

Article

Fusion of Infrared and Visible Images Based on Optimized Low-Rank Matrix Factorization with Guided Filtering

Jingyu Ji ¹, Yuhua Zhang ¹, Zhilong Lin ^{1,*}, Yongke Li ¹, Changlong Wang ¹, Yongjiang Hu ^{1,*}, Fuyu Huang ² and Jiangyi Yao ³

¹ Department of UAV, Army Engineering University, Shijiazhuang 050003, China; whale982022@163.com (J.J.); zyh842022@163.com (Y.Z.); lyk80001@163.com (Y.L.); wcl_lgd@163.com (C.W.)

² Department of Electronic and Optical Engineering, Army Engineering University, Shijiazhuang 050003, China; hfy_lgd@163.com

³ Equipment Simulation Training Center, Army Engineering University, Shijiazhuang 050003, China; yjy_jjy@163.com

* Correspondence: lzl912022@163.com (Z.L.); hyj70001@163.com (Y.H.)

Abstract: In recent years, image fusion has been a research hotspot. However, it is still a big challenge to balance the problems of noiseless image fusion and noisy image fusion. In order to improve the weak performance and low robustness of existing image fusion algorithms in noisy images, an infrared and visible image fusion algorithm based on optimized low-rank matrix factorization with guided filtering is proposed. First, the minimized error reconstruction factorization is introduced into the low-rank matrix, which effectively enhances the optimization performance, and obtains the base image with good filtering performance. Then using the base image as the guide image, the source image is decomposed into the high-frequency layer containing detail information and noise, and the low-frequency layer containing energy information through guided filtering. According to the noise intensity, the sparse reconstruction error is adaptively obtained to fuse the high-frequency layers, and the weighted average strategy is utilized to fuse the low-frequency layers. Finally, the fusion image is obtained by reconstructing the pre-fused high-frequency layer and the pre-fused low-frequency layer. The comparative experiments show that the proposed algorithm not only has good performance for noise-free images, but more importantly, it can effectively deal with the fusion of noisy images.

Keywords: image fusion; low-rank matrix factorization; guided filtering; minimized error reconstruction factorization; noisy image



Citation: Ji, J.; Zhang, Y.; Lin, Z.; Li, Y.; Wang, C.; Hu, Y.; Huang, F.; Yao, J. Fusion of Infrared and Visible Images Based on Optimized Low-Rank Matrix Factorization with Guided Filtering. *Electronics* **2022**, *11*, 2003. <https://doi.org/10.3390/electronics11132003>

Academic Editor: Manohar Das

Received: 9 June 2022

Accepted: 25 June 2022

Published: 26 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Infrared and visible image fusion is an important part of the image processing field, and its main purpose is to merge the complementary information in different images into one picture through technical algorithms [1]. Since the fused image can mostly maintain the significant features and energy information from various sensors, the fusion result can be utilized by subsequent processing tasks or decision-making assistance. Therefore, it can provide strong support in target detection and tracking, military fields, computer vision, remote sensing, and medical treatment [2].

In recent years, convolutional neural networks have developed rapidly and have been widely used in many fields [3]. Since deep learning can effectively extract and express salient features, it has developed rapidly in the field of image fusion. Liu et al. [4] proposed a convolutional neural network (CNN)-based infrared and visible image fusion method, using a twin convolutional network to obtain a weight map that integrates pixel activity information from two source images. Liu et al. [5] proposed a new method called convolutional sparse representation, which combines the advantages of convolutional neural networks and sparse representation for image fusion. Luo et al. [6] proposed an infrared and visible image fusion method based on Nonsubsampled Contourlet Transform

(NSCT) and stacked autoencoders. The image was decomposed into high-frequency and low-frequency layers using NSCT, and the low-frequency coefficients were calculated by stacking autoencoders to achieve image fusion. Hui et al. [7] used a deep learning network to extract salient features to obtain a better image fusion performance. Daniela et al. [8] designed an automated solution for facial feature recognition, enabling further applications of infrared and visible image fusion. However, judging from the fusion results of these papers, the performance of deep learning-based algorithms is not always better than traditional algorithms, and the results are even worse due to insufficient training samples. Moreover, due to the high computational complexity, powerful hardware support is required. In addition, due to the lack of actual ground data, both infrared and visible deep learning-based methods belong to unsupervised learning. Therefore, compared with the traditional method, the deep learning methods only rely on the composition of the network architecture and the design of the loss function, and it is difficult to obtain an overwhelming fusion result.

So far, the fusion methods based on multi-scale decomposition have been deeply and meticulously researched, and a good fusion performance has been achieved. For example, Singh et al. [9] designed two different infrared and visible image fusion schemes in the wavelet domain and feature space domain, respectively, and achieved good results in practice. Wang et al. [10] proposed an image fusion method based on an improved pulse-coupled neural network (PCNN) and multi-scale decomposition, which can produce good visual effects. Zhou et al. [11] proposed a new hybrid multi-scale image fusion method based on gradient-guided filtering. The fusion result can fully show the advantages in contrast and detail preservation. Ma et al. [12] proposed a multi-scale decomposition image fusion method through the combination of rolling guided filters and Gaussian filters. In order to improve the fusion performance of the detail layer, this method also proposed an optimized weighted least squares scheme. To overcome the limitations of the edge-preserving filter and reduce the artifacts at the edges of the image, Zhang et al. [13] used a new edge-preserving technology to achieve image fusion, and the co-occurrence filter can extract and fuse images. Therefore, a good image fusion effect was obtained. Duan et al. [14] proposed a new decomposition method called the multi-scale decomposition of the double exponential edge preservation smoother. This method can fully extract multi-scale structural information, and has good performance in terms of natural visual effects and detail preservation.

Most of the algorithms mentioned above ignore a key issue. Generally, the image fusion performance obtained by different image sensors is easily affected by imaging equipment and environmental factors, and there may be some noise in the images. However, traditional algorithms cannot take into account the fusion of noisy images and noise-free images simultaneously. To solve this problem, a new image fusion algorithm based on optimized low-rank matrix decomposition and guided filtering based on the traditional algorithm is proposed, which can effectively remove the noise in the image and obtain a good fusion image. In addition, the proposed algorithm also has good edge and detail preservation ability as well as good robustness. The main contributions of this article are as follows:

(1) To achieve a good denoising performance, the minimized error reconstruction factor is introduced. The effect of low-rank matrix decomposition is optimized, and image denoising can be achieved through update iteration;

(2) In order to effectively separate the noise information and energy structure information from the source image, guided filtering is utilized to decompose the source image at two scales. Among them, due to the optimization of low-rank matrix decomposition, an image with good denoising performance can be obtained, which can be utilized as a guided image, and then the good filtering performance of guided filtering can be utilized to better decompose the source image into a high-frequency layer with noise information and a low-frequency layer with energy structure information;

(3) In order to realize the denoising fusion of the high-frequency layer, an adaptive sparse error reconstruction method is proposed, which can adaptively change the denoising ability according to the intensity of the noise, avoiding excessive denoising or insufficient denoising.

The rest of this paper is organized as follows: Section 2 introduces some key theoretical algorithms used in this paper; Section 3 introduces the proposed algorithm; Section 4 introduces the comparative test and parameter setting; finally, the conclusion is described in Section 5.

2. Key Theories

2.1. Low-Rank Matrix Factorization Based on Minimizing Errors

The matrix D can be regarded as the low-rank part A and the sparse part E which can be modeled as the following optimization problem [15]:

$$\min_{A,E} \text{rank}(A) + \lambda \|E\|_0, s.t. D = A + E \tag{1}$$

where $\text{rank}(A)$ and $\|E\|_0$ are both nonlinear and non-convex, so optimization is difficult.

