = \left[\sum_{m=-1}^1 \sum_{n=-1}^1 \omega_1(m+2, n+2) D_l^k(i+m, j+n) \right]^2, \quad (44)$$

$$E_2^{l,k}(i, j) = \left[\sum_{m=-1}^1 \sum_{n=-1}^1 \omega_2(m+2, n+2) D_l^k(i+m, j+n) \right]^2, \quad (45)$$

Step 4: Similar to the fifth step in Section 5.2, iterate the Eqs. (9)–(15) and Eqs. (29)–(31) until all the neurons are ignited, and then the final high-frequency fusion coefficients $D_l^k(i, j)$ is obtained.

Step 5: Similar to the sixth step in Section 5.2, use the optimization model to combine the output coefficients:

$$\arg \min \left\{ \|D_l^k(i, j) - D_1^{l,k}(i, j)\|_2^2 + \lambda \|D_l^k(i, j) - D_2^{l,k}(i, j)\|_2^2 \right\}. \quad (46)$$

5.4. Supplementary instructions

The above-mentioned adaptive dual-channel PCNN model with triple-linking strength is called as the **ATD-PCNN** model. The consequent of the linking strength calculated from the above three operators is shown in Fig. 6. Obviously, Fig. 6(b) is similar to the source image, so the local structural information factor can represent the energy characteristic of the image very well; Fig. 6(c) shows the edge information of the source image clearly, so the texture information of the image can be delivered in detail to the ATD-PCNN model by the sum of direction gradient; compared with Fig. 6(b) and (c), Fig. 6(d) contains both the detail and the subject information, so the Laplacian edge energy operator is very suitable to act as an auxiliary linking strength. In general, the local structure information (LSI) factor contains the energy information of the image, so it is similar to the essence of the low-frequency components, and then it can be used as the linking strength of the low-frequency components of PCNN model. Similarly, the sum of directional gradient (SDG) operator can capture the texture details of the image, so it is similar to the essence of the high-frequency components, and then it can be used as the linking strength of the high-frequency components of PCNN model. Unlike other types of image fusion, the purpose of infrared and visible light image fusion is to combine the spatial energy distribution information of the infrared image with the texture detail information of the visible light image, so the spectral differences of them should be considered, and the LSI and SDG can precisely play the role. Since the research in this paper is based on the fusion of infrared and visible light image, the above two linking strengths are set for them, so the LSI and SDG are not suitable for other types of image fusion.

6. Experimental results and analysis

To verify the superiority of the proposed method, our method is compared with the following methods: NSCT-DUAL-PCNN [10], NSST-SF-PCNN [13], GFF [26], VSM-WLS [27], CNN [28], IFE-VIP [29], GTF [30], PCNN, and NSCT-ATD-PCNN. To show the breadth of the algorithm, six different environments are selected. The first group shows a man in the doorway, the second group depicts a soldier in a trench, the third group shows a jeep in front of a house, the fourth group shows a man entering a building, the fifth group depicts a ship on a lake, and the sixth group shows a man hidden in the forest. The original image size of the first two groups is 768×576 , the size of the middle two groups is 620×450 , and the size of the last two groups is 505×510 . In this paper, our method takes “maxflat” and [2, 2, 3, 3] as the direction filter and the pyramid filter of LNSST, respectively, and $\lambda = 0.002$, $A = 32$; NSCT-ATD-PCNN use “9/7”, “pkva”, and [0, 2, 3, 3] as the direction filters and the pyramid filter, respectively; and PCNN method are set as: $\alpha_L = 0.06931$, $\alpha_\theta = 0.2$, $V_L = 1$, $V_\theta = 20$, $\theta = 0.2$, $N = 200$, and $W = [0.707, 1.0.707; 1.0, 1; 0.707, 1.0.707]$. The source images of each group of infrared and visible light are strictly registered, and they can be downloaded from the site:

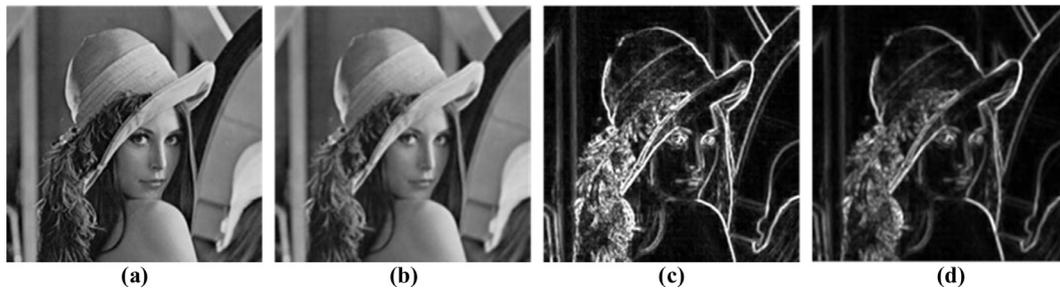


Fig. 6. The images calculated by the linking strength. (a) The source image. (b) The image calculated by the local structural information operator. (c) The image calculated by the sum of direction gradient. (d) The image calculated by the Laplacian edge energy operator.

