Infrared and low-light-level image fusion based on ℓ_2-energy minimization and mixed-ℓ_1-gradient regularization

Boyang Cheng, Longxu Jin, Guoning Li

PII: S1350-4495(18)30536-X
DOI: https://doi.org/10.1016/j.infrared.2018.11.023
Reference: INFPHY 2776

To appear in: *Infrared Physics & Technology*

Received Date: 20 July 2018
Revised Date: 25 October 2018
Accepted Date: 26 November 2018

Please cite this article as: B. Cheng, L. Jin, G. Li, Infrared and low-light-level image fusion based on ℓ_2-energy minimization and mixed-ℓ_1-gradient regularization, *Infrared Physics & Technology* (2018), doi: https://doi.org/10.1016/j.infrared.2018.11.023

Infrared and low-light-level image fusion based on ℓ_2-energy minimization and mixed-ℓ_1-gradient regularization

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*

**Corresponding author: boyangwudi@163.com*

Abstract

In order to compensate for the visual defect of the low-light-level image and combine the saliency features of the infrared image, this paper proposes an infrared and low-light-level image fusion model based on ℓ_2-energy minimization and mixed-ℓ_1-gradient regularization. First, this novel model uses the non-subsampled shearlet transform (NSST) as a multi-scale decomposition tool to capture the low and high-frequency components of the source images. Because the NSST has good localization characteristics, excellent directional selectivity, parabolic edge characteristics, and translation invariance, it is more suitable for image decomposition and reconstruction. Secondly, for the low-frequency components that reflect the energy information, an optimization model based on ℓ_2-energy minimization is adopted as its fusion rule. This new rule allows the fused image to have similar pixel intensities to the given infrared image, thus improving the visual observation of the fused image and reducing the influence of the brightness defect under weak light. Thirdly, considering that the ℓ_1-norm encourages the sparseness of the gradients, this paper uses the ℓ_1-gradient regularization to guide the fusion of high-frequency components. This method can greatly restore the gradient features hidden in the source images to the fused image so that the fused image will have clearer edge details. In order to verify the effectiveness of the proposed algorithm, we adopted 6 × 6 independent fusion experiments. The final experimental results show that the proposed algorithm has better visual effects in the fusion problem of low-light-level environment, and the performance of objective evaluation is also good, which is better than other existing typical methods.

Keywords: image fusion; low-light-level image; NSST; ℓ_2-energy minimization; mixed-ℓ_1-gradient regularization;

1. Introduction

Image fusion technology combines images from multiple sensors into a new image that contains multiple types of information features to more accurately describe and understand the scene or target information within the image [1]. At present, image fusion has a wide range of applications in both the civil and defense fields. Among them, the fusion of visible light (VL) and infrared (IR) images is currently the most applied in relation. Under poor light conditions, the image quality of the visible light image is hardly satisfying, and the target is easily concealed [2]. However, the IR image generates an image according to the radiation of the object itself and can actively acquire the target information in the scene, so the hidden hot targets can be well displayed. Through the effective combination of these two kinds of images, we can overcome the natural defects of the visual characteristics of the human eyes under the weak light in a certain situation and extend the visual band range of the human eyes [3].

Various fusion approaches have been proposed recently, especially those aiming at pixel level-based image fusion, which can be roughly divided into the following two categories: the space domain-based method [4], and the transform domain-based method [5]. The spatial domain method directly processes the spatial pixel elements of the image, and the representative methods such as sparse decomposition (SR) [6-8], pulse-coupled neural network (PCNN) [9-11], multi-scale top hat transformation (MTH) [12-14], structure tensor [15-17] have been successfully

used to deal with the issue of image fusion. Since the spectra of IR and VL image have different wavelengths, the spectral difference between these two kinds of images is large. When the two source images are directly fused by the above methods, the large spectral differences of the pixel elements cannot be compensated, so the fusion images often have a problem of low contrast ratio. Different from the space domain-based method, the multi-scale geometric analysis (MGA) [18-23] is used as the principal research tool in the transform domain method. It is worth mentioning that the non-subsampled shearlet transform (NSST) has the best fusion effect among these MGA tools. Because of its simple mathematical structure, good direction selectivity and anisotropy, and removing the sampling operation during the decomposition process, it has received extensive attention from scholars [24]. In view of its various advantages, this paper uses NSST as an image multi-scale decomposition tool. The NSST can decompose a source image into a series of sub-band images with different scales and frequencies, thereby laying the foundation for subsequent fusion rules. The most direct method for applying to fusion rules is the weighted-average method [25], which simply spatially superimposes the brightness of the source image, leading to a decrease in the resolution of the final fused image and losing a number of texture details information. In order to obtain better fusion effects, scholars [26-28] have successively utilized some spatial domain methods as the fusion rules for NSST to compensate for the defects of the weighted average method. Therefore, it is very meaningful to study fusion algorithms based on the NSST method.

In general, when the existing spatial domain methods are used as the fusion rules for the NSST method, most of their applicable scenarios are directed to the visible light environment with moderate illumination. However, for the special fusion of low-light-level (LL) and infrared (IR) images, the existing research methods are relatively few. First, the IR image is a kind of image that reflects the energy distribution and does not change greatly as the brightness of the external light changes. Therefore, when a LL image is fused with an IR image, the brightness of the final fused image is generally greatly reduced, which affects the visual perception. In order to solve the low brightness disadvantage brought by LL image, the fused image should have similar pixel intensities with the given IR image. In addition, the gradient distribution in the two kinds of source images should be also transferred into the fused image. As we all know, the low-frequency sub-image contains the energy information, and its fusion rules also determine the final fusion visual effect. In order to make use of the brightness advantage of the IR image to make up for the visual defect under the weak light, this paper proposes an optimization model based on ℓ_2-energy minimization as the fusion rule of low-frequency sub-image. This optimization model uses the ℓ_2-norm [29] to characterize the energy information of the source image, which ensures to a certain degree that the fused image can meet the sensitive observation range of human eyes. In addition, the high-frequency components of the image highlight the feature of texture gradient, so this paper proposes an optimization model based on the ℓ_1-gradient regularization as its fusion rule for high-frequency sub-images. The model uses ℓ_1-norm [30] to encourage the sparseness of the gradients, so the gradient information of the source image is restored to the fused image. The above two new fusion rules will be discussed in detail in the third chapter.

On the basis of the preceding review, this paper proposes a fusion model based on ℓ_2-energy minimization and mixed-ℓ_1-gradient regularization for the fusion of low-light-level and infrared images. Firstly, the NSST is adopted as the MGA tool for the decomposition of infrared and low-light-level images. Secondly, we use the ℓ_2-energy minimization and the ℓ_1-gradient regularization as the fusion rules of the low-frequency and high-frequency sub-image, respectively. In short, the fusion rules are turned into optimization problems, namely energy minimization and gradient regularization. Because this model contains only one regularization term parameter, it has better self-adaptability. As far as we know, such a fusion method has not been studied yet.

The rest of the proposed paper is structured as follows. The relevant theory of the NSST and gradient descent is shown in Chapter 2. The proposed algorithm based on the new fusion rule is

presented in Chapter 3. The part of experimental results and analysis is shown in Chapter 4. The summary of this paper is given in Chapter 5.

2. Relevant Theory

2.1. Non-subsampled shearlet transform [31-32]

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

$$A_{AB}(y)=\{y_{j,l,k}=\left|\det A\right|^{j/2} y\left(B^{l}A^{j}x-k\right);\ j,l\in\mathbb{Z},k\in\mathbb{Z}^{2}\}. \tag{1}$$

where $y\in L^2(R^2)$; A and B are 2×2 reversible matrices; j and l represent any integer, and k represents the number set of any integer square; A^j is associated with scale transformation, and B^l is associated with geometric transformation; $|\det B|=1$. If $A_{AB}(y)$ has a Parseval framework as the following, the elements of this system are called synthetic wavelets, i.e., for any $f\in L^2(R^2)$:

$$\sum_{j,l,k}\left|\left\langle f,y_{j,l,k}\right\rangle\right|^{2}=\left\|f\right\|^{2}. \tag{2}$$

When $A_a=\begin{bmatrix}a & 0\\ 0 & \sqrt{a}\end{bmatrix}$, and $B_s=\begin{bmatrix}1 & s\\ s & 1\end{bmatrix}$, the following system is generated as the shearlet system:

$$\{y_{ask}(x)=a^{-\frac{3}{4}}y\left(A_a^{-1}B_s^{-1}x-k\right),a\in\mathbb{R}^{+},s\in\mathbb{R},k\in\mathbb{R}^{2}\}. \tag{3}$$

where a denotes any positive real number, s denotes any real number, k is the number set of any real square, and $y_{ask}(x)$ is the shearlet.

