

Article

Entropy-Based Image Fusion with Joint Sparse Representation and Rolling Guidance Filter

Yudan Liu ¹, Xiaomin Yang ^{1,*} , Rongzhu Zhang ¹, Marcelo Keese Albertini ² , Turgay Celik ³ and Gwanggil Jeon ^{4,5,*} 

¹ College of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China; shanke27@126.com (Y.L.); zhang_rz@scu.edu.cn (R.Z.)

² Department of Computer Science, Federal University of Uberlandia, Uberlandia, MG 38408-100, Brazil; albertini@ufu.br

³ School of Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg 2000, South Africa; celikturgay@gmail.com

⁴ School of Electronic Engineering, Xidian University, Xi'an 710071, China

⁵ Department of Embedded Systems Engineering, Incheon National University, Incheon 22012, Korea

* Correspondence: arielyang@scu.edu.cn (X.Y.); gjeon@inu.ac.kr (G.J.)

Received: 19 December 2019; Accepted: 16 January 2020; Published: 18 January 2020



Abstract: Image fusion is a very practical technology that can be applied in many fields, such as medicine, remote sensing and surveillance. An image fusion method using multi-scale decomposition and joint sparse representation is introduced in this paper. First, joint sparse representation is applied to decompose two source images into a common image and two innovation images. Second, two initial weight maps are generated by filtering the two source images separately. Final weight maps are obtained by joint bilateral filtering according to the initial weight maps. Then, the multi-scale decomposition of the innovation images is performed through the rolling guide filter. Finally, the final weight maps are used to generate the fused innovation image. The fused innovation image and the common image are combined to generate the ultimate fused image. The experimental results show that our method's average metrics are: mutual information (MI)—5.3377, feature mutual information (FMI)—0.5600, normalized weighted edge preservation value ($Q^{AB/F}$)—0.6978 and nonlinear correlation information entropy ($NCIE$)—0.8226. Our method can achieve better performance compared to the state-of-the-art methods in visual perception and objective quantification.

Keywords: image entropy; joint bilateral filter; image fusion; rolling guidance filter; joint sparse representation; multi-scale decomposition

1. Introduction

Image fusion combines multiple source images of the same scene together to make the fusion image more suitable for human visual perception or computer processing [1]. The source images are obtained from different sensors or imaging conditions. Each source image contains redundant and complementary information about the scene. The purpose of image fusion is to integrate the redundant and complementary information to make the fused image contain more relevant information