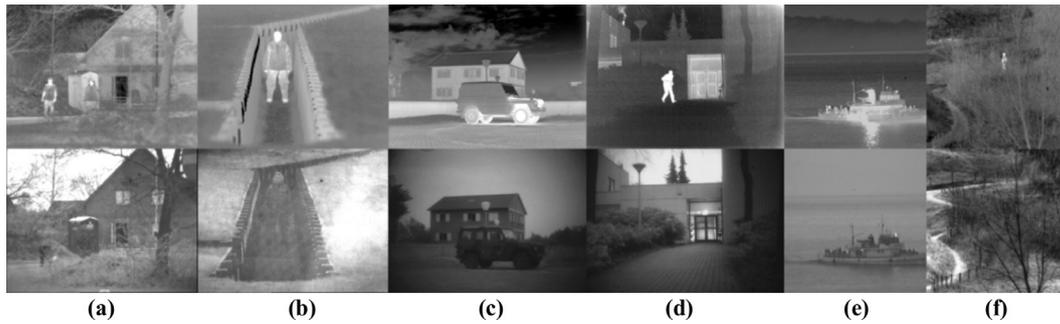


Fig. 7. Source images.

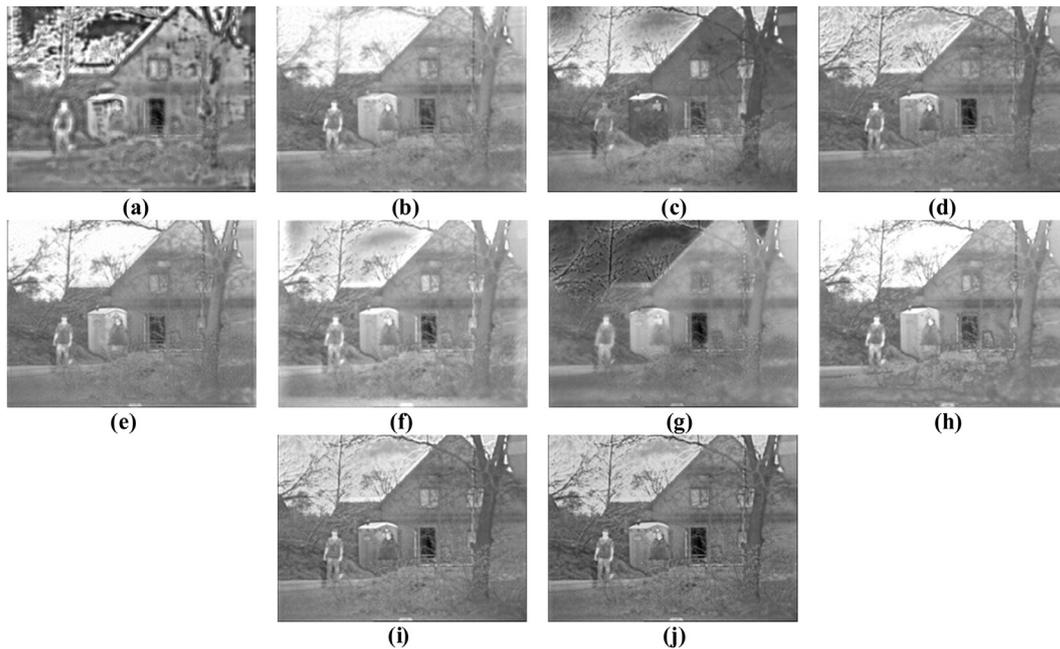


Fig. 8. The first group of fusion experiment.

https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029. All experiments are conducted in MATLAB 2012a on a PC with an Intel Core i7/3.4GHz/4G processor. The six sets of source images are presented in Fig. 7(a)–(f) from the first to the sixth group.

The NSCT-DUAL-PCNN, NSST-SF-PCNN, GFF, VSM-WLS, CNN, IFE-VIP, GTF, PCNN, NSCT-ATD-PCNN methods and the proposed method are shown from Fig. 8(a)–(j) to Fig. 13(a)–(j). The first group to the sixth group of experiments are shown from top to bottom. In each experiment, the NSCT-DUAL-PCNN method shows a large amount of black pseudo-noise. Without the best selection for the linking strength, this method causes the fused image to

be blurred, which significantly affects the perceptibility. When the dual-channel PCNN model is not applied, the image contrast and edge details of the NSST-PCNN method are not clarified. The GFF and GTF methods poorly integrate visible light information into the fused image, which results in a fused image that is close to the original infrared image. The CNN method is better for the fusion of infrared targets, although the background of the infrared information is not well integrated into the fused image. The IFE-VIP method has issues with high brightness, which results in a reduction in image contrast and inability to retrieve important information. The overall perception of VSM-WLS method is gloomy,