The traditional shearlet transform has a down-sampling operation, so the Gibbs effect is easily generated during image fusion, which affects the final fusion effect. To this end, the non-subsampled shearlet transform (NSST) is proposed on the basis of shearlet transform, which is mainly divided into two steps:

Multi-scale decomposition: Since there is no down-sampling operation in the NSST, the image f is decomposed by a non-sampling pyramid (NSP), and finally k+1 sub-band images of the same size as f are obtained.

Directional localization: The NSST uses shearlet filters to realize the directional localization of high-frequency images, and the specific implementation process is as follows:

1) Mapping the pseudo-polarization coordinates to the Cartesian coordinates.

2) Generating the shearlet filter by using the "Meyer" window function.

3) Convolution operation is performed on the k band-pass sub-band images and the "Meyer" window function. The decomposition framework NSST is shown in Fig. 1.

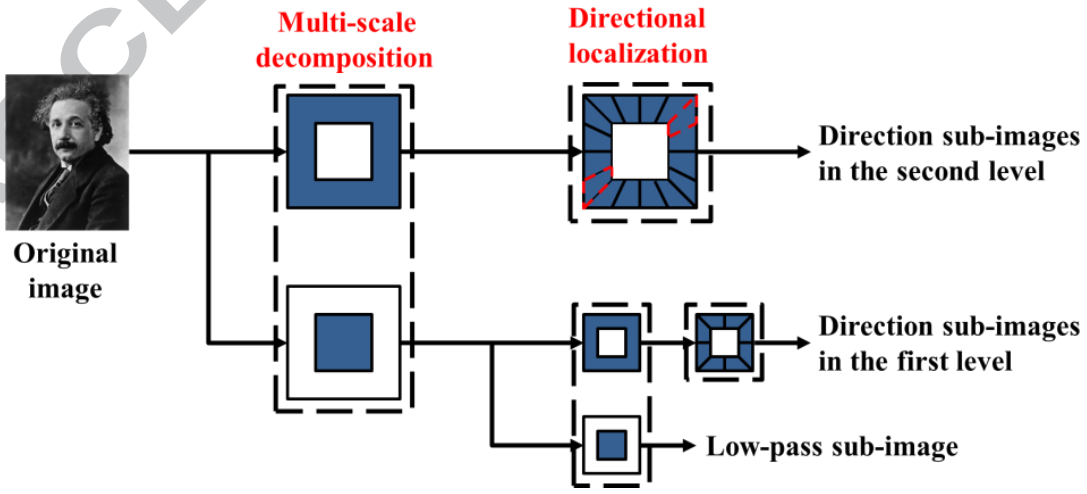


Fig. 1. The decomposition framework of NSST

2.2. Gradient Descent

The gradient descent method [33] was used to solve the unconstrained optimization problem. For a binary function $u(x, y)$, consider the extremum problem of its functional analysis, the

expression is as follows:

$$J[u(x,y)] = \iint_D F(x,y,u(x,y),u_x(x,y),u_y(x,y))dxdy \tag{4}$$

where $J[u(x, y)]$ is called the energy functional of $u(x, y)$, F is a function of three-differentiated, and $u(x, y)$ requires a second-order continuous partial derivative within the closed area D of the plane. Perform variational operations on $J[u(x, y)]$, and we can get the following expression:

$$\begin{cases} \delta J = \dfrac{\partial}{\partial \alpha} J(u(x,y,\alpha))\big|_{\alpha=0} = \iint_D (F_u \delta u + F_{u_x}\delta u_x + F_{u_y}\delta u_y)dxdy \\ F_u = \dfrac{\partial F}{\partial u}, F_{u_x} = \dfrac{\partial F}{\partial u_x}, F_{u_y} = \dfrac{\partial F}{\partial u_y}. \end{cases} \tag{5}$$

due to:

$$\begin{cases} \dfrac{\partial}{\partial x}(F_{u_x}\delta u) = F_{u_x}\delta u_x + \dfrac{\partial}{\partial x}(F_{u_x})\delta u \\ \dfrac{\partial}{\partial y}(F_{u_y}\delta u) = F_{u_y}\delta u_y + \dfrac{\partial}{\partial y}(F_{u_y})\delta u \end{cases} . \tag{6}$$

put the Eq. (6) into the Eq. (5), and we can get that:

$$\delta J = \iint_D \left(F_u \delta u - \frac{\partial}{\partial x}(F_{u_x})\delta u - \frac{\partial}{\partial y}(F_{u_y})\delta u + \frac{\partial}{\partial x}(F_{u_x}\delta u) + \frac{\partial}{\partial y}(F_{u_y}\delta u) \right) dxdy . \tag{7}$$

use the **Green formula**:

$$\begin{aligned} &\iint_D \left(\frac{\partial}{\partial x}(F_{u_x}\delta u) + \frac{\partial}{\partial y}(F_{u_y}\delta u) \right) dxdy \\ &= \int_{\partial D} \left(F_{u_x} dy - F_{u_y} dx \right) \delta u d\delta D = 0 \end{aligned} . \tag{8}$$

because of the δu=0 at the boundary, Eq. (7) becomes as the following:

$$\delta J = \iint_D \left(F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}) \right) \delta u dxdy . \tag{9}$$

the necessary condition for the extreme value of the energy functional is δJ=0, and due to the arbitrariness of δu, the following formula can be obtained:

$$F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}) = 0 . \tag{10}$$

where the Eq. (10) is the ***Euler-Lagrange*** equation. From the above analysis, we can see that the unconstrained optimization problem is usually converted into the Euler equation solution problem. However, it is very difficult to directly solve Euler equations, and the most commonly used method is the ***Gradient Descent***.

The basic idea of the Gradient Descent method is to start from the initial value u_0 and finally reach the local minimum point of $J(u)$ along the opposite direction of the gradient direction. Eq. (10) can be viewed as the gradient of Eq. (4), which is expressed as follows:

$$\begin{cases} \dfrac{\partial u}{\partial t} = -F_u - \dfrac{\partial}{\partial x}(F_{u_x}) - \dfrac{\partial}{\partial y}(F_{u_y}) \\ u(x,y,0) = u_0(x,y) \end{cases} . \tag{11}$$

where the time auxiliary parameter t is introduced in the Eq. (11) to represent the evolution of convex optimization equation. The parameter t represents the process of the continuous approximation of the function $u(x, y)$. When $\partial u/\partial t$=0, the Euler equation solution to the unconstrained optimization problem is obtained. Proceeding from the initial function $u_0(x, y)$, we iterate the calculation of Eq. (11) until the steady state is achieved. The termination condition of the iteration can be determined if the number of iterations exceeds a certain set value or the condition below is satisfied:

$$\frac{\sqrt{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(u^{n+1}(x,y)-u^{n}(x,y)\right)^{2}}}{M\times N}<\varepsilon. \tag{12}$$

where n and n+1 denote the t and t+1 moments respectively, ε is a small positive number, and $M\times N$ is the size of the original image. If the above conditions are satisfied, it means that the iteration reaches a steady state, so the iterative process can be ended.

3. Fusion Rules

The fusion rules for the low and high-frequency components of the image will be discussed in detail in Section 3.1 to Section 3.3, and the innovations of this paper are summarized as follows:

1. This paper combines the idea of image feature optimization with multi-scale decomposition to propose the IR and LL image fusion based on hybrid-norm optimization and NSST.

2. In order to improve the overall brightness of the fused image and compensate for the brightness defect of the LL image, the low-frequency component of the image adopts an optimization model based on the ℓ_2-energy minimization as its fusion rule.

3. The high-frequency components of the image are guided by an optimization model based on ℓ_1-gradient regularization, which can greatly restore the gradient features inside the two kinds of source images.

Through the combination of the above two optimization models, the feature components contained in the source images can be transferred to the final fused image. The Schematic of the proposed fusion model is shown in Fig. 2.