Therefore, it is necessary to use the rank and norm of the matrix to perform convex relaxation, so that the above formula is relaxed into the following convex optimization problem. In order to obtain a better optimization effect, the minimum error reconstruction factor β is introduced, and the above formula becomes:

$$\min_{A,E} \|A\|_* + \lambda \|E\|_{1,1} + \beta \|D - A - E\|_{1,1}, s.t. D = A + E \tag{2}$$

Solving convex optimization problems can be optimized by an iterative threshold algorithm, accelerated near-end gradient method, dual method, etc. In this paper, an augmented Lagrange multiplier algorithm (alternating direction multiplier method [16]) is used for optimization. First, construct the augmented Lagrange function:

$$L(A, E, Y, u) = \|A\|_* + \lambda \|E\|_{1,1} + \beta \|D - A - E\|_{1,1} + \langle Y, D - A - E \rangle + \frac{u}{2} \|D - A - E\|_F^2 \tag{3}$$

when $Y = Y_k, u = u_k$, the alternate algorithm is used to solve the optimization problem:

$$\min_{A,E} L(A, E, Y_k, u_k) \tag{4}$$

The exact Lagrange multiplier method is utilized to alternately iterate the matrices A and E until the termination condition is met. If $E = E_{k+1}^j$, then

$$\begin{aligned} A_{k+1}^{j+1} &= \underset{A}{\operatorname{argmin}} L(A, E_{k+1}^j, Y_k, u_k) = \underset{A}{\operatorname{argmin}} \|A\|_* + \beta \|D - A - E_{k+1}^j\|_{1,1} + \frac{u_k}{2} \left\| A - \left(D - E_{k+1}^j + \frac{Y_k}{u_k} \right) \right\|_F^2 \\ &= D_{\frac{1}{u_k}, \beta} \left(D - E_{k+1}^j + \frac{Y_k}{u_k} \right) \end{aligned} \tag{5}$$

Then update the matrix E according to A_{k+1}^{j+1} :

$$\begin{aligned} E_{k+1}^{j+1} &= \underset{E}{\operatorname{argmin}} L(A_{k+1}^{j+1}, E, Y_k, u_k) = \underset{E}{\operatorname{argmin}} \lambda \|E\|_{1,1} + \beta \|D - A_{k+1}^{j+1} - E\|_{1,1} + \frac{u_k}{2} \left\| E - \left(D - A_{k+1}^{j+1} + \frac{Y_k}{u_k} \right) \right\|_F^2 \\ &= S_{\frac{1}{u_k}, \beta} \left(D - A_{k+1}^{j+1} + \frac{Y_k}{u_k} \right) \end{aligned} \tag{6}$$

Let A_{k+1}^* and E_{k+1}^* be the exact values of A_{k+1}^{j+1} and E_{k+1}^{j+1} , respectively, then the update formula of matrix Y is:

$$Y_{k+1} = Y_k + u_k (D - A_{k+1}^* - E_{k+1}^*) \tag{7}$$

The parameter u_k can be updated as follows:

$$u_{k+1} = \begin{cases} \rho u_k, & \frac{u_k \|E_{k+1}^* - E_k^*\|_F}{\|D\|_F} < \varepsilon \\ u_k, & \text{otherwise} \end{cases} \tag{8}$$

where $\rho > 1$ is a constant, and $\varepsilon > 0$ is a small positive number.

The above-mentioned accurate Lagrange multiplier method (ALM) requires multiple updates in the inner loop and performs multiple singular value decompositions. Therefore, an inaccurate Lagrange multiplier method is proposed, which does not require the exact solution of $\min_{A,E} L(A, E, Y_k, u_k)$ before the external loop starts; that is, the inner loop of the ALM method is removed, and the update formula becomes the following form:

$$A_{k+1} = \operatorname{argmin}_A L(A, E_{k+1}, Y_k, u_k) = D_{\frac{1}{u_k}, \beta} (D - E_{k+1} + \frac{Y_k}{u_k}) \tag{9}$$

$$E_{k+1} = \operatorname{argmin}_E L(A_{k+1}, E, Y_k, u_k) = S_{\frac{\lambda}{u_k}, \beta} (D - A_{k+1} + \frac{Y_k}{u_k}) \tag{10}$$

where $D_{\frac{1}{u_k}, \beta}$ and $S_{\frac{\lambda}{u_k}, \beta}$ are singular value threshold operators and soft threshold operators, respectively.

2.2. Guided Filtering

Traditional edge-preserving smoothing filters including the weighted least squares filter [17] or bilateral filter [18] are widely utilized in the field of image processing, which can avoid ringing artifacts and achieve the effect of not blurring the edges during the decomposition process. Guided filtering [19] also belongs to an edge-preserving filtering algorithm, which can obtain an edge-preserving smooth image through a guided image. Guided filtering is a local linear model of guided image G_i and filter output O_i :

$$O_i = p_n G_i + q_n, i \in \theta_n \tag{11}$$

where p_n and q_n are constants in the window θ_n at the pixel n . The idea of optimized regression is used to solve I_o . Then the cost function is defined, and a regular term ϵ is added to the cost function through the ridge regression method to prevent overfitting.

$$E(p_n, q_n) = \sum_{i \in \theta_n} ((p_n G_i + q_n - I_i)^2 + \epsilon p_n^2) \tag{12}$$

where I_i is the input image. Through this formula, p_n and q_n that minimize $E(p_n, q_n)$ can be obtained

$$p_n = \frac{\frac{1}{|\alpha|} \sum_{i \in \theta_n} G_i I_i - \theta_n \hat{I}_n}{\sigma_n^2 + \epsilon} \tag{13}$$

$$q_n = \hat{I}_n - p_n \theta_n \tag{14}$$

where θ_n and σ_n^2 are the mean value and variance of I in θ_n , respectively, \hat{I}_n is the mean value of I in θ_n , and $|\alpha|$ is the number of pixels in θ_n . The p_n and q_n in each window are obtained by traversing the image through the window. However, each pixel may be contained in multiple windows, which leads to multiple calculations of p_n and q_n . Therefore, in order to simplify the calculation, take the average value \hat{p}_n, \hat{q}_n of p_n and q_n , and then obtain:

$$O_i = \hat{p}_n G_i + \hat{q}_n \tag{15}$$

Guided filtering is different from most filtering methods in that it requires direct convolution, and its calculation time has nothing to do with the filter parameters. Because of its good edge retention and structure transfer characteristics, it is widely utilized in the fields of image decomposition, image smashing, and image fusion. Figure 1 is a schematic diagram of guided filtering.

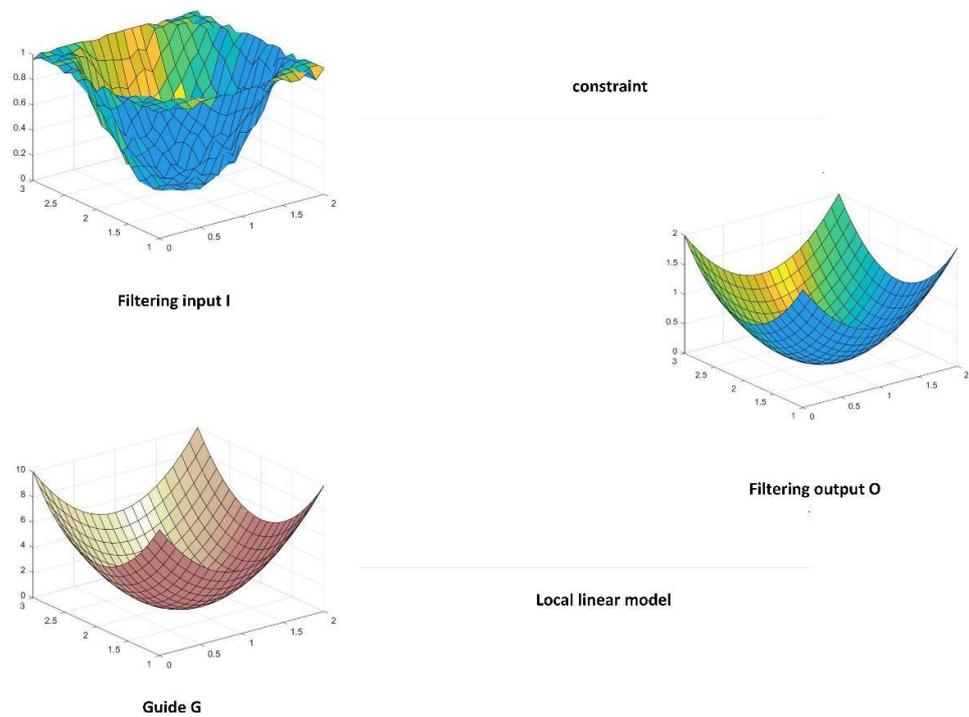


Figure 1. The principle diagram of guided filtering.