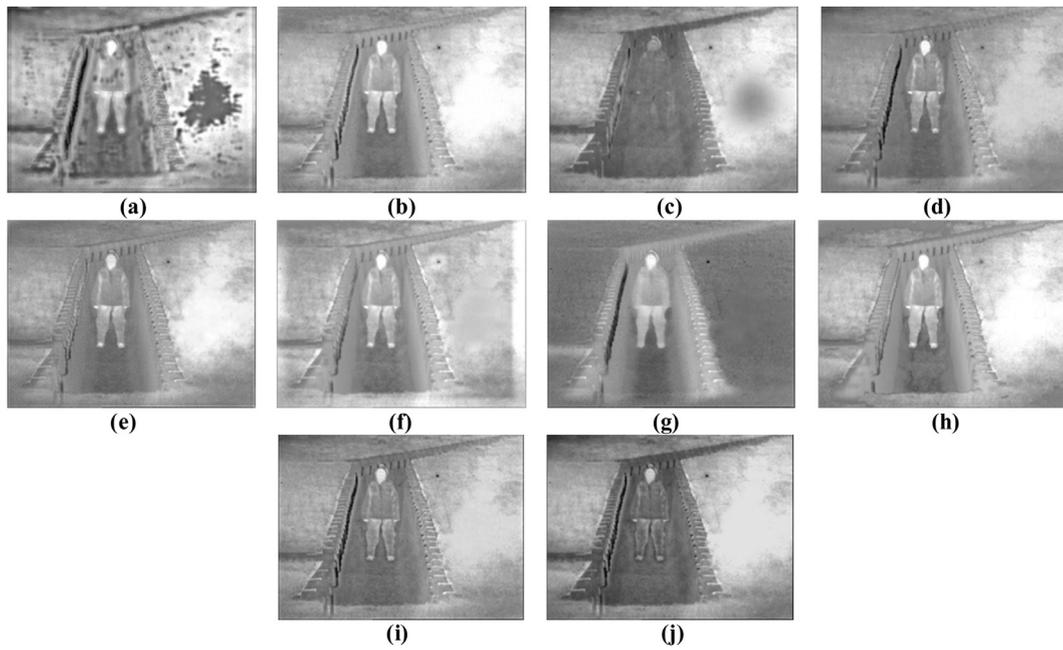


Fig. 9. The second group of fusion experiment.

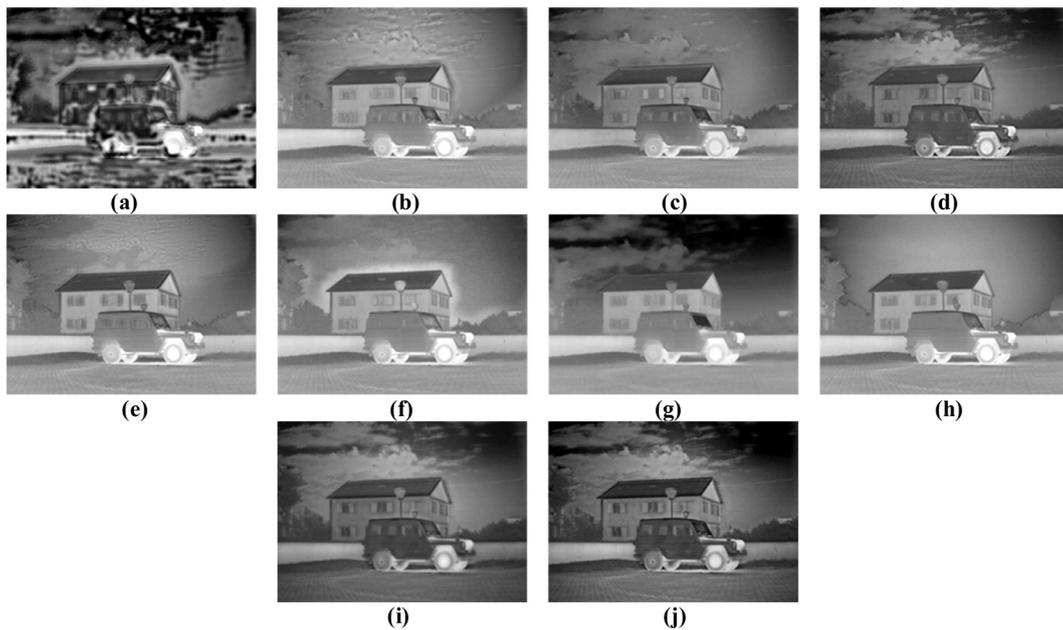


Fig. 10. The third group of fusion experiment.

so the contrast ratio of the fused images is relatively low. Because the traditional PCNN method does not combine with MGA tools, neither adopt the setting of adaptive parameters, the final fused images lose many details and have poor impressions. The NSCT-ATD-PCNN method uses the same ATD-PCNN model as the proposed algorithm in this paper. However, because the decomposition effect of NSCT is not as good as the LNSST, the final fused images are inferior to the proposed algorithm in detail and contrast ratio. Compared with the above mentioned algorithms, the fused images based on our algorithm have the highest contrast ratio and can fuse the IR target and background information under the human visual field. Furthermore, the proposed algorithm does not lose the edge detail information in the fusion process; thus, its performance is the best.