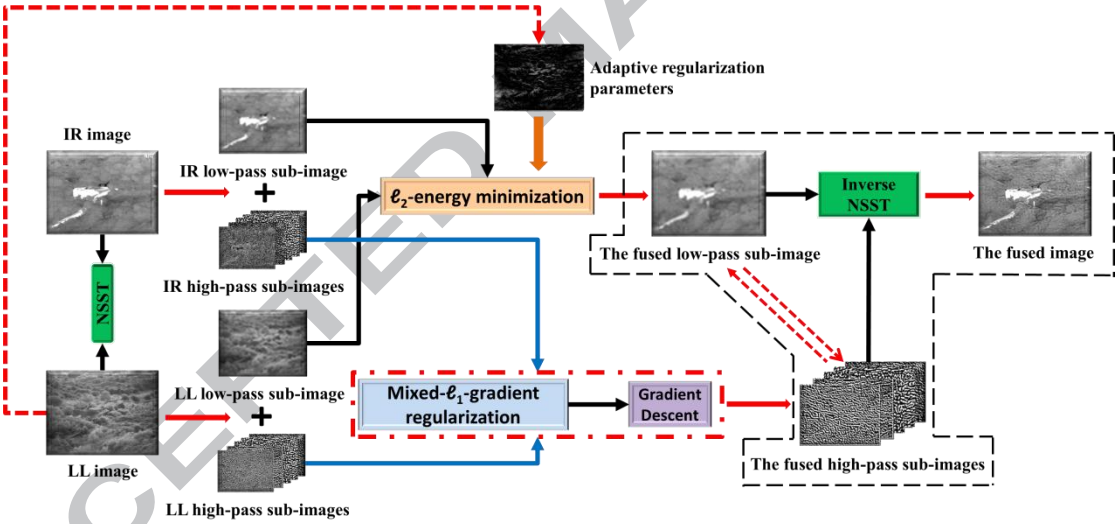


Fig.2. The Schematic of the proposed fusion model

3.1. Fusion rule of low-frequency sub-images

In order to take advantage of the brightness advantages of the IR image to compensate for visual imperfections under the weak light, we hope that the fused image should have similar pixel intensities as the given IR source image. Since the ℓ_2-norm can represent the energy information of the image, this paper proposes an optimization model based on the ℓ_2-energy minimization, and its expression is shown as follows:

$$\arg\min\left\{\left\|C_{l,k}(i,j)-C_{l,k}^{IR}(i,j)\right\|_2^2+\lambda\cdot\left\|C_{l,k}(i,j)-C_{l,k}^{LL}(i,j)\right\|_2^2\right\}. \tag{13}$$

where $C^{IR}_{l,k}(i,j)$, $C^{LL}_{l,k}(i,j)$ and $C_{l,k}(i,j)$ denote the low-frequency coefficients of the IR image, LL image, and fusion image, respectively, l represents the decomposition number, and k represents the direction of decomposition of each layer. In addition, λ represents the regularization parameter, and $\|\ \|_2$ denotes the ℓ_2-norm. The first term of the model represents the data fidelity, which ensures that the low-frequency components of the fused image are maximally close to the low-frequency components of the IR image. At the same time, the low-frequency coefficients of

the low-light-level image are added to the regularization terms of this model, which ensures to a certain degree that the fused image can meet the sensitive observation range of human eyes. The ***Euler-Lagrange*** equation of Eq. (13) can be calculated by Eq. (10):

$$\left(C_{l,k}(i,j)-C_{l,k}^{IR}(i,j)\right)+\lambda\cdot\left(C_{l,k}(i,j)-C_{l,k}^{LL}(i,j)\right)=0. \tag{14}$$

since there is no gradient operator inside the above equation, a corresponding linear solution can be obtained. In order to express the contribution of the source images to the fused image, we propose the following expression:

$$C_{l,k}(i,j)=\omega_1 C_{l,k}^{IR}(i,j)+\omega_2 C_{l,k}^{LL}(i,j). \tag{15}$$

where ω_1 and ω_2 denote the weight parameters, which represent the contribution of the source images to the fused image. Combine the Eq. (14) with the Eq. (15), the solution to the Eq. (13) is that:

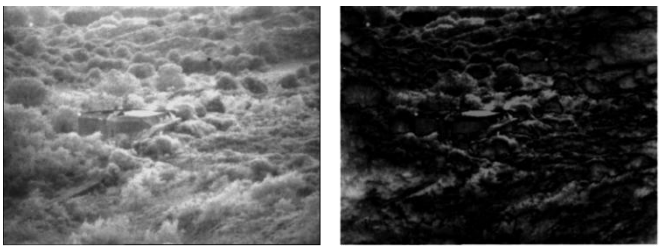
$$\begin{cases} C_{l,k}(i,j)=\dfrac{1}{1+\lambda}C_{l,k}^{IR}(i,j)+\dfrac{\lambda}{1+\lambda}C_{l,k}^{LL}(i,j) \\ \omega_1=\dfrac{1}{1+\lambda},\ \omega_2=\dfrac{\lambda}{1+\lambda} \end{cases}. \tag{16}$$

by the above method, the regularization parameter λ is put into the weight parameter and the final fusion effect can be determined by controlling the size of λ. When λ is a fixed constant, Eq. (13) is transformed into the weighted average method. Because this kind of fusion rule cannot take into account the difference of the two spectra, it will cause the contrast ratio of the fused image to be reduced to a certain extent and affect the final perception. This obviously contradicts the original intention of this paper. For this reason, λ should also be a function that can change according to the change in the gray value of the source image. At the same time, in view of the role of λ as a correction parameter, it can also characterize the spatial information distribution of the LL image, so that the subjective sense of the fused image is closer to the IR image, which makes up for the lack of brightness of the LL image. For this reason, this paper proposes a novel adaptive regularization parameter, and its expression is shown as follows:

$$\lambda(i,j)=\sum_{p=0}^{L-1}M_p\left|LL(i,j)-LL_p\right|. \tag{17}$$

where $LL(i,j)$ represents the intensity of the pixel (i,j) in the LL original image, the value of LL_p is in the range of [0, 255], p represents the pixel intensity, that is $LL_p = p$, M_p represents the number of pixels whose gray-value is equal to LL_p, and L is the number of gray levels (in this paper it is 256).

An example of the visual effect of λ is shown in Fig. 3, where Fig. 3(a) is the original LL image and Fig. 3(b) is the corresponding effect diagram. λ is changed from the originally fixed constant to an adaptive function, which solves the problem of an adaptive setting of peripheral parameters to some extent. At the same time, λ can extract the area that is relatively more important or sensitive to human eyes in an image and compensates for the data fidelity term of ℓ_2-energy minimization.



(a) **(b)**

Fig. 3. The visual effect of the adaptive regularization parameter

3.2. Fusion rule of high-pass sub-images

The high-frequency components of the image reflect the edge and texture features, and its

fusion rules determine the final level of detail. In view of the fact that the ℓ_1-norm encourages the sparseness of the gradients, this paper proposes a fusion rule based on mixed-ℓ_1-gradient regularization, and its expression is as follows:

$$\arg\min\left\{\left[\frac{1}{2}\left\|D_{l,k}(i,j)-D_{l,k}^{LL}(i,j)\right\|_2^2+\alpha\cdot\left\|\nabla D_{l,k}(i,j)-\nabla D_{l,k}^{IR}(i,j)\right\|_1\right]^p \times\left[\frac{1}{2}\left\|D_{l,k}(i,j)-D_{l,k}^{IR}(i,j)\right\|_2^2+\beta\cdot\left\|\nabla D_{l,k}(i,j)-\nabla D_{l,k}^{LL}(i,j)\right\|_1\right]^{1-p}\right\} \tag{18}$$

$$\left\|\nabla D_{l,k}-\nabla D_{l,k}^{IR}\right\|_1\approx\sum_{i,j}\sqrt{\left(\nabla_i D_{l,k}-\nabla_i D_{l,k}^{IR}\right)^2+\left(\nabla_j D_{l,k}^{IR}-\nabla_j D_{l,k}^{IR}\right)^2} \tag{19}$$

$$\left\|\nabla D_{l,k}-\nabla D_{l,k}^{IR}\right\|_1\approx\sum_{i,j}\sqrt{\left(\nabla_i D_{l,k}-\nabla_i D_{l,k}^{IR}\right)^2+\left(\nabla_j D_{l,k}^{IR}-\nabla_j D_{l,k}^{IR}\right)^2} \tag{20}$$