3. Fusion Framework

In order to solve the problem of effectively retaining details while denoising, a new fusion model is introduced. Different from traditional decomposition methods, the source image is first denoised and decomposed by using an optimized low-rank matrix in order to achieve better denoising effect. At this time, the source image is decomposed into base component and detail component, and most of the disturbance can be completely preserved in the detail component. Then the base component is used as the guide image, the source image is used as the input image, and the source image is decomposed into a high-frequency layer and a low-frequency layer through a guided filter. The high-frequency layer contains detail and noise components, and the low-frequency layer contains energy and structure information. Different fusion methods are introduced to obtain pre-fusion layers based on the characteristics of the two layers. Among them, for the high-frequency layer fusion, fusion denoising is effectively realized by combining the relationship between sparse representation and noise intensity; for the low-frequency layer, a weighted average fusion strategy is used for pre-fusion. Finally, the final fusion image is realized by reconstructing the two pre-fusion layers. Figure 2 shows the main flow of the algorithm in this paper.

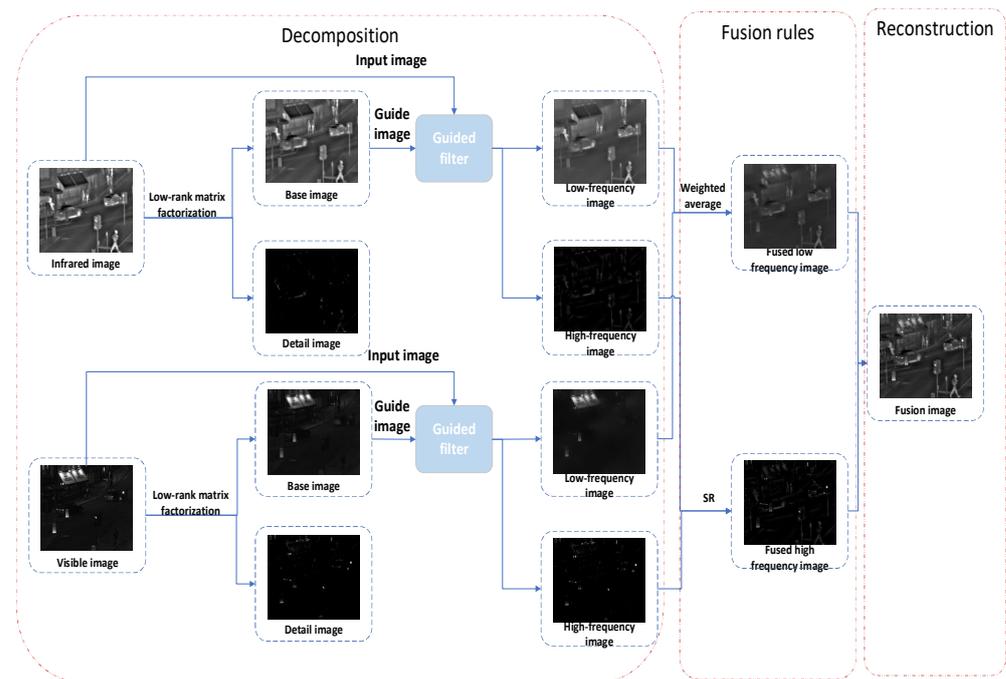


Figure 2. The framework of the fusion algorithm.

3.1. The Decomposition Model

In order to separate the noise in the source image in a targeted manner, using the optimized low-rank matrix’s better denoising effect, first process the source image:

$$\left(I_n^b, I_n^d \right) = LRF(I_n, \mu, \lambda) \tag{16}$$

where I_n is the n th source image, $n \in \{1, 2, \dots, N\}$, μ and λ are the iteration error and the number of iterations, respectively, $LRF(\cdot)$ is the low-rank matrix factorization operator, and I_n^b is the base component after the noise removal. Next, I_n is used as the input image, I_n^b is used as the guide image, and the low-frequency layer of the image is obtained through the guiding filter:

$$I_n^l = GF\left(I_n, I_n^b, \sigma_s, \sigma_r\right) \tag{17}$$

where σ_s, σ_r are filter parameters, $GF(\cdot)$ is the guided filter operator, and I_n^l represents the low-frequency layer of I_n . After the guided filtering, most of the noise has been removed from the image, and the important details and structural information in the image are retained in the low-frequency layer. The high-frequency layer of the image can be obtained by the following formula:

$$I_n^h = I_n - I_n^l \tag{18}$$

Each group of images in Figure 3 contains a noise-free image and a noisy image with $\sigma = 20$. These two groups of images test the reliability of the proposed decomposition model, especially for noisy images. It can be seen from Figure 3:

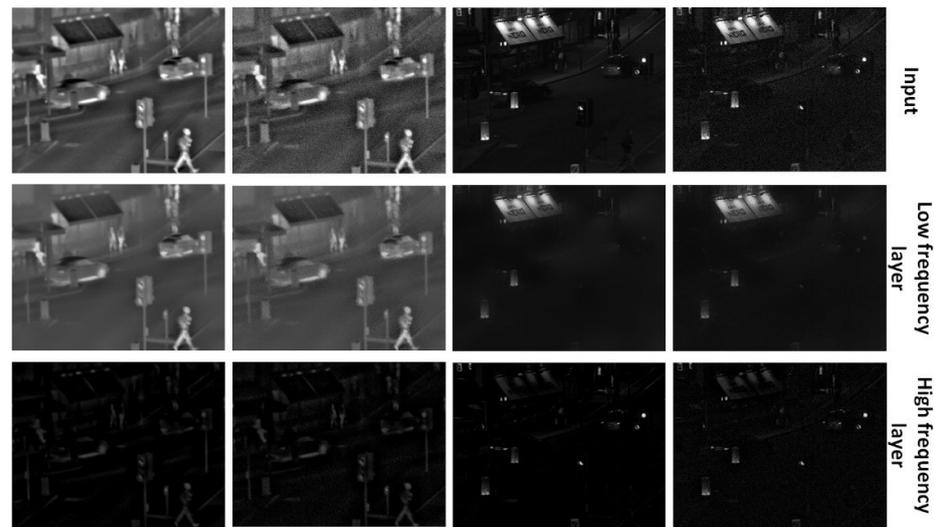


Figure 3. Two-scale decomposition image.

(1) After decomposition by the proposed algorithm, most of the noise and details are preserved to the high-frequency layer. At the same time, it can be seen that some details still exist in the low-frequency layer;

(2) The low-frequency layer produced by noise-free and noisy images is very similar; that is, the noise information almost completely exists in the high-frequency layer.