6.2. Objective evaluation

Usually, image fusion results can be evaluated in subjectively and objectively. Under most circumstances, fusion results present limited differences; thus, correctly subjectively evaluating fusion results is difficult. 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 [31], (2) information entropy (IE) [32], (5) edge retentiveness ($Q^{AB/F}$) [33] (3) space frequency (SF) [34], and (4) standard deviation (SD) [35].

A detailed quantitative evaluation of the six groups of IR and VI images is shown in Tables 1–3. The values in boldface

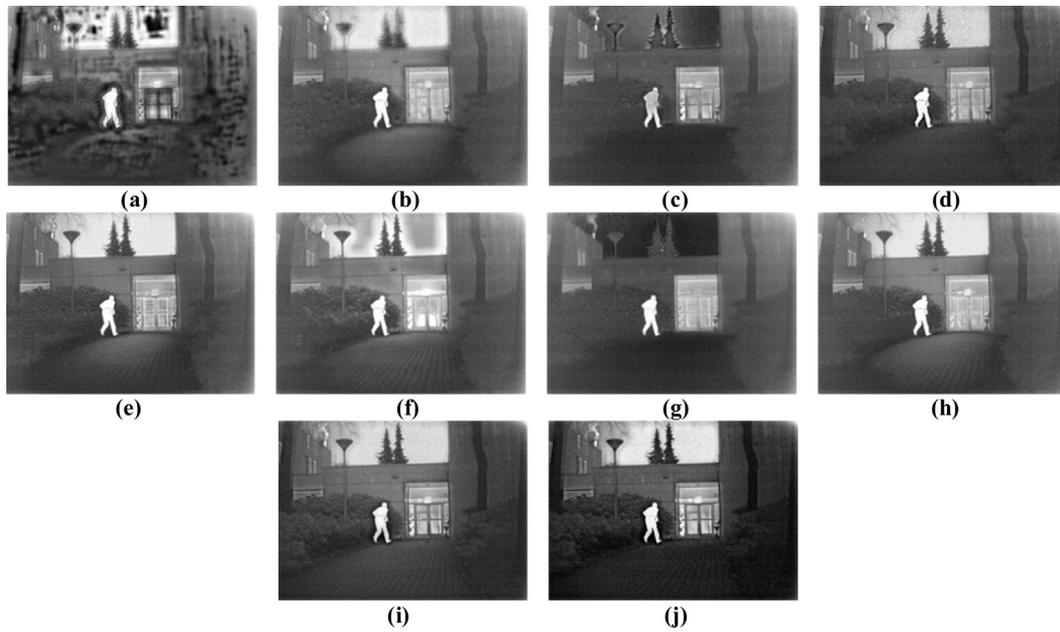


Fig. 11. The fourth group of fusion experiment.