where $D^{IR}{}_{l,k}(i,j)$, $D^{LL}{}_{l,k}(i,j)$ and $D_{l,k}(i,j)$ are the representative of the high-frequency sub-coefficients of the IR image, LL image, and fusion image, respectively, l represents the decomposition number, k represents the direction of decomposition of each layer, α and β represent the regularization parameter, $\|\ \|_2$ denotes the ℓ_2-norm, $\|\ \|_1$ denotes the ℓ_1-norm and ∇ is the gradient operator. In addition, we call the parameter p as the gradient factor, which measures the importance of the two high-frequency components in spatial distribution. Its expression is as follows:

$$p=\begin{cases}1 & \left|D_{l,k}^{LL}(i,j)\right|\geq\left|D_{l,k}^{IR}(i,j)\right|\\ 0 & \left|D_{l,k}^{LL}(i,j)\right|\leq\left|D_{l,k}^{IR}(i,j)\right|\end{cases} \tag{21}$$

where p=1 means that at the pixel (i, j), the high-frequency sub-coefficients of the LL image can highlight the final gradient feature of the fused image. For this reason, the high-frequency sub-coefficients of the IR image should play a corrective and supplemental role, so it is placed in the regularization term of the model. Then, the ***Euler-Lagrange*** equation of Eq. (18) is changed into:

$$\left(D_{l,k}(i,j)-D_{l,k}^{LL}(i,j)\right)+\alpha\cdot div\left(\frac{\nabla D_{l,k}(i,j)-\nabla D_{l,k}^{IR}(i,j)}{\left|\nabla D_{l,k}(i,j)-\nabla D_{l,k}^{IR}(i,j)\right|}\right)=0 \tag{22}$$

where *div* represents the divergence operator. Since the direct solution of the above equation is very difficult, we use the gradient descent method mentioned in section 2.2 to solve it. The process is as follows:

Step 1: Introduce the time variable parameter t, and the gradient descending flow corresponding to the Eq. (20) is that:

$$\frac{\partial D_{l,k}(i,j)}{\partial t}=-\left(D_{l,k}(i,j)-D_{l,k}^{LL}(i,j)\right)+\alpha\cdot div\left(\frac{\nabla D_{l,k}(i,j)-\nabla D_{l,k}^{IR}(i,j)}{\left|\nabla D_{l,k}(i,j)-\nabla D_{l,k}^{IR}(i,j)\right|}\right) \tag{23}$$

Step 2: Set the initialization of the equation parameters and use edge repeat expansion processing for the boundary conditions:

$$\begin{cases}D_{l,k}(i,j,0)=D_{l,k}^{LL}(i,j)\\ D_{l,k}^{n+1}(i,j)=D_{l,k}^{n}(i,j)+\Delta t\cdot\dfrac{\partial D_{l,k}^{n}(i,j)}{\partial t}\\ \partial D_{l,k}(i,j)/\partial n=0\end{cases} \tag{24}$$

where n is the number of iterations, and Δt is the time step.

Step 3: Bring the Eq. (23) into the Eq. (24). The stationary solution of $D_{l,k}(i,j)$ is obtained by time difference iteration:

$$D_{l,k}^{n+1} = D_{l,k}^{n}(i,j) - \Delta t \cdot \left(D_{l,k}^{n}(i,j) - D_{l,k}^{LL}(i,j) \right) + \alpha \cdot \Delta t \cdot div \left(\frac{\nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{IR}(i,j)}{\left| \nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{IR}(i,j) \right|} \right) \tag{25}$$

Step 4: when p=0, it indicates that the IR image contains more gradient information at this time. Repeat step 1 to step 3, then the stationary solution of $D_{l,k}(i,j)$ will be changed into:

$$D_{l,k}^{n+1} = D_{l,k}^{n}(i,j) - \Delta t \cdot \left(D_{l,k}^{n}(i,j) - D_{l,k}^{IR}(i,j) \right) + \alpha \cdot \Delta t \cdot div \left(\frac{\nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{LL}(i,j)}{\left| \nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{LL}(i,j) \right|} \right) \tag{26}$$

Step 5: Combine the Eq. (28) with the Eq. (29) and introduce the gradient factor p to obtain the final fused high-frequency sub-coefficients $D_{l,k}(i,j)$ as:

$$\begin{aligned} D_{l,k}^{n+1} = {} & D_{l,k}^{n}(i,j) - \Delta t \cdot \left(D_{l,k}^{n}(i,j) - D_{l,k}^{LL}(i,j) \right)^{p} \cdot \left(D_{l,k}^{n}(i,j) - D_{l,k}^{IR}(i,j) \right)^{1-p} \\ & + \alpha \cdot \Delta t \cdot div \left(\frac{\nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{IR}(i,j)}{\left| \nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{IR}(i,j) \right|} \right)^{p} \cdot div \left(\frac{\nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{LL}(i,j)}{\left| \nabla D_{l,k}^{n}(i,j) - \nabla D_{l,k}^{LL}(i,j) \right|} \right)^{1-p} \end{aligned} \tag{27}$$

Step 6: Use the Eq. (12) as the termination criterion of the iteration together with the constraint of iteration number, and its expression is as follows:

$$\frac{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(D_{l,k}^{n+1}(i,i) - D_{l,k}^{n}(i,j) \right)^{2}}}{M \times N} < \varepsilon \tag{28}$$

3.3. The acquisition of fused images

Through the construction of Section 3.1 to Section 3.2, we finally obtain the fused low-frequency and high-frequency sub-coefficients $C_{l,k}(i,j)$ and $D_{l,k}(i,j)$ of the image. Then, these two kinds of coefficients are reconstructed by inverse NSST, and the fused image will be obtained.

4. Experimental Results and Analysis

To verify the superiority of the proposed method, our method is compared with the following methods: NSST[34], GFF[35], HMSD[36], IFE[37], GTF[38]. To show the adaptivity of the algorithm, we tested it on six different picture groups presented in Fig. 4(a)-(f). The first group shows soldiers guns, the second group depicts the street at night, the third shows a bunker under the dark light, the fourth shows a soldier hidden behind the smoke, the fifth depicts a figure in the gully, and the last group depicts a pedestrian in the forest. The visible light images of the above six groups of experiments are all under low-light-level conditions, which are slightly different from the applicable scenes of the traditional infrared and visible light image fusion algorithms. The original image sizes of the six groups are 917×678, 632×496, 768×576, 768×576, 768×576, 360×270。The experimental parameters of each algorithm are set as follows:

In the NSST method, the level of multi-scale decompositions is [2, 4, 4]. The size of the shearing window is set to be 3. Δ and V_{θ} are set to be 0.01 and 10, respectively. W_{ijkl} is [0.1035, 0.1465, 0.1035; 0.1465, 0.0000, 0.1465; 0.1035, 0.1465, 0.1035].

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

In the HMSD method, we use a four-level decomposition of the hybrid-MSD, in which we set $\sigma_{s,0}$=2, $\sigma_{r,0}$=0.1, λ=30.

In the IFE method, we set the parameters as: QuadNormDim = 512; QuadMinDim = 32; GaussScale = 9; MaxRatio = 0.001; StdRatio = 0.5.

In the GTE method, we fix λ to 4 as an empirical value.

Our method takes "maxflat" as the pyramid filter of NSST. The number of decomposition layers is four, that is, [2, 2, 3, 3], and the window size of the shearlet filter is [15, 15, 15, 15]. With the above empirical parameters setting, the NSST can often obtain better decomposition effect. In addition, we adopt α=β=0.6, ε=5×10^{-4}, and Δt=5×10^{-3}. The source images of each group of infrared and low-light-level are strictly registered, and they can be downloaded from the site:

https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029.All experiments are conducted with MATLAB 2012a programming using a PC with Intel Core i7/3.4 GHz/4G processor.