3.2. Fusion Rules

3.2.1. High-Frequency Layers Pre-Fusion

The method based on SR can well realize the fusion denoising of the detail layers. It includes two stages: dictionary learning and sparse representation. In the first step, the high-frequency layer of training data is generated by Equation (17), 8×8 blocks are collected from detail images, and the final training collection is constructed. The Kernel-based singular value decomposition (KSVD) [20] algorithm can be used to obtain a complete dictionary D . In the second step, an 8×8 block of each source image is taken and normalized. To obtain SR parameters of high-frequency layers, an Orthogonal Matching Pursuit (OMP) [21] algorithm is utilized by Equation (18):

$$\min_{\alpha_n^k} \|\alpha_n^k\|_0, \quad s.t. \|V_n^k - D\alpha_n^k\|_0 < \varepsilon \quad (19)$$

where V_n^k is the k -th block of I_n , and α_n^k is the related sparse vector. ε is the sparse reconstruction error, represented as:

$$\varepsilon = \begin{cases} P, & \sigma = 0 \\ 0.005 + 8E\sigma, & \sigma > 0 \end{cases} \quad (20)$$

where σ is the Gaussian standard deviation, P is a constant, and $E > 0$ influences ε when $\sigma > 0$. Next the “absolute-maximum” is utilized to obtain fusion sparse representation coefficients:

$$\alpha_{F^h}^k = \alpha_{\hat{n}}^k, \quad \hat{n} = \operatorname{argmax}_n \left\{ \|\alpha_n^k\|_1 \mid n = 1, 2, \dots, N \right\} \quad (21)$$

The fusion high-frequency vector $\bar{\alpha}_{F^h}^k$ can be obtained by:

$$\bar{\alpha}_{F^h}^k = D\alpha_{F^h}^k \quad (22)$$

Finally, the fusion high-frequency layer is obtained by reshaping each $\bar{\alpha}_{F^h}^k$ into 8×8 blocks and then arranging them according to the initial location to generate the pre-fused F^h .

3.2.2. Low-Frequency Layers Pre-Fusion

The low-frequency layer of the source image contains more global structural information and energy information. Therefore, this paper uses a weighted average strategy [22] for low-frequency layer fusion:

$$F^l = \omega_1 I_1^h + \omega_2 I_2^h \quad (23)$$

where ω_1 and ω_2 represent the weight value. In order to maintain the global structure and energy information and reduce redundant information, let $\omega_1 = 0.5$ and $\omega_2 = 0.5$.

After obtaining these two pre-fusion components, the final fusion image F is:

$$F = F^h + F^l \quad (24)$$

4. Discussion

In this section, after setting the parameters of the proposed algorithm, a comparative experiment is carried out, including the experiment of the noiseless image and the experiment of the image with noise. Qualitative and quantitative analyses were carried out, respectively.

4.1. Experimental Setup

The experimental dataset is selected from the website https://figshare.com/articles/TN_Image_Fusion_Dataset/1008029 (accessed on 15 May 2022) to verify the proposed algorithm. Six pairs of images are shown in Figure 4. Five recent methods are compared in the same experimental environment for verification, including CBF [23], CNN [4], GTF [24], IFEVIP [25], and TIF [26]. Furthermore, the fusion performance is quantitatively evaluated by six indicators, including entropy (EN) [27], edge information retention ($Q_{AB/F}$) [28], Chen-Blum's index (Q_{CB}) [29], mutual information (MI) [30], structural similarity (SSIM) [31], and peak signal-to-noise ratio (PSNR) [32].



Figure 4. Six pairs of source images.

EN is used to measure the amount of information contained in the source image in the fusion image. $Q_{AB/F}$ utilizes local metrics to estimate how well salient information from source images is represented in fused images. Q_{CB} is used as a human visual evaluation index to measure the quality of fused images. MI is used to measure the amount of information transferred from the source image into the fused image. SSIM is used to measure the structural similarity between the fused image and the source image. PSNR is used to measure the ratio between the effective information of the image and the noise, which can reflect whether the image is distorted. In summary, these metrics are chosen to evaluate the fused images obtained by the proposed algorithm from different perspectives.

4.2. Parameter Settings

The controlled variable method is utilized to analyze the two free parameters in the model: the number of iterations λ in Equation (15) and the sparse reconstruction error parameter E in Equation (20). In addition, μ in Equation (15) is set to 10^{-8} , σ_s in Equation (17) is set to 0.1, and parameter P in Equation (20) is set to 0.001.

The discussion of E

First, fix $\lambda = 200$ and then use the two indicators SSIM and MI to analyze the performance under different E . The experimental results are shown in the Figure 5. It can be seen that both indicators are better when $E = 0.003$, and the fusion performance decreases when $E < 0.003$ or $E > 0.003$. Therefore, after comprehensive consideration, the best value for E is 0.003.

The discussion of λ

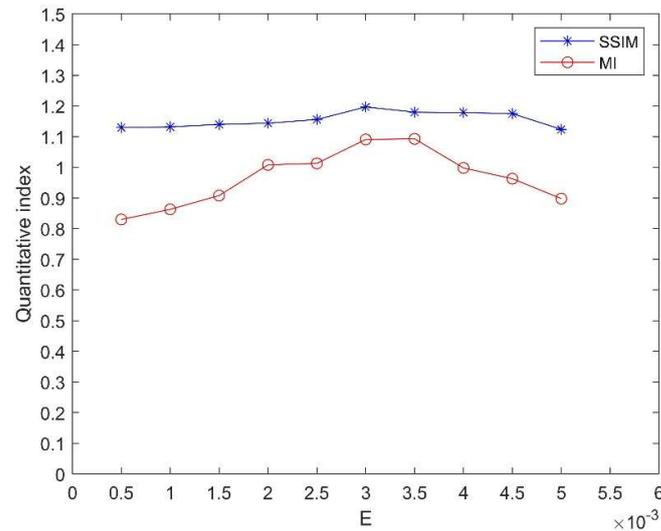


Figure 5. Quantitative evaluation of the fused images produced by different E .

If λ is too small, it will affect the denoising effect, but if λ is too large, it will affect the speed of the experiment. Therefore, it is very necessary to choose an appropriate number of iterations. Fix $E = 0.003$. In order to better determine the number of iterations, SSIM, MI, and time T are used to observe the fusion performance and speed. It can be seen from Figure 6 that when $\lambda < 300$, as the number of iterations increases, the image fusion effect gets better and better. When $\lambda > 300$, the fusion effect hardly changes. In addition, the fusion speed gradually slows down with an increase in the number of iterations. Taking two factors into consideration, the best value for λ is 300.

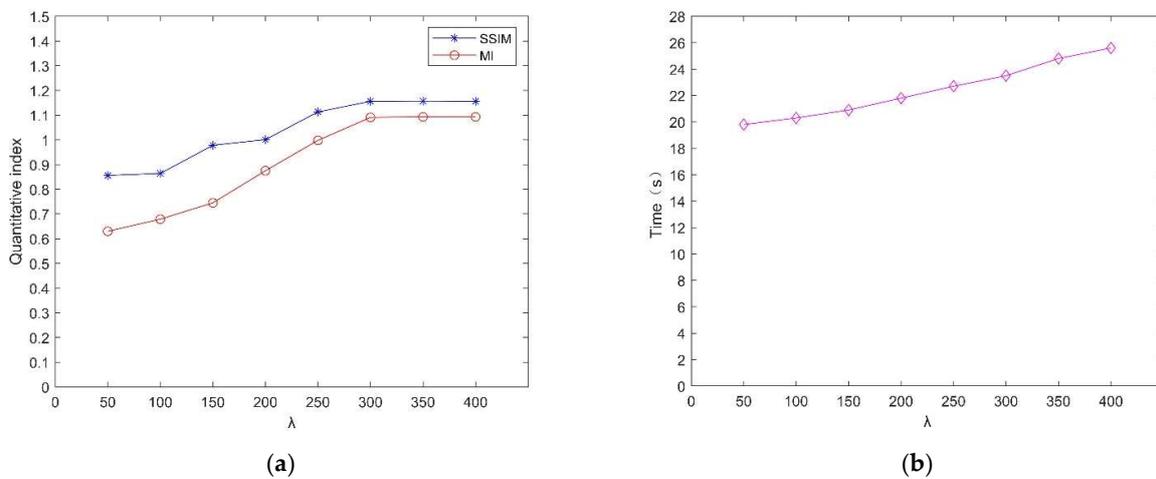


Figure 6. Quantitative evaluation of the fused images produced by different λ . ((a) represents the quality evaluation value of different λ , and (b) represents the consumption time of different λ .)