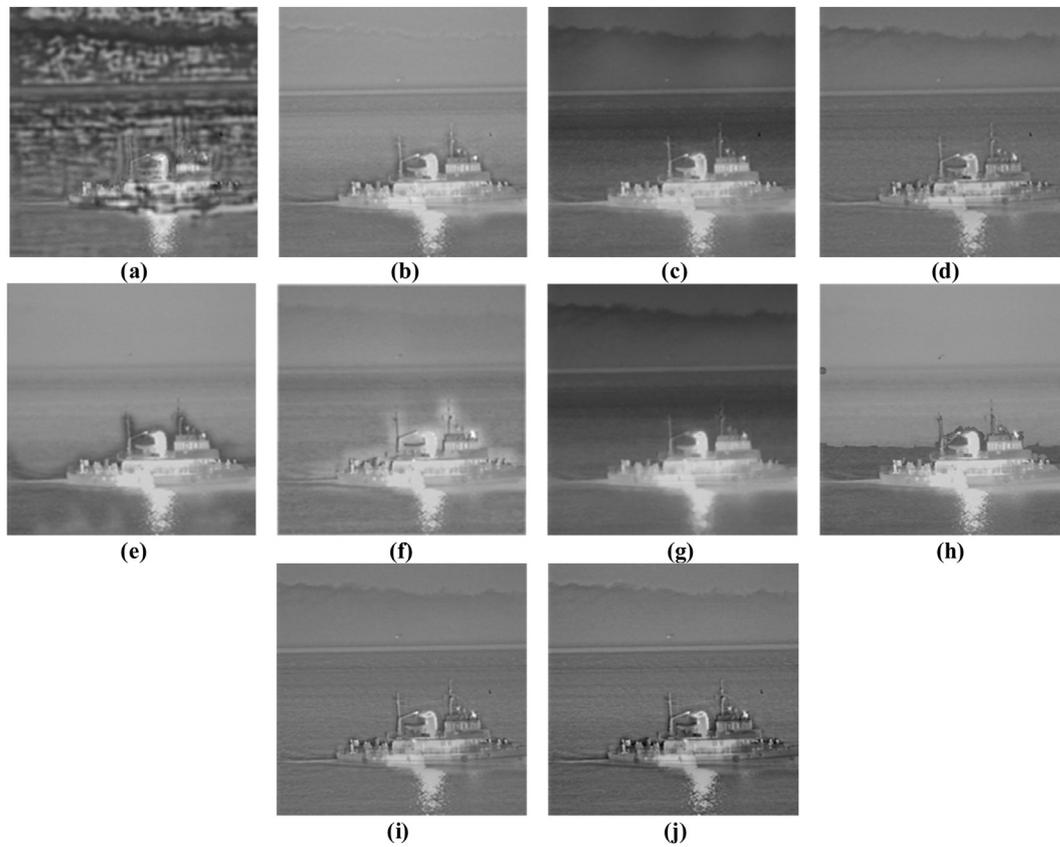


Fig. 12. The fifth group of fusion experiment.

represent the best results using the same index in the aforementioned method.

The SD of our algorithm is higher than that of other algorithms as indicated in Table 1, which shows that the gray value distribution of our algorithm is uniform; thus, our algorithm can achieve the best detailed transformation and clarity in various cases. In the first group of experiments, the $Q^{AB/F}$ value of the proposed method

is slightly lower than that of the GFF algorithm, and the SF value of the proposed method is slightly lower than that of the VSM-WLS algorithm, which is mainly related to the fusion of the sky background. In the second group of experiments, the evaluation parameters of the proposed algorithm are higher than those of the other algorithms; thus, the superiority of the proposed algorithm is again demonstrated. In the third group of experiments, the value

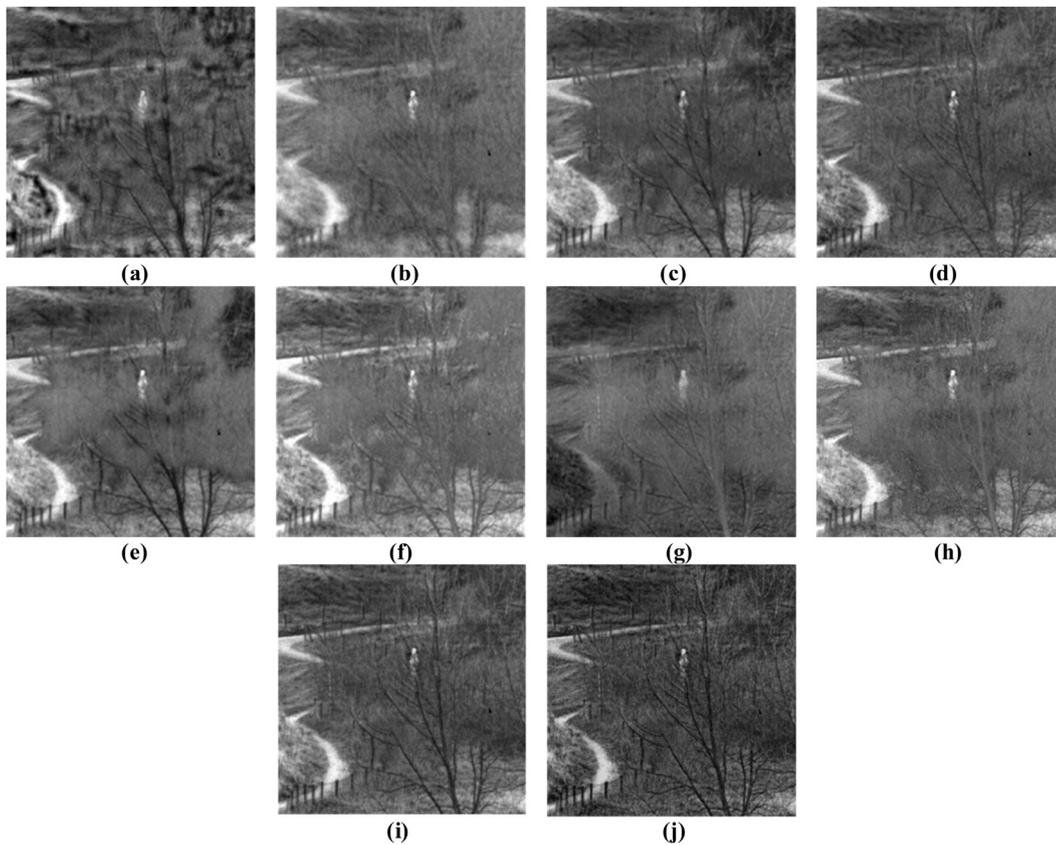


Fig. 13. The sixth group of fusion experiment.