(a) (b) (c) (d) (e) (f)

Fig. 4. Source images

4.1. Subjective evaluation

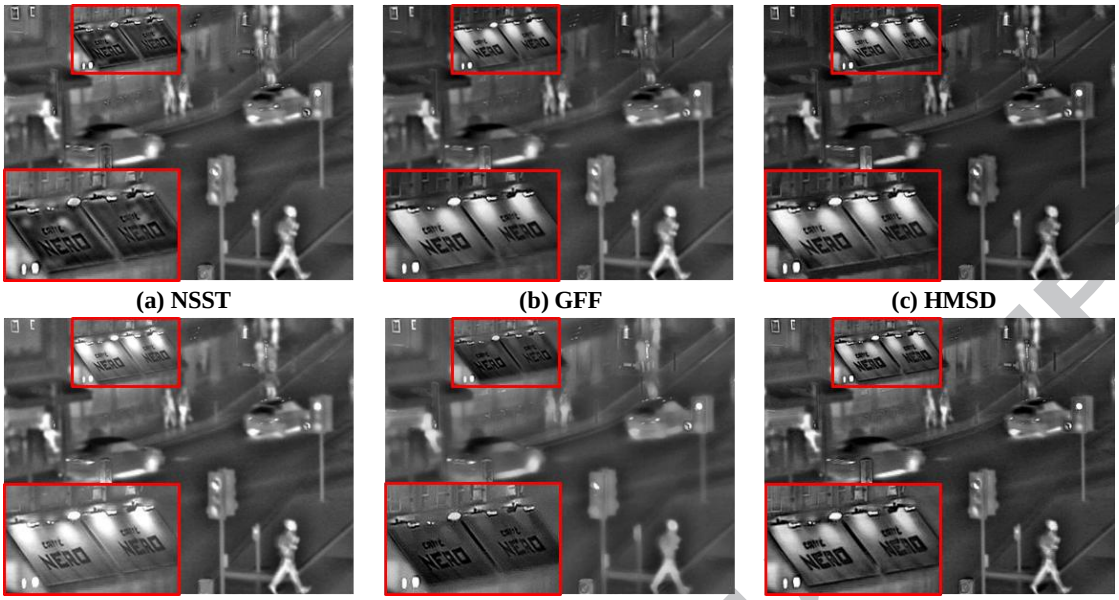
The NSST, GFF, HMSD, IFE, GTF methods, and the proposed method are shown from Figs. 5(a) ~ 5(f) to Fig. 10(a) ~ 10(f). We magnify the details of each picture so that it is easier for observers to compare the visual effects between the above algorithms. In the first group of experiments, the brightness of all fused images basically reach the human eye's observation range, so these fusion algorithms can compensate for the weakness of the low light to some extent. However, after zooming in on the details, the NSST method has artifact noise in the cloud, resulting in an unnatural transition of the image gray-value. The cloud details of the GFF, HMSD, IFE, and GTF methods are smoothed obviously, which are far less clear than the proposed algorithm. This is partly due to the mixed-ℓ_1-gradient regularization, which can restore the gradient features of the source images to the greatest extent.

(a) NSST **(b) GFF** **(c) HMSD**

(e) IFE **(f) GTF** **(h) Proposed**

Fig. 5. The first group of fusion experiment.

In the second set of experiments, we directly compared the enlarged details. The fusion details of the NSST and GTF methods are dim and the fonts are vague, which are difficult to distinguish. The fusion image of the IFE method is too bright and exceeds the optimal viewing range of human eyes. The fusion details of the GFF method are similar to those in Fig. 6(h), but the font outline of the proposed algorithm seems clearer. The overall brightness of the HMSD method is not as good as other algorithms, and its details appear uneven black noise at the bottom. Overall, the proposed algorithm has a better visual perception while the edge texture is still clear.

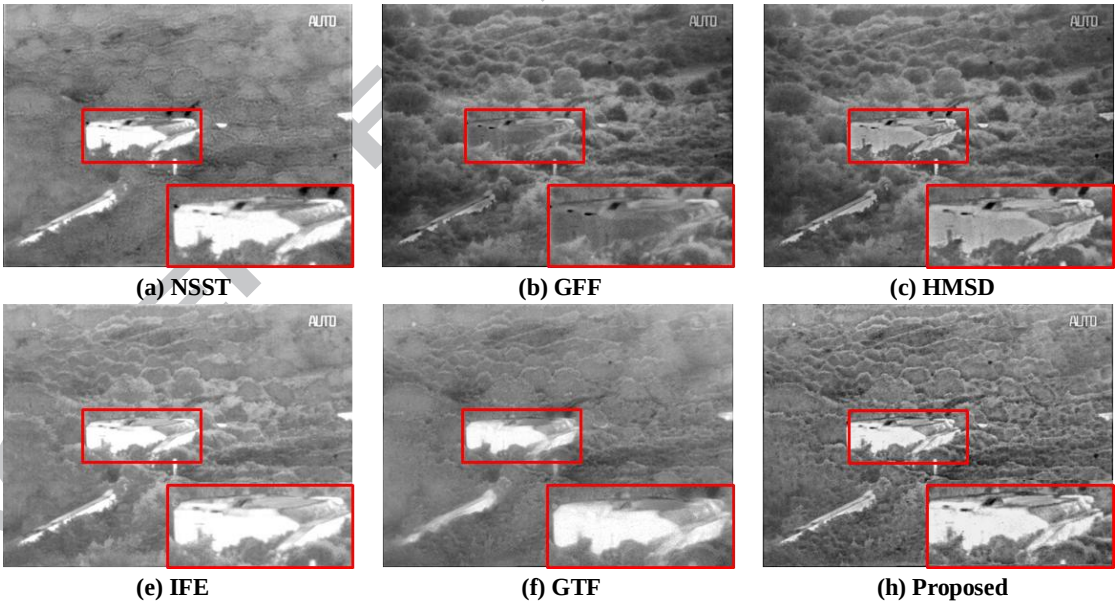


(a) NSST (b) GFF (c) HMSD

(e) IFE (f) GTF (h) Proposed

Fig. 6. The second group of fusion experiment.

In the third group of experiments, the image brightness of the GFF and HMSD methods are still low, so these fusion algorithms do not compensate for the disadvantages of the low-light-level image. The brightness of the IFE and GTF methods is indeed very high, but maybe overexposed, resulting in blurred texture details of the background. The NSST method has a similar look and feel to the method in this paper, but the gradient characteristics of the trees are far less clear as shown in Fig. 7(h).

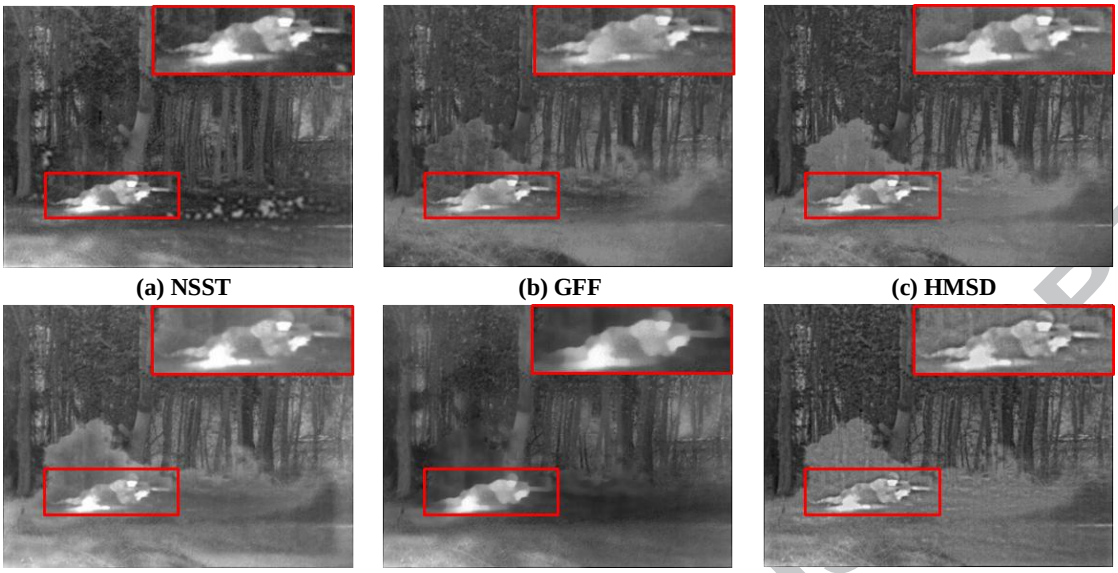


(a) NSST (b) GFF (c) HMSD

(e) IFE (f) GTF (h) Proposed

Fig. 7. The third group of fusion experiment

In the fourth group of experiments, the soldiers' edges of the NSST and GTF methods are more pronounced, but the smoke information of the LL image is lost. Although this still achieves the effect of target recognition, it does not match the idea of fusion. The IFE method does not show overexposure in this group of experiments, but compared to the remaining three contrast algorithms, the soldier's outline is unclear and the texture is lost. The GFF and HMSD methods have a similar look and feel to the Fig. 8(h), indicating that the above two fusion methods are excellent and feasible in this kind of scene. However, comparing the tree textures of the image background, the trees of the proposed algorithm are still relatively dense.