4.3. Noise-Free Image Fusion and Evaluation

Figure 7 shows the fusion results of the proposed algorithm and the comparison algorithm. The first column contains infrared images, and the second column contains visible images; the remaining images are the fusion images obtained by various methods.

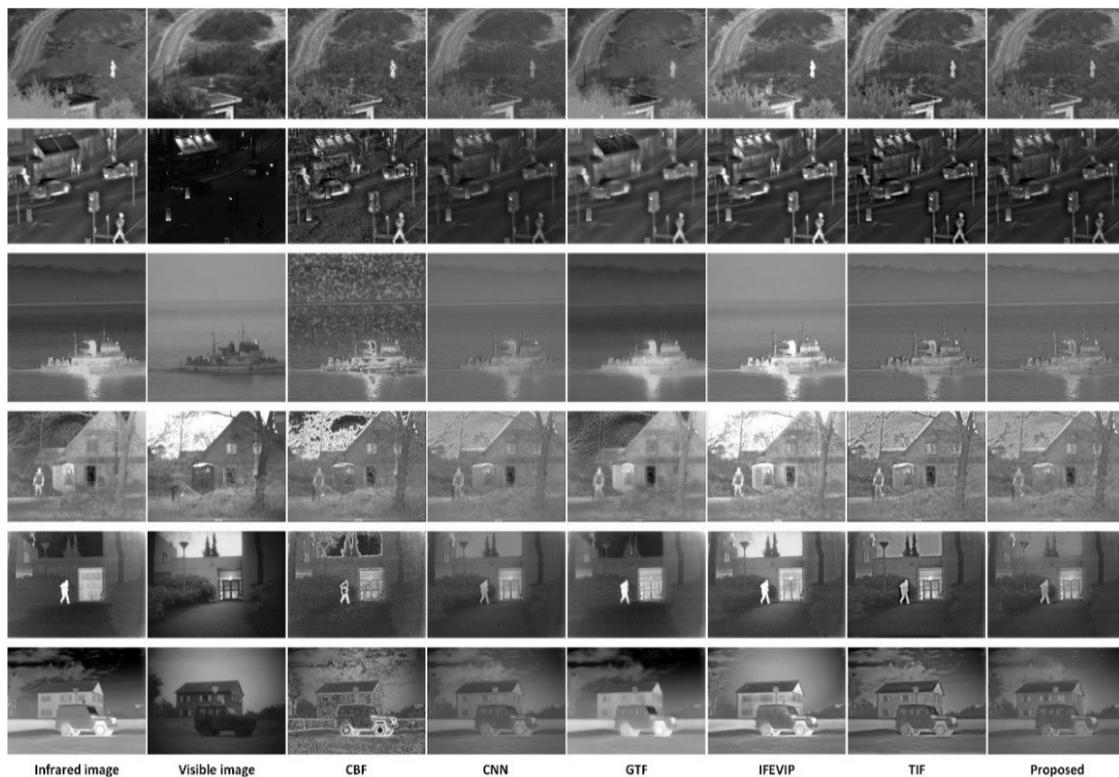


Figure 7. Fusion results of noise-free images.

4.3.1. Subjective Evaluation

It can be seen from Figure 7 that the proposed method can retain more detail information, and there is less manual information. This is because the proposed two-scale decomposition algorithm can well separate the noise information and other main detail information, and the fusion rules are set appropriately. However, the salient features of the images obtained by CBF are not obvious and contain more artificial noise information. Although the fusion images generated by the CNN have lower brightness than the images generated by the proposed algorithm, the structures are better preserved. The GTF and IFEVIP methods maintain a good brightness, but the visual effects are too strengthened, resulting in obvious errors in the results. The TIF method has the phenomenon of fuzzy internal characteristics. Therefore, in the fusion results, the proposed algorithm can preserve the important content of the source image and obtain the best visual performance in terms of brightness and structural details, which means that the proposed algorithm can produce better subjective performance.

4.3.2. Objective Evaluation

Figure 8 shows the different objective evaluation values of the fusion results in the six pairs of images. From each subgraph in Figure 8, we can see that the index values of the proposed algorithm are almost all the highest, especially for the four indexes of $Q_{AB/F}$, Q_{CB} , MI, and PSNR; the proposed algorithm is always better compared to other algorithms. For the EN indicator, the proposed algorithm only performs poorly in the boat image. In addition, compared with other algorithms, the proposed algorithm has obvious advantages in the $Q_{AB/F}$ index, and the value in Figure 8 is significantly higher than other methods.

Among various evaluation indicators, the proposed algorithm is not optimal in only a few places, but it can still be proven that the proposed algorithm has good performance.

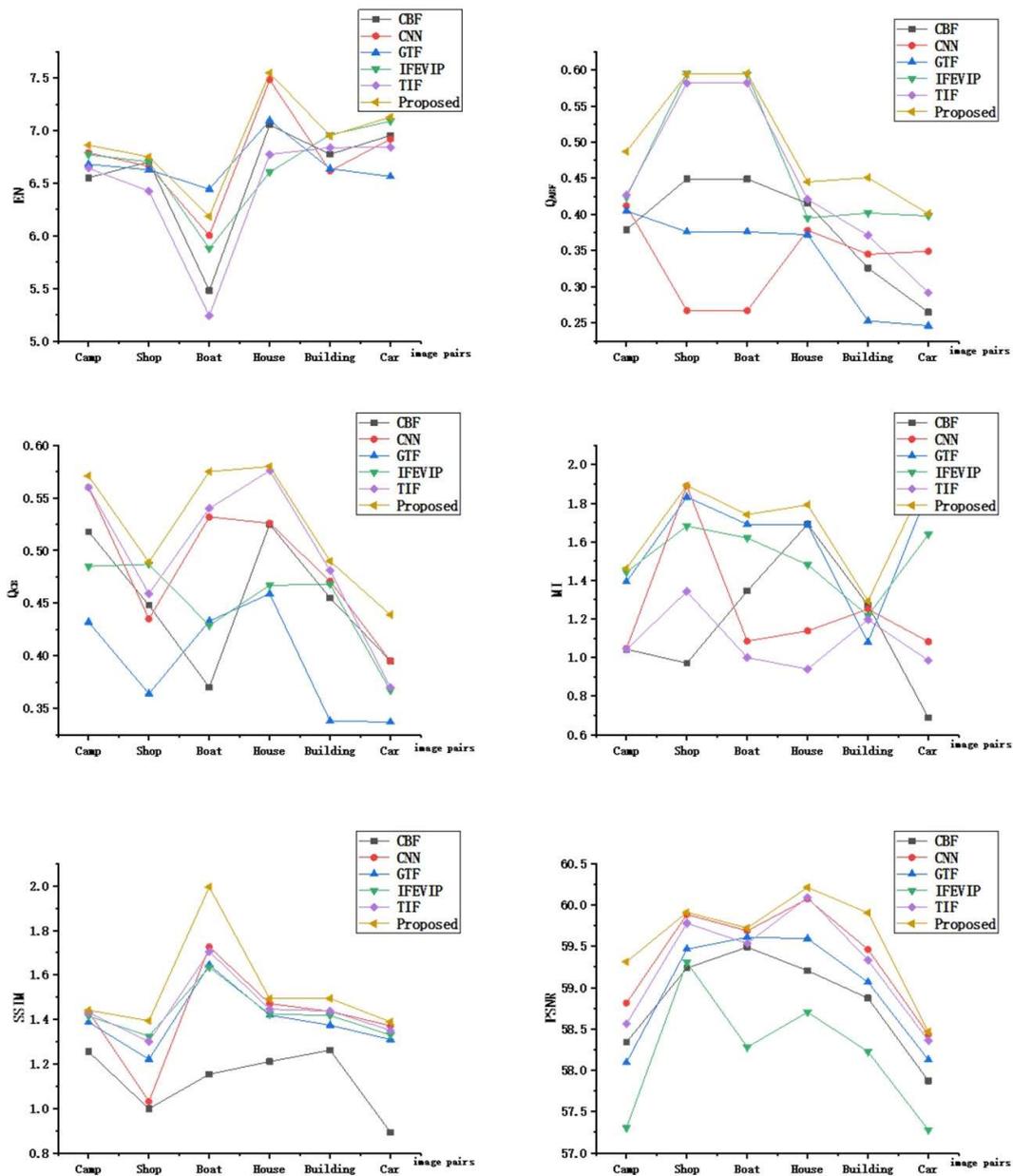


Figure 8. Objective evaluation metrics for fusion results.