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

Group	Fusion methods	Evaluation index				
		AVG	IE	QA ^{B/F}	SF	SD
1	NSCT-PCNN	4.9012	7.2616	0.3573	12.7019	41.8997
	NSST-PCNN	4.1972	7.1106	0.3496	9.8456	33.9946
	GFF	4.8740	7.1255	0.5292	11.7246	41.9639
	VSM-WLS	5.3883	6.9847	0.4658	13.6016	39.0442
	CNN	5.3250	6.8158	0.4846	12.7526	37.5135
	IFE-VIP	4.5750	6.9247	0.5208	10.2045	36.8846
	GTF	5.0240	6.9137	0.5004	11.4568	38.4232
	PCNN	4.0534	6.7437	0.4307	11.2735	35.6836
	NSCT-ATD-PCNN	5.3843	7.1747	0.5142	12.4678	48.5212
	Proposed	5.4955	7.2817	0.5254	12.8566	50.3672
	2	NSCT-PCNN	6.0503	7.4272	0.4784	13.5230
NSST-PCNN		5.3765	6.6793	0.5100	13.4397	29.6847
GFF		5.4261	7.3881	0.6071	13.0951	46.2241
VSM-WLS		6.9869	6.9736	0.5813	13.8672	34.1787
CNN		5.8501	7.1369	0.5829	14.4781	40.5863
IFE-VIP		4.7424	6.9278	0.5756	10.7570	33.3696
GTF		5.5264	6.8724	0.5920	12.7345	40.4321
PCNN		4.4315	6.7389	0.5065	11.8681	28.4707
NSCT-ATD-PCNN		6.8375	7.2835	0.6131	13.7536	46.5432
Proposed		6.9650	7.4347	0.6376	14.5025	48.0779

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

Group	Fusion methods	Evaluation index				
		AVG	IE	QA ^{B/F}	SF	SD
3	NSCT-PCNN	4.9858	7.6696	0.3820	6.1437	53.0802
	NSST-PCNN	3.3474	7.2022	0.5686	6.9689	56.4587
	GFF	2.8348	7.1193	0.5831	6.1894	35.6565
	VSM-WLS	5.3423	7.1823	0.4487	9.1661	36.9416
	CNN	3.3754	7.1753	0.5005	7.0172	38.2479
	IFE-VIP	3.2863	7.1574	0.5659	6.8684	37.6036
	GTF	3.4752	7.1932	0.4952	7.0242	41.9325
	PCNN	3.0752	7.1083	0.3665	6.0675	40.7127
	NSCT-ATD-PCNN	5.4631	7.2143	0.5543	9.0351	55.5432
	Proposed	5.5721	7.2959	0.5761	9.3107	57.1198
	4	NSCT-PCNN	5.0073	7.3628	0.3908	9.2786
NSST-PCNN		4.3398	6.9941	0.4344	8.1744	39.4377
GFF		4.1216	6.8889	0.6072	7.9503	34.3422
VSM-WLS		5.0290	6.9702	0.4720	9.5800	49.3080
CNN		4.8633	7.1839	0.5074	9.3732	51.8397
IFE-VIP		4.2545	7.1547	0.5324	8.6974	49.3352
GTF		4.3243	6.8742	0.4634	8.0216	50.8255
PCNN		4.0434	6.6638	0.4235	7.8865	40.2263
NSCT-ATD-PCNN		4.8254	7.2753	0.5447	9.5237	55.7344
Proposed		4.9107	7.3250	0.5668	9.6053	57.2323

of $Q^{A/B}$ of the proposed algorithm is slightly lower than that of the GFF algorithm, but the difference is still small. However, the IE value of the proposed algorithm is lower than that of the NSCT-DUAL-PCNN because the NSCT-DUAL-PCNN algorithm causes many black artifacts in the fusion, which leads to an abnormal increase in the gray value; thus, the IE is unrealistic. In the fourth group of experiments, the AVG value of the proposed algorithm is slightly lower than that of the VSM-WLS algorithm, and the other values are higher than those of the other algorithms. The fifth group of

experiments is similar to the fourth group, and only the SF value of the proposed algorithm is lower than that of the VSM-WLS algorithm, which shows that the VSM-WLS algorithm is similar to our proposed algorithm. In the sixth group of experiments, the $Q^{A/B}$ value of the proposed algorithm is slightly lower than that of the CNN algorithm, and the other parameters are higher compared with the other methods. Based on the objective evaluation parameters, the final result is the same as that of the superior. The proposed algorithm is superior to the others algorithms in terms of