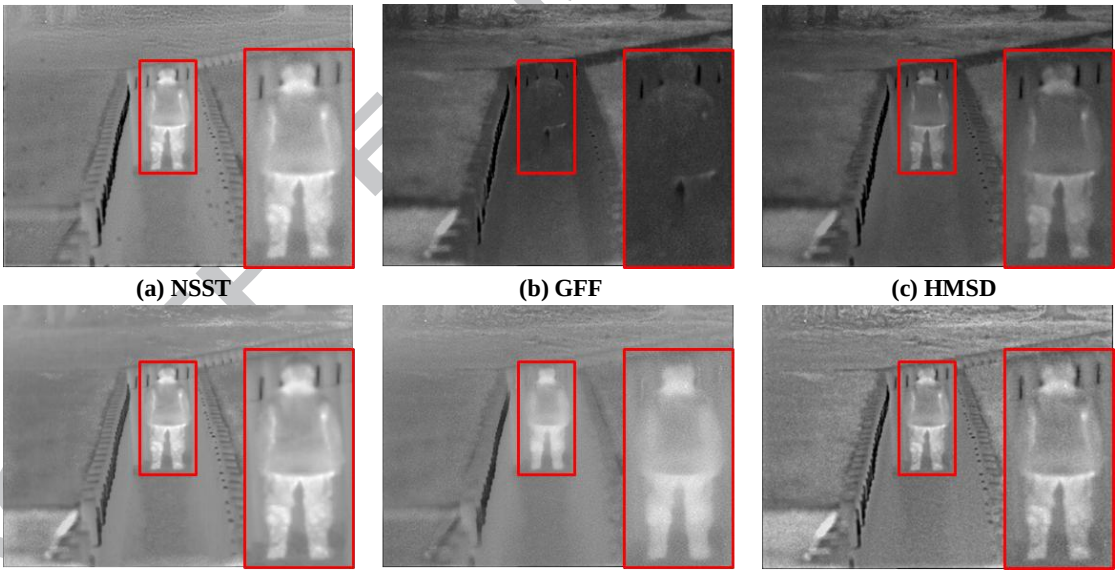


(a) NSST (b) GFF (c) HMSD

(e) IFE (f) GTF (h) Proposed

Fig. 8. The fourth group of fusion experiment.

In the fifth group of experiments, the visibility of the GFF and HMSD methods is lower, indicating that the low-light-level disadvantage of the fused image still exists. The figure in the GTF-fused image is vague and difficult to distinguish. The fused image brightness of the NSST and IFE methods is moderate, but their perception is too close to the IR source image, thus losing the texture detail information in the LL image. In contrast, the figure in Fig. 9(h) is the most legible and the gradient features are also most pronounced.

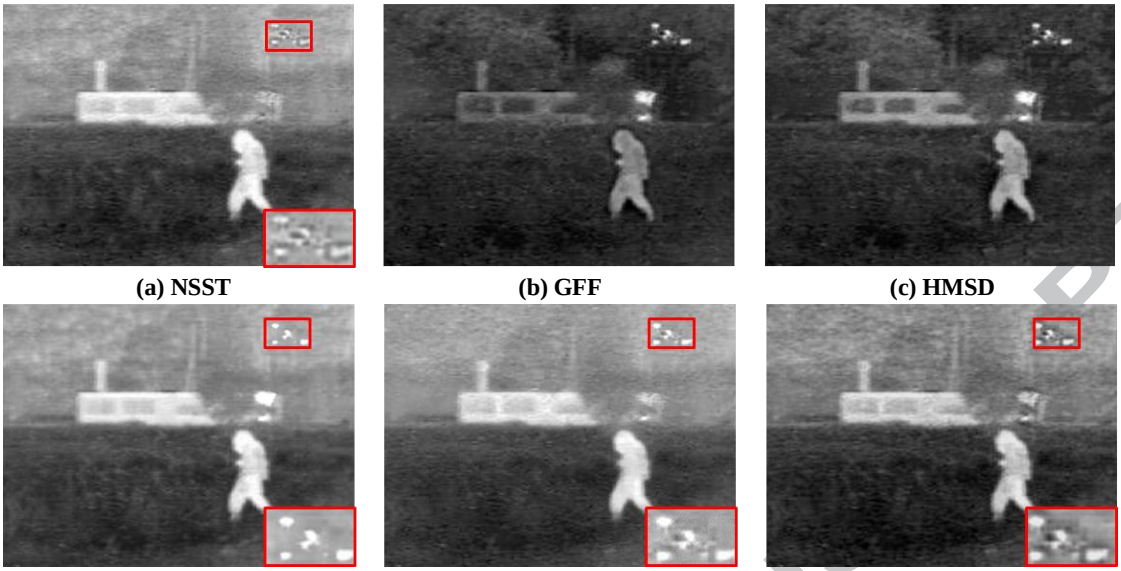


(a) NSST (b) GFF (c) HMSD

(e) IFE (f) GTF (h) Proposed

Fig. 9. The fifth group of fusion experiment.

In the sixth group of experiments, the fusion images of the GFF and HMSD methods are lower in brightness than other images. Although they also accord with the human eyes' perception, they do not exert the maximum brightness advantage of the IR source image. The Figs. 10(a), (e)-(h) have similar look and feel with the fusion image of the proposed algorithm. But after zooming in on the details, it is found that the bright point of Fig. 10(h) is the clearest, and there is no shortcoming of brightness overexposure.



(a) NSST (b) GFF (c) HMSD

(e) IFE (f) GTF (h) Proposed

Fig. 10. The sixth group of fusion experiment.

In summary, the above comparison algorithms can better deal with the fusion problem and have achieved certain academic attainments. However, for these special images fusion of the low-light-level and infrared spectrums, the fusion visual effect of the proposed algorithm seems to be more in line with the human eyes. The "ℓ_2-energy minimization" that we take at low-frequency components will maximize the brightness of the fused image while avoiding overexposure defects; at the same time, the mixed ℓ_1-gradient regularization can more comprehensively add the gradient features inside the source image to the fused images. Therefore, the brightness of the final fused image is in line with the human vision and has finer texture detail information. Through the above description, the new fusion method proposed in this paper has less peripheral parameters, moderate brightness, high contrast, and best subjective perception.

4.2. Objective evaluation

In general, the performance of image fusion results can be evaluated in a subjective and objective way. In most cases, the difference between the fusion results is small, and it is difficult to correctly evaluate the fusion results in a subjective way. Therefore, it is necessary to evaluate the fusion effect on the basis of objective quality evaluation. This paper selects the following eight objective quality indicators as the evaluation criteria:

1. Average Gradient (AVG) [39]: The AVG value reflects the clarity of the image.

2. Information Entropy (IE) [40]: The IE value reflects the richness of the information in the image.

3. Edge Information Preservation Value ($Q^{AB/F}$) [41]: The $Q^{AB/F}$ value is used to describe the fused image to preserve the visually important edge intensity information and direction information in each source image.

4. Spatial Frequency (SF) [42]: The SF value is used to reflect the overall activity of the image in the spatial domain.

5. Standard Deviation (SD) [43]: The SD value is used to reflect the distribution of pixel gray-values.

6. Gradient-based fusion metric (Q^G) [44]: The Q^G value evaluates the amount of gradient information transferred from LL and IR images to the fused images.

7. Mutual Information (MI) [45]: The MI value of an image reflects the degree of correlation between the two images.

For the above seven indicators, the larger the value, the better the performance of the fusion algorithm, and the larger the value, the more the fused image contains the source images information. However, when the brightness of the image is overexposed or there is artifact noise,

these parameters will be artificially high and become unreal. Therefore, the objective parameters also need to be evaluated together with subjective perception. The detailed quantitative evaluation of Figs. 5-10 is demonstrated in Tables 1-3. The values in boldface represent the best results.