In summary, the proposed algorithm performs well both qualitatively and quantitatively for the fusion of noise-free infrared and visible images.

4.4. Noisy Image Fusion and Evaluation

Figures 9 and 10 are examples of six pairs of noisy infrared and visible images. The noise intensity of the source images in Figures 9 and 10 are 10 and 20, respectively. The first column of Figures 9 and 10 contains the infrared images, and the second column contains the visible images; the remaining images are the fusion images obtained by various methods.

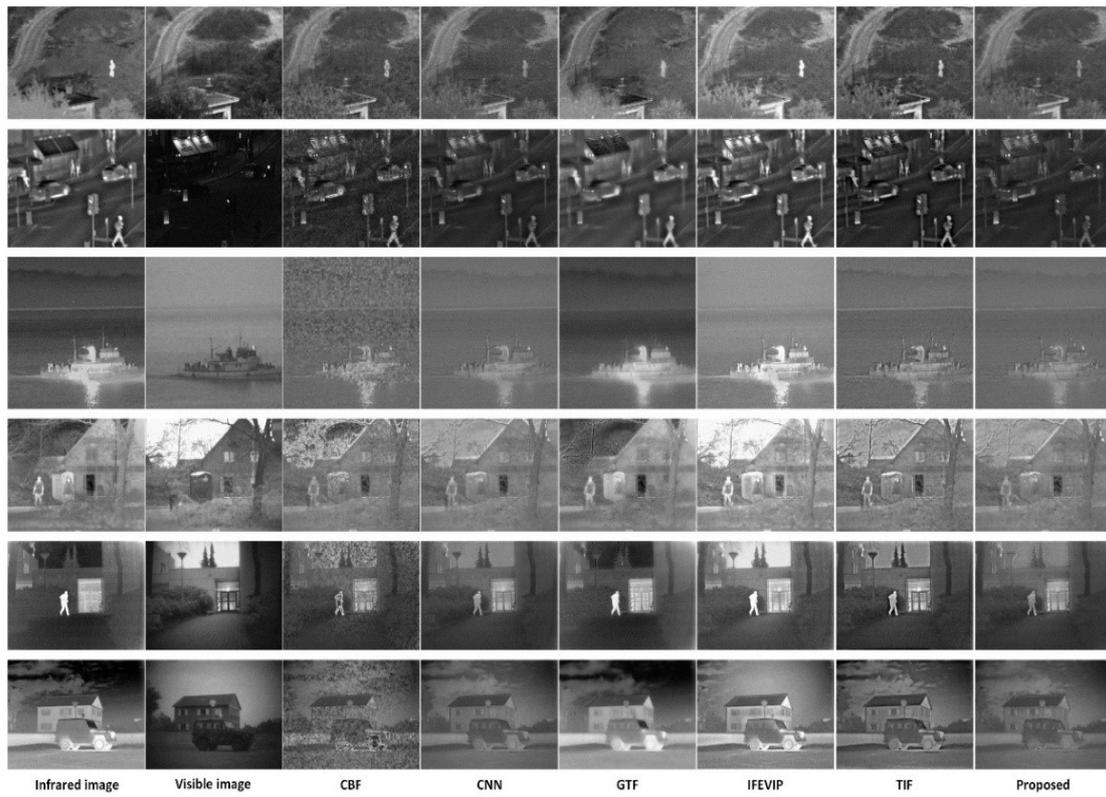


Figure 9. Fusion results of noisy images ($\sigma = 10$).

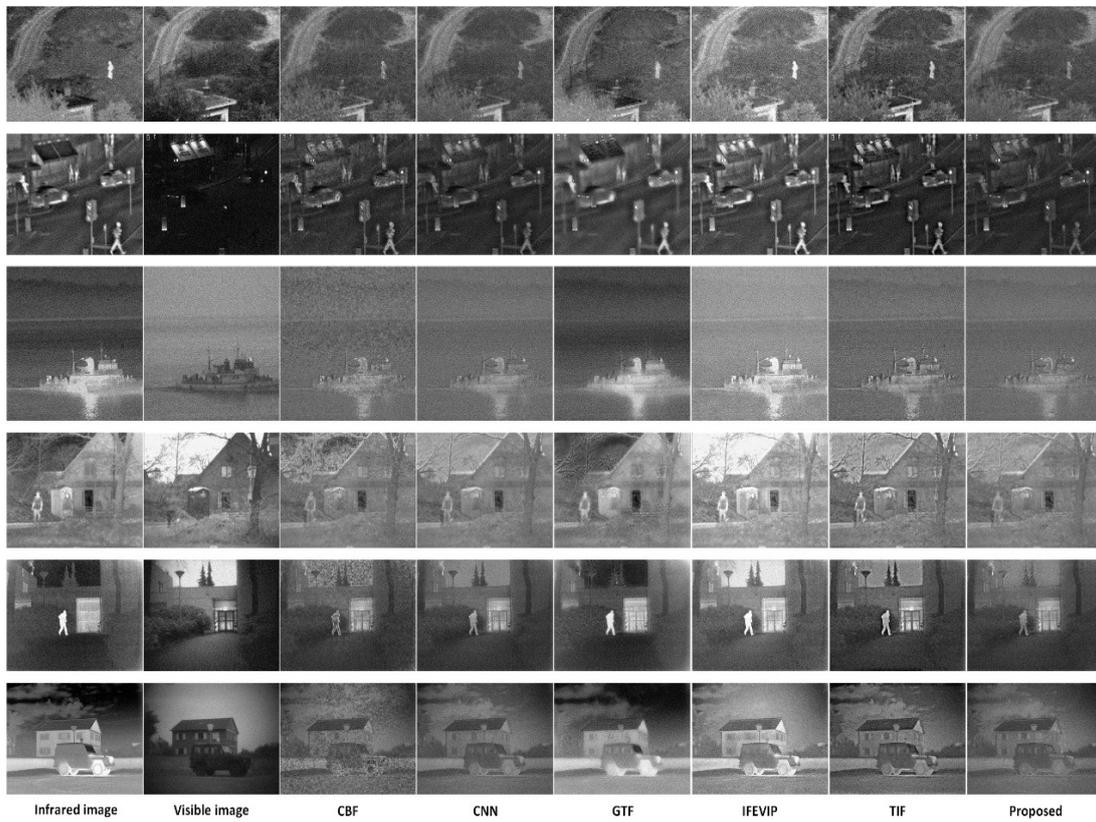


Figure 10. Fusion results of noisy images ($\sigma = 20$).

4.4.1. Subjective Evaluation

When the noise intensity is 10, the noise-removing capabilities of CBF and TIF methods are insufficient, and their fusion results lack useful information. The CNN method has a certain denoising effect, but the result is too low in contrast. GTF and IFEVIP methods can denoise effectively to a certain extent, but the contrast is too high and the image is unnatural. These two methods can be fused in a noisy environment, but some irrelevant information will be introduced, resulting in unreal visual effects. Compared with other algorithms, the proposed algorithm has the best fusion performance in detail reservation, and the noise in the fusion results is significantly reduced simultaneously, so it has good performance in denoising.