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

Group	Fusion methods	Evaluation index				
		AVG	IE	QA ^{B/F}	SF	SD
5	NSCT-PCNN	2.8453	5.7484	0.3542	7.3054	19.2228
	NSST-PCNN	2.7707	5.9209	0.5654	7.1446	19.8795
	GFF	2.4447	5.4152	0.5568	6.4845	20.4683
	VSM-WLS	3.1300	5.4133	0.5699	8.0035	16.9300
	CNN	1.9172	5.9175	0.5704	6.0754	19.4387
	IFE-VIP	2.4104	5.7929	0.4641	6.2482	19.8898
	GTF	2.0647	5.4326	0.5089	6.5424	17.7524
	PCNN	2.0432	5.3456	0.5255	6.5378	17.4567
	NSCT-ATD-PCNN	2.8789	5.9345	0.6022	7.6578	21.4590
	Proposed	2.9062	6.0003	0.6221	7.7212	22.7661
6	NSCT-PCNN	6.6058	6.9485	0.3350	11.0123	35.2992
	NSST-PCNN	5.1420	6.6522	0.3193	8.5655	29.8836
	GFF	5.6612	6.8422	0.5514	9.7923	34.4608
	VSM-WLS	6.5141	6.5431	0.4075	12.8170	27.9779
	CNN	5.7914	6.9653	0.5598	10.2403	34.8340
	IFE-VIP	6.7687	6.7255	0.4969	11.6890	32.5045
	GTF	5.9627	6.6752	0.4216	11.5228	34.8340
	PCNN	5.2560	6.5256	0.3678	10.1345	30.0908
	NSCT-ATD-PCNN	7.3215	7.0256	0.5267	12.3456	37.7876
	Proposed	7.3764	7.1396	0.5384	13.2142	39.6234

the image gray value distribution, edge detail, and clarity. Furthermore, the proposed algorithm has good perception, which prevents the limitations of black artifacts caused by the large difference between the VI and IR. The contrast of the image after fusion is good, the overall gray value transition is natural, and good perceptibility is observed.

7. Conclusion

This paper presented a novel fusion framework for VI and IR images based on the proposed LNSST-ATD-PCNN method. In our method, the LNSST is used as a multi-scale decomposition tool for the image based on the NSST, and the adaptive dual-channel PCNN model is also improved. To address the large spectral difference between IR and VI, three novel operators are treated as the adaptive linking strength. Furthermore, the ISML of the low-pass sub-image and the IAVG of the high-pass sub-images are input to stimulate the ATD-PCNN. To verify the fusion performance, six different scenarios are used. The outcome shows that our algorithm can effectively fuse VI and IR images with a high contrast while retaining a wealth of textural and detail information without any artifacts. The experimental results of the subjective and objective evaluation indicate that our algorithm has better fusion performance than typical fusion techniques.

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).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neucom.2018.05.028](https://doi.org/10.1016/j.neucom.2018.05.028).

References

[1] X. Bai, Morphological center operator based infrared and visible image fusion through correlation coefficient, *Infrared Phys. Technol.* 76 (2016) 546–554.
 [2] J. Zhao, G. Cui, X. Gong, Y. Zang, S. Tao, D. Wang, Fusion of visible and infrared images using global entropy and gradient constrained regularization, *Infrared Phys. Technol.* 81 (2017) 201–209.