Group	Fusion methods	Evaluation index							
		AVG	IE	$Q^{AB/F}$	SF	SD	Q^{G}	MI	Time/s
1	NSST	2.8004	7.4459	0.3632	6.3537	49.1421	0.0646	5.4140	105.8
	GFF	2.4423	7.2676	0.3124	5.9182	43.2459	0.6723	3.7476	6.34
	HMSD	2.5168	6.7592	0.4671	5.7548	30.4111	0.5517	1.5337	14.48
	IFE	2.4734	7.3563	0.3081	5.8852	45.4516	0.6721	**5.7207**	0.77
	GTF	2.2992	7.5368	0.3150	6.1268	38.0316	0.2805	4.6863	15.55
	Proposed	**3.0443**	**7.5396**	**0.5083**	**6.7040**	**49.6053**	**0.6883**	5.0234	11.75
2	NSST	4.4854	6.7408	0.3947	12.2818	33.9181	0.5975	4.1492	53.15
	GFF	4.1166	6.5331	0.3250	12.1622	32.4020	0.6690	1.1995	3.50
	HMSD	4.5604	6.4007	0.4141	13.1399	33.4755	0.5920	1.1573	9.26
	IFE	4.0015	**6.7250**	0.3037	11.0240	**36.8924**	0.6800	**6.3218**	0.19
	GTF	3.4195	6.6624	0.4017	8.5376	32.9902	0.3706	2.7908	7.96
	Proposed	**5.3176**	6.6728	**0.4936**	**14.3709**	33.4962	**0.7001**	2.6782	6.02

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

Group	Fusion methods	Evaluation index							
		AVG	IE	$Q^{AB/F}$	SF	SD	Q^{G}	MI	Time/s
3	NSST	4.7297	6.6296	0.4610	9.2349	**30.4771**	0.3087	1.9014	75.62
	GFF	4.8164	6.8550	0.3733	9.2982	29.1646	0.6180	2.9697	4.74
	HMSD	5.4068	**6.9079**	0.5146	10.4805	30.0983	0.4876	2.4668	11.70
	IFE	4.1919	6.5826	0.4846	8.0533	28.4200	0.4595	3.0663	0.41
	GTF	2.6488	6.5386	0.4390	5.7551	29.8855	0.4133	1.3486	13.22
	Proposed	**5.9604**	6.7091	**0.6559**	**11.6891**	29.3733	**0.6261**	**3.1145**	8.18
4	NSST	4.0823	6.7336	0.5724	8.0539	**30.6569**	0.4026	2.5955	75.15
	GFF	3.5122	6.5458	0.5922	7.5658	25.7660	0.4010	1.6022	4.75
	HMSD	4.2453	6.7088	0.5526	8.6093	29.3511	0.4444	1.6590	11.22
	IFE	3.5143	**6.7961**	0.5478	6.4459	29.7513	0.4283	2.3118	0.40
	GTF	2.9368	6.7144	0.3990	5.2216	30.4435	0.3815	1.7595	14.25
	Proposed	**4.6509**	6.7142	**0.6222**	**8.9115**	28.8217	**0.4705**	**2.6522**	8.18

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

Group	Fusion methods	Evaluation index							
		AVG	IE	$Q^{AB/F}$	SF	SD	Q^{G}	MI	Time/s
5	NSST	4.8621	6.0796	0.4392	10.1306	20.7228	0.5644	1.7982	75.70
	GFF	4.6467	6.7834	0.2747	10.0294	29.4084	0.7459	3.6741	4.67
	HMSD	4.8185	6.6081	0.3158	10.2921	26.2318	0.7098	3.2535	11.60
	IFE	2.9955	6.2046	0.5158	7.1281	21.4265	0.4446	**3.9178**	0.21
	GTF	1.1509	5.8292	**0.6143**	3.2977	18.7667	0.6184	1.6860	11.70
	Proposed	**5.2458**	**7.0054**	0.4574	**10.6800**	**29.4313**	**0.6513**	3.2867	8.12
6	NSST	3.1173	7.4812	0.6000	5.7239	**55.5390**	0.3772	3.4720	39.52
	GFF	2.8302	6.4602	0.4703	5.3966	25.3406	0.5445	1.9612	2.78
	HMSD	3.3878	6.6189	0.5422	6.4034	32.0218	0.4781	1.7390	6.95
	IFE	2.4378	7.2408	0.4867	4.8878	53.6844	0.4718	3.3544	0.36
	GTF	2.1225	7.3043	0.4825	4.3037	54.7586	0.4666	3.1219	6.62

Proposed	**3.6821**	**7.3438**	**0.6072**	**6.7685**	50.3084	**0.4817**	**3.6604**	4.51

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

In general, we can clearly see that the AVG, SF, $Q^{AB/F}$, and Q^{G} values of the proposed algorithm are the highest in each group. Since the proposed low-frequency fusion rule can better target the low-light-level environment, the brightness of the fused image is in line with the human eyes and has the highest degree of clarity. In addition, the mixed-ℓ_1-gradient regularization makes the edge information more comprehensive, so the amount of gradient information transferred to the fused image is more, and the edge of the image is clearer. In view of the sharpness of the image and the increase of the amount of edge information, the overall activity of the fusion algorithm is greatly improved. Unfortunately, the IE and MI values of the algorithm in this paper do not reach the highest in each group of experiments, and they were behind the IFE algorithm in the first, second, fourth, and fifth groups of experiments. Unfortunately, the IE and MI values of the algorithm in this paper do not reach the highest in each group of experiments, and they are behind the IFE algorithm in the first, second, fourth, and fifth groups of experiments. We find that the fusion images of the IFE algorithm seem to have the disadvantage of excessive brightness. Although this can make up for the disadvantage of LL images, the high saturation brightness will cause the richness of the image information increasing abnormally, and the above two evaluation values are not true. In addition, the SD value of the proposed algorithm is lower than the NSST algorithm in the 3rd, 4th, and 6th groups. Although our fused images have achieved good visual results in these sets of experiments, the algorithm in this paper needs to be improved in the distribution of pixel gray-values.

We also simulated the running time of each algorithm and compared the timeliness of the algorithm by comparison. It is found that the timeliness of the IFE algorithm is the best and it is worthy of recognition. The NSST method uses the PCNN model as its fusion rule, and the running time is much larger than other comparison algorithms. The running time of the proposed algorithm is in the middle, slightly higher than the GFF method, which indicates that the timeliness of the algorithm needs to be improved. However, the good visual effects of the fusion algorithms often sacrifice certain timeliness, so the running time of the proposed algorithm is still acceptable.

In summary, we find that the results of the objective evaluation are basically consistent with the subjective visual effects. Although some evaluation values are not the highest in some individual experiments, the contrast ratio of the fused images are better, the transition of gray-value is natural, and the impression is better. Therefore, for the fusion problem in the low-light-level environment, the proposed algorithm seems to be slightly better.

5. Conclusion

In this paper, we propose a novel model based on ℓ_2-energy minimization and mixed-ℓ_1-gradient regularization for the fusion of infrared and low-light-level images. The ℓ_2-energy minimization can ensure that the brightness of the fused image is not affected by the low-light image, while the mixed-ℓ_1-gradient regularization restores the gradient feature of the source images as much as possible. With the construction of two new fusion rules, the fused images enhance brightness and preserve richer image features, so they are more in line with human visual observation. In order to verify the superiority of the fusion performance, we used six different low-light-level scenes for fusion experiments, and each group of experiment used five different excellent fusion algorithms for comparison. Finally, in terms of visual effects and parameter evaluation, the proposed algorithm is superior to other comparison algorithms, so it can better solve the problem of fusion for low-light-level and infrared images.