When the noise intensity reaches 20, the contours of the fusion results obtained by the CBF, CNN, and TIF methods have been severely damaged, and a large number of outstanding mistakes have been taken into the fusion results. In the IFEVIP method, the contrast is too high. Although the GTF method can denoise, the result is too smooth and lacks detail information. In contrast, the fusion results of the proposed algorithm not only preserve the detail content, contrast, and structure preservation of the source image, but its denoising effect is also remarkable.

4.4.2. Objective Evaluation

The objective evaluation of fusion results is shown in Tables 1 and 2. Compared with CBF, CNN, GTF, IFEVIP, and TIF methods, the proposed method can obtain better objective analysis results, and is basically consistent with the objective evaluation results of noise-free image fusion. So, it proves the usefulness and superiority of the method proposed in this paper.

In summary, for the fusion of noisy infrared and visible images, the proposed algorithm has a good performance both qualitatively and quantitatively. This is because the two-scale decomposition algorithm designed in this paper can well separate the noise information and structural information in the source image, which are reflected in the high-frequency layer and the low-frequency layer, respectively. Through the adaptive sparse fusion algorithm, the denoising fusion of the high-frequency layer can be adaptively realized according to the intensity of the noise, and there will be no phenomenon of excessive denoising or insufficient denoising, which lays the foundation for the final fusion effect.

Table 1. Quantitative index of image fusion results ($\sigma = 10$).

Source Images	Index	CBF	CNN	GTF	IFEVIP	TIF	Proposed
Camp	EN	6.601	6.761	6.820	6.901	6.403	6.797
	Q_{ABF}	0.317	0.422	0.459	0.380	0.359	0.480
	Q_{CB}	0.517	0.550	0.475	0.520	0.561	0.568
	MI	0.889	0.905	0.933	0.786	0.945	1.080
	SSIM	1.213	1.109	1.090	1.224	1.200	1.297
	PSNR	58.467	58.548	57.782	56.807	58.362	58.933
Shop	EN	6.559	6.807	6.739	6.883	6.608	6.890
	Q_{ABF}	0.301	0.453	0.408	0.474	0.408	0.497
	Q_{CB}	0.447	0.438	0.294	0.384	0.446	0.472
	MI	0.818	1.225	0.878	1.479	1.050	1.595
	SSIM	0.980	1.050	0.764	1.120	1.018	1.194
	PSNR	59.637	59.889	59.222	59.177	59.712	59.997
Boat	EN	6.141	6.756	6.788	6.283	6.608	6.867
	Q_{ABF}	0.273	0.481	0.475	0.471	0.317	0.496
	Q_{CB}	0.439	0.569	0.469	0.488	0.547	0.576
	MI	0.474	0.771	1.315	1.381	0.540	1.378
	SSIM	1.145	1.200	1.095	1.217	1.229	1.295
	PSNR	59.674	59.833	59.159	58.148	59.804	59.826

Table 1. Cont.

Source Images	Index	CBF	CNN	GTF	IFEVIP	TIF	Proposed
House	EN	6.783	6.640	6.512	6.989	6.871	7.142
	Q_{ABF}	0.305	0.453	0.456	0.394	0.368	0.456
	Q_{CB}	0.474	0.474	0.470	0.508	0.568	0.574
	MI	0.727	0.896	1.027	1.535	0.791	1.696
	SSIM	1.128	1.173	1.100	1.224	1.202	1.293
	PSNR	59.720	60.172	59.458	58.560	60.068	60.198
Building	EN	6.935	6.882	7.114	7.272	7.031	7.349
	Q_{ABF}	0.278	0.476	0.440	0.407	0.341	0.540
	Q_{CB}	0.467	0.485	0.435	0.509	0.532	0.556
	MI	0.807	1.036	1.169	1.140	0.965	1.182
	SSIM	1.117	1.131	0.991	1.213	1.159	1.294
	PSNR	59.175	59.349	58.736	58.004	59.235	59.943
Car	EN	6.787	6.627	7.113	7.144	6.906	7.506
	Q_{ABF}	0.230	0.421	0.412	0.455	0.351	0.527
	Q_{CB}	0.414	0.374	0.366	0.424	0.468	0.476
	MI	0.421	0.672	0.844	0.671	0.726	0.926
	SSIM	0.941	1.030	0.878	1.138	1.062	1.189
	PSNR	58.131	58.371	57.875	57.137	58.315	58.494

Table 2. Quantitative index of image fusion results ($\sigma = 20$).

Source Images	Index	CBF	CNN	GTF	IFEVIP	TIF	Proposed
Camp	EN	6.942	6.890	7.131	7.264	7.123	7.277
	Q_{ABF}	0.285	0.322	0.496	0.343	0.311	0.429
	Q_{CB}	0.474	0.456	0.498	0.492	0.552	0.553
	MI	0.870	0.863	0.711	1.008	0.926	1.091
	SSIM	1.160	1.132	0.948	1.144	1.070	1.197
	PSNR	57.926	57.995	56.992	55.801	57.671	58.291
Shop	EN	6.878	6.972	6.983	7.237	6.881	7.519
	Q_{ABF}	0.326	0.325	0.407	0.449	0.338	0.520
	Q_{CB}	0.441	0.426	0.302	0.472	0.467	0.495
	MI	0.960	0.748	0.700	1.510	0.866	1.330
	SSIM	1.009	0.854	0.615	1.063	0.924	1.094
	PSNR	59.593	59.645	58.742	58.782	59.508	59.928
Boat	EN	6.612	6.764	6.225	6.855	6.653	6.995
	Q_{ABF}	0.268	0.384	0.515	0.332	0.283	0.521
	Q_{CB}	0.485	0.493	0.487	0.501	0.524	0.542
	MI	0.452	0.550	0.957	0.831	0.461	0.977
	SSIM	1.136	1.063	0.910	1.128	1.064	1.195
	PSNR	59.324	59.322	58.362	57.566	59.267	59.689
House	EN	6.910	6.925	7.314	7.211	7.129	7.494
	Q_{ABF}	0.276	0.346	0.421	0.348	0.309	0.481
	Q_{CB}	0.471	0.448	0.487	0.500	0.424	0.551
	MI	0.492	0.603	0.568	0.425	0.585	0.649
	SSIM	1.128	1.100	0.945	1.147	1.064	1.193
	PSNR	59.680	59.083	58.975	58.188	59.787	59.928

Table 2. Cont.

Source Images	Index	CBF	CNN	GTF	IFEVIP	TIF	Proposed
Building	EN	7.152	7.150	7.339	7.545	7.300	7.062
	Q_{ABF}	0.267	0.325	0.476	0.356	0.303	0.486
	Q_{CB}	0.464	0.442	0.449	0.469	0.519	0.538
	MI	0.735	0.880	0.711	0.867	0.786	0.965
	SSIM	1.088	1.063	0.840	1.137	1.027	1.193
	PSNR	58.978	59.066	58.121	57.501	58.904	59.885
Car	EN	6.927	6.897	7.755	7.362	7.107	7.780
	Q_{ABF}	0.252	0.342	0.458	0.405	0.300	0.535
	Q_{CB}	0.431	0.387	0.375	0.428	0.464	0.493
	MI	0.416	0.500	1.175	1.118	0.542	1.342
	SSIM	1.008	0.983	0.718	1.078	0.953	1.089
	PSNR	58.048	58.197	57.378	56.827	58.108	58.435

4.5. Computational Efficiency

In order to test the real-time performance of the algorithm in this paper, the various methods were placed in the same experimental environment for comparison, and the average execution time comparison is shown in Table 3. Since the experiment needs to perform multiple iterations and achieve partial fusion through sparse representation, the efficiency of the proposed method is not very high. Therefore, in future research, improving algorithm performance and increasing computational efficiency are important research directions.

Table 3. Computational efficiency of different methods.