[3] Q. Zhang, Y. Liu, R.S. Blum, J. Han, D. Tao, Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review, *Inf. Fusion* 40 (2018) 57–75.
 [4] F. Meng, B. Guo, M. Song, X. Zhang, Image fusion with saliency map and interest points, *Neurocomputing* 177 (2016) 1–8.
 [5] Y. Liu, S. Liu, Z. Wang, A general framework for image fusion based on multi-scale transform and sparse representation, *Inf. Fusion* 24 (2015) 147–164.
 [6] J. Zhao, Q. Zhou, Y. Chen, H. Feng, Z. Xu, Q. Li, Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition, *Infrared Phys. Technol.* 56 (2013) 93–99.
 [7] Y. Chai, H.F. Li, J.F. Qu, Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain, *Opt. Commun.* 283 (2010) 3591–3602.
 [8] G. Easley, D. Labate, W. Lim, Sparse directional image representations using the discrete shearlet transform, *Appl. Comput. Harmonic Anal.* 25 (1) (2008) 25–46.
 [9] W.Q. Lim, The discrete shearlet transform: a new directional image representation and compactly supported shearlet frames, *IEEE Trans. Image Process.* 19 (5) (2010) 1166–1180.
 [10] T. Xiang, L. Yan, R. 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.
 [11] Y. Chai, H.F. Li, J.F. Qu, Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain, *Opt. Commun.* 283 (2010) 3591–3602.
 [12] Z. Liu, Y. Feng, Y. Zhang, X. 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.
 [13] W. Kong, L. Zhang, Y. Lei, Novel fusion method for visible light and infrared images based on NSST-SF-PCNN, *Infrared Phys. Technol.* 65 (2014) 103–112.
 [14] M. Yin, P. Duan, W. Liu, X. 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.
 [15] J. Zhao, Q. Zhou, Y. Chen, H. Feng, Z. Xu, Q. Li, Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition, *Infrared Phys. Technol.* 56 (2013) 93–99.
 [16] Y. Chen, S. Tong, F. Cong, J. Xu, Symmetrical singular value decomposition representation for pattern recognition, *Neurocomputing* 214 (2017) 143–154.
 [17] W. Kong, B. Wang, Y. Lei, Technique for infrared and visible image fusion based on non-subsampled shearlet transform and spiking cortical model, *Infrared Phys. Technol.* 71 (2015) 87–98.
 [18] W. Wua, Z. Qiu, M. Zhao, Q. Huang, Y. Lei, Visible and infrared image fusion using NSST and deep Boltzmann machine, *Optik* 157 (2018) 334–342.
 [19] M. Marquez, Y. Mejia, Henry Arguello, Compressive spectral image super-resolution by using singular value decomposition, *Opt. Commun.* 404 (2017) 163–168.
 [20] Y. Li, Y. Song, Y. Zhao, S. Zhao, Songyuan Tang, An infrared target detection algorithm based on lateral inhibition and singular value decomposition, *Infrared Phys. Technol.* 85 (2017) 238–245.
 [21] X. Xu, D. Shan, G. Wang, X. Jiang, Multimodal medical image fusion using PCNN optimized by the QPSO algorithm, *Appl. Soft Comput.* 46 (2016) 588–595.
 [22] A.K. Helmy, G.S. El-Taweel, Image segmentation scheme based on SOM-PCNN in frequency domain, *Appl. Soft Comput.* 40 (2016) 405–415.
 [23] L. Tang, J. Qian, L. Li, J. Hu, X. Wu, Multimodal medical image fusion based on discrete Tchebichef moments and pulse coupled neural network, *Int. J. Imaging Syst. Technol.* 27 (1) (2017) 57–65.
 [24] L. Qiong, M. Yide, A new algorithm for noise reducing of image based on PCNN time matrix, *Electron. Inf. Technol.* 8 (2008) 1869–1873.
 [25] X. Liu, W. Mei, H. Du, Structure tensor and nonsubsampling shearlet transform based algorithm for CT and MRI image fusion, *Neurocomputing* 235 (2017) 131–139.
 [26] S. Li, X. Kang, J. Hu, Image fusion with guided filtering, *IEEE Trans. Image Process.* 22 (2013) 2864–2875 a.
 [27] J. Ma, Z. Zhou, B. Wang, H. Zong, Infrared and visible image fusion based on visual saliency map and weighted least square optimization, *Infrared Phys. Technol.* 82 (2017) 8–17.
 [28] Y. Liu, X. Chen, H. Peng, Z. Wang, Multi-focus image fusion with a deep convolutional neural network, *Inf. Fusion* 36 (2017) 191–207.
 [29] Y. Zhang, L. Zhang, X. Bai, L. Zhang, Infrared and visible image fusion through infrared feature extraction and visual information preservation, *Infrared Phys. Technol.* 83 (2017) 227–237.
 [30] J. Ma, C. Chen, C. Li, J. Huang, Infrared and visible image fusion via gradient transfer and total variation minimization, *Inf. Fusion* 31 (2016) 100–109.
 [31] X. Bai, Infrared and visible image fusion through feature extraction by morphological sequential toggle operator, *Infrared Phys. Technol.* 71 (2015) 77–86.
 [32] Y. Ma, J. Chen, C. Chen, F. Fan, J. Ma, Infrared and visible image fusion using total variation model, *Neurocomputing* 202 (2016) 12–19.
 [33] W. Kong, Y. Lei, M. Ren, Fusion method for infrared and visible images based on improved quantum theory model, *Neurocomputing* 212 (2016) 12–21.
 [34] A.K. Helmy, G.S. El-Taweel, Image segmentation scheme based on SOM-PCNN in frequency domain, *Appl. Soft Comput.* 40 (2016) 405–415.
 [35] X. Bai, Y. Zhang, F. Zhou, B. Xue, Quadtree-based multi-focus image fusion using a weighted focus-measure, *Inf. Fusion* 22 (2015) 105–118.



Boyang Cheng received his B.S. degree from the applied physics, Jilin University, Jilin, China, in 2011. He is now a Ph.D. student of optical engineering in Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His major research interests include image processing and image fusion.



Guoning Li received his B.S. degree and M.S. degree from Changchun University of Science and Technology in 2002 and 2005, respectively. He received his Ph.D. degree from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences in 2008. Currently he is a researcher in Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include image acquisition and image processing.



Longxu Jin received his B.S. degree from Changchun University of Science and Technology in 1987. He received his M.S. degree and Ph.D. degree from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences in 1993 and 2003, respectively. Currently he is a researcher in Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include aerospace payload electronic control system, image acquisition, and image compression.