Acknowledgments

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

References

[1] Alex Pappachen James, Belur V. Dasarathy, Medical image fusion: A survey of the state of the art. Information Fusion 19 (2014) 4-19.
[2] Zhiqin Zhu, Hongpeng Yin, Yi Chai, Yanxia Li, Guanqiu Qi, A novel multi-modality image fusion method based on image decomposition and sparse representation. Information Sciences 432 (2018) 516-529.
[3] Wenda Zhao, Huimin Lu, and Dong Wang, Multisensor Image Fusion and Enhancement in Spectral Total Variation Domain. IEEE Transactions on Image Processing 20 (4) (2018) 866-879.
[4] Jiayi Ma, Yong Ma, Chang Li, Infrared and visible image fusion methods and applications: A survey. Information Fusion 45 (2019) 153-178.
[5] Jiao Du, Weisheng Li, Ke Lu, BinXiao, An overview of multi-modal medical image fusion. Neurocomputing 215 (2016) 3-20.
[6] Yu Liu, Shuping Liu, Zengfu Wang, A general framework for image fusion based on multi-scale transform and sparse representation. Information Fusion 24 (2015) 147-164.
[7] Yu Liu, Xun Chen, Rabab K, Z. Jane Wang, Image Fusion With Convolutional Sparse Representation. IEEE Signal Processing Letters 23 (12) (2016) 1882-1886.
[8] Zhiqin Zhu, Yi Chai, Hongpeng Yin, Yanxi Li, Zhaodong Liu, A novel dictionary learning approach for multi-modality medical image fusion. Neurocomputing 214 (2016) 471-482.
[9] Tianzhu Xiang, Li Yan, Rongrong Gao, A fusion algorithm for infrared and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain. Infrared Physics & Technology 69 (2015) 53–61.
[10] Boyang Cheng, Longxu Jin, Guoning Li, A novel fusion framework of visible light and infrared images based on singular value decomposition and adaptive DUAL-PCNN in NSST domain. Infrared Physics & Technology 91 (2018) 153-163.
[11] Boyang Cheng, Longxu Jin, Guoning Li, Adaptive fusion framework of infrared and visual image using saliency detection and improved dual-channel PCNN in the LNSST domain. Infrared Physics & Technology 92 (2018) 30-43.
[12] Xiangzhi Bai, Fugen Zhou, Bindang Xue, Infrared image enhancement through contrast enhancement by using multiscale new top-hat transform. Infrared Physics & Technology 54 (2011) 61-69.
[13] Xiangzhi Bai, Xiaowu Chen, Fugen Zhou, Zhaoying Liu, Bindang Xue, Multiscale top-hat selection transform based infrared and visual image fusion with emphasis on extracting regions of interest. Infrared Physics & Technology 60 (2013) 81-93.
[14] Pan Zhu, Xiaoqing Ma, Zhanhua Huang, Fusion of infrared-visible images using improved multi-scale top-hat transform and suitable fusion rules. Infrared Physics & Technology 81 (2017) 282-295.
[15] Zhiqiang Zhou, Sun Li, Bo Wan, Multi-scale weighted gradient-based fusion for multi-focus images. Information Fusion 20 (2014) 60-72.
[16] Wenda Zhao, Zhijun Xu, Jian Zhao, Gradient entropy metric and p-Laplace diffusion constraint-based algorithm for noisy multispectral image fusion. Information Fusion 27 (2016) 138-149.
[17] Gemma Piella, Image Fusion for Enhanced Visualization: A Variational Approach. Int J Comput Vis 83 (2009): 1-11.
[18] Shutao Li, Xudong Kang, Leyuan Fang, Jianwen Hu, HaitaoYin, Pixel-level image fusion: A survey of the state of the art. Information Fusion 33 (2017) 100-112.
[19] M. N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.
[20] A. L. Da Cunha, Zhou Jianping, M. N. Do, The nonsubsampled contourlet transform: theory, design, and applications. IEEE Trans. Image Process. 15 (10) (2006) 3089–3101.
[21] I. Selesnick, R. Baraniuk, N. Kingsbury, The dual-tree complex wavelet transform, IEEE Signal Process. Mag. 22 (6) (2005) 123–151.

[22] G. Easley, D. Labate, W. Lim, Sparse directional image representations using the discrete shearlet transform, Applied and Computational Harmonic Analysis, 25 (1) (2008) 25–46.
[23] LIM W Q. The Discrete Shearlet Transform: A New Directional Image Representation and Compactly Supported Shearlet Frames. IEEE Transactions on Image Processing, 19 (5) (2010) 1166-1180.
[24] Liu Zhanwen, Feng Yan, Chen Hang, Jiao Licheng, A fusion algorithm for infrared and visible based on guided filtering and phase congruency in NSST domain. Optics and Lasers in Engineering 97 (2017) 71-77.
[25] Xingbin Liu, Wenbo Mei, Huiqian Du, Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion. Neurocomputing 235 (2017) 131–139.
[26] Zhang Baohua, Lu Xiaoqi, Pei Haiquan, Zhao Ying, A fusion algorithm for infrared and visible images based on saliency analysis and non-subsampled Shearlet transform. Infrared Physics & Technology 73 (2015) 286–297.
[27] Zhanwen Liu, Yan Feng, Yifan Zhang, Xu Li, A fusion algorithm for infrared and visible images based on RDU-PCNN and ICA-bases in NSST domain. Infrared Physics & Technology 79 (2016) 183–190.
[28] Weiwei Kong, Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform. Infrared Physics & Technology 63 (2014) 110-118.
[29] Jinming Duan, Zhaowen Qiu, Wenqi Lu, Guodong Wang, Zhenkuan Pan, Li Bai, An edge-weighted second order variational model for image decomposition. Digital Signal Processing 49 (2016) 162-181.
[30] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2) (2009) 210-227.
[31] Bhawna Goyal, Ayush Dogra, Sunil Agrawal, B. S. Sohi, Two-dimensional gray scale image denoising via morphological operations in NSST domain & bitonic filtering. Future Generation Computer Systems 82 (2018) 158-175.
[32] Wei Wua, Zongming Qiu, Min Zhao, Qiuhong Huang, Yang Lei, Visible and infrared image fusion using NSST and deep Boltzmann machine. Optik 157 (2018) 334-342.
[33] Wei-Wei Wang, Peng-Lang Shui, and Xiang-Chu Feng, Variational Models for Fusion and Denoising of Multifocus Images. IEEE Signal Processing Letters 15 (2008) 65-68.
[34] Weiwei Kong, Longjun Zhang, Yang Lei, Novel fusion method for visible light and infrared images based on NSST-SF-PCNN. Infrared Physics & Technology 65 (2014) 103-112.
[35] S. Li, X. Kang, J. Hu, Image fusion with guided filtering. IEEE Trans. Image Process. 22 (2013a) 2864-2875.
[36] Zhiqiang Zhou, Bo Wang, Sun Li, Mingjie Dong, Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters. Information Fusion 30 (2016) 15-26.
[37] Yu Zhang, Lijia Zhang, Xiangzhi Bai, Li Zhang, Infrared and visual image fusion through infrared feature extraction and visual information preservation. Infrared Physics & Technology 83 (2017) 227–237.
[38] Jiayi Ma, Chen Chen, Chang Li, Jun Huang, Infrared and visible image fusion via gradient transfer and total variation minimization. Information Fusion 31 (2016) 100-109.
[39] Boyang Cheng, Longxu Jin, Guoning Li, Adaptive fusion framework of infrared and visual image using saliency detection and improved dual-channel PCNN in the LNSST domain. Infrared Physics & Technology 92 (2018) 30-43.
[40] Xiangzhi Bai, Yu Zhang, Fugen Zhou, Bindang Xue, Quadtree-based multi-focus image fusion using a weighted focus-measure. Information Fusion 22 (2015) 105-118.
[41] Yu Liu, Xun Chen, Hu Peng, Zengfu Wang, Multi-focus image fusion with a deep convolutional neural network. Information Fusion 36 (2017) 191-207.
[42] Weiwei Kong, Yang Lei, Minmin Ren, Fusion method for infrared and visible images based

on improved quantum theory model. Neurocomputing 212 (2016) 12-21.
[43] Boyang Cheng, Longxu Jin, Guoning Li, General fusion method for infrared and visual images via latent low-rank representation and local non-subsampled shearlet transform. Infrared Physics and Technology 92 (2018) 68-77.
[44] C. Xydeas, V. Petrović, Objective image fusion performance measure, Electron. Lett. 36 (4) (2000) 308–309.
[45] Yong Ma, Jun Chen, Chen Chen, Fan Fan, Jiayi Ma, Infrared and visible image fusion using total variation model. Neurocomputing 202 (2016) 12-19.

There are no conflicts of interest.

- The NSST is a multi-scale analysis tool for efficiently decomposing and reconstructing images.
- A fusion model based on image feature reconstruction and multi-scale decomposition is proposed in this paper.
- An optimization model based on ℓ_2-energy minimization is treated as the fusion rule for low-frequency sub-image.
- The fusion of high-frequency components is guided by the ℓ_1-gradient regularization.