Method	CBF	CNN	GTF	IFEVIP	TIF	Proposed
Time/s	10.73	23.16	2.91	1.34	1.03	22.03

5. Conclusions

In this paper, an infrared and visible image fusion algorithm based on optimized low-rank matrix decomposition and guided filtering is proposed. The proposed algorithm takes advantage of the filtering effect of low-rank matrix decomposition on noisy images, and introduces a reconstruction factor to minimize the error to improve the decomposition efficiency and performance. The final two-scale decomposition is achieved through guided filtering, and the noise information and structure information are better separated to obtain a better fusion performance. A large number of fusion results show that the proposed algorithm is obviously superior to the existing fusion methods in visual and quantitative evaluation, and can obtain strong anti-noise performance. Furthermore, the method can be effectively extended to image fusion problems of other modalities.

Author Contributions: Methodology, Y.L.; software, J.Y.; validation, C.W.; investigation, Y.Z.; resources, Y.H.; data curation, Z.L.; writing—original draft preparation, J.J.; writing—review and editing, J.J., C.W., Y.Z., Y.L., Y.H., Z.L., J.Y. and F.H.; funding acquisition, F.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 62171467, and the Natural Science Foundation of Hebei Province, grant number F2021506004.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Hu, Y.; Xu, C.; Li, Z.; Lei, F.; Feng, B.; Chu, L.; Nie, C.; Wang, D. Detail enhancement multi-exposure image fusion based on homomorphic filtering. *Electronics* **2022**, *11*, 1211. [[CrossRef](#)]
- Su, Y.; Tang, C.; Li, B.; Qiu, Y.; Zheng, T.; Lei, Z. Greyscale image encoding and watermarking based on optical asymmetric cryptography and variational image decomposition. *J. Mod. Opt.* **2018**, *66*, 377–389. [[CrossRef](#)]

3. Kowsher, M.; Alam, M.A.; Uddin, M.J.; Ahmed, F.; Ullah, M.W.; Islam, M.R. Detecting third umpire decisions & automated scoring system of cricket. In Proceedings of the 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 11–12 July 2019; pp. 1–8.
4. Liu, Y.; Chen, X.; Cheng, J.; Peng, H.; Wang, Z. Infrared and visible image fusion with convolutional neural networks. *Int. J. Wavelets Multiresolut. Inf. Process.* **2018**, *16*, 1850018. [[CrossRef](#)]
5. Liu, Y.; Chen, X.; Ward, R.K.; Wang, Z.J. Image fusion with convolutional sparse representation. *IEEE Signal Process. Lett.* **2016**, *23*, 1882–1886. [[CrossRef](#)]
6. Luo, X.; Li, X.; Wang, P.; Qi, S.; Guan, J.; Zhang, Z. Infrared and visible image fusion based on NSCT and stacked sparse autoencoders. *Multimed. Tools Appl.* **2018**, *77*, 22407–22431. [[CrossRef](#)]
7. Hui, L.; Wu, X.J.; Kittler, J. Infrared and visible image fusion using a deep learning framework. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018.
8. Cardone, D.; Spadolini, E.; Perpetuini, D.; Filippini, C.; Chiarelli, A.M.; Merla, A. Automated warping procedure for facial thermal imaging based on features identification in the visible domain. *Infrared Phys. Technol.* **2020**, *112*, 103595. [[CrossRef](#)]
9. Singh, S.; Gyaourova, A.; Bebis, G.; Pavlidis, I. *Infrared and visible image fusion for face recognition. Biometric Technology for Human Identification*; SPIE: Reno, NV, USA, 2004; Volume 5404, pp. 585–597. [[CrossRef](#)]
10. Wang, N.Y.; Wang, W.L.; Guo, X.R. A new image fusion method based on improved PCNN and multiscale decomposition. *Adv. Mater. Res.* **2014**, *834–836*, 1011–1015. [[CrossRef](#)]
11. Zhu, J.; Jin, W.; Li, L.; Han, Z.; Wang, X. Multiscale infrared and visible image fusion using gradient domain guided image filtering. *Infrared Phys. Technol.* **2018**, *89*, 8–19. [[CrossRef](#)]
12. Ma, J.; Zhou, Z.; Wang, B.; Zong, H. Infrared and visible image fusion based on visual saliency map and weighted least square optimization. *Infrared Phys. Technol.* **2017**, *82*, 8–17. [[CrossRef](#)]
13. Zhang, P.; Yuan, Y.; Fei, C.; Pu, T.; Wang, S. Infrared and visible image fusion using co-occurrence filter. *Infrared Phys. Technol.* **2018**, *93*, 223–231. [[CrossRef](#)]
14. Duan, C.; Wang, Z.; Xing, C.; Lu, S. Infrared and visible image fusion using multi-scale edge-preserving decomposition and multiple saliency features. *Optik* **2020**, *228*, 165775. [[CrossRef](#)]
15. Candes, E.J.; Li, X.; Ma, Y.; Wright, J. Robust principal component analysis? *arXiv* **2009**, arXiv:0912.3599. [[CrossRef](#)]
16. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2011**, *3*, 1–122. [[CrossRef](#)]
17. Jiang, Y.; Wang, M. Image fusion using multiscale edge-preserving decomposition based on weighted least squares filter. *IET Image Process.* **2014**, *8*, 183–190. [[CrossRef](#)]
18. Salehi, H. Image de-speckling based on the coefficient of variation, improved guided filter, and fast bilateral filter. *Int. J. Image Graph.* **2021**, *21*, 2250036. [[CrossRef](#)]
19. Li, S.; Kang, X.; Hu, J. Image fusion with guided filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875.
20. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [[CrossRef](#)]
21. Donoho, D.L.; Tsaig, Y.; Drori, I.; Starck, J.L. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **2012**, *58*, 1094–1121. [[CrossRef](#)]
22. Yang, F.; Li, J.; Xu, S.H.; Pan, G.F. The research of a video segmentation algorithm based on image fusion in the wavelet domain. In Proceedings of the 5th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Smart Structures and Materials in Manufacturing and Testing, Dalian, China, 26–29 April 2010; Volume 7659, pp. 279–285.
23. Shreyamsha Kumar, B.K. Image fusion based on pixel significance using cross bilateral filter. *Signal Image Video Process.* **2015**, *9*, 1193–1204. [[CrossRef](#)]
24. Ma, J.; Chen, C.; Li, C.; Huang, J. Infrared and visible image fusion via gradient transfer and total variation minimization. *Inf. Fusion* **2016**, *31*, 100–109. [[CrossRef](#)]
25. Zhang, Y.; Zhang, L.; Bai, X.; Zhang, L. Infrared and visual image fusion through infrared feature extraction and visual information preservation. *Infrared Phys. Technol.* **2017**, *83*, 227–237. [[CrossRef](#)]
26. Bavirisetti, D.P.; Dhuli, R. Two-scale image fusion of visible and infrared images using saliency detection. *Infrared Phys. Technol.* **2016**, *76*, 52–64. [[CrossRef](#)]
27. Chibani, Y. Additive integration of SAR features into multispectral SPOT images by means of the à trous wavelet decomposition. *ISPRS J. Photogramm. Remote Sens.* **2006**, *60*, 306–314. [[CrossRef](#)]
28. Xydeas, C.S.; Pv, V. Objective image fusion performance measure. *Electron. Lett.* **2000**, *56*, 181–193. [[CrossRef](#)]
29. Chen, Y.; Blum, R.S. A new automated quality assessment algorithm for night vision image fusion. In Proceedings of the 2007 41st Annual Conference on Information Sciences and Systems, Baltimore, MD, USA, 14–16 March 2007; pp. 518–523.
30. Qu, G.; Zhang, D.; Yan, P. Information measure for performance of image fusion. *Electron. Lett.* **2002**, *38*, 313–315. [[CrossRef](#)]
31. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
32. Jagalingam, P.; Hegde, A.V. A Review of Quality Metrics for Fused Image. *Aquat. Procedia* **2015**, *4*, 133–142. [[CrossRef](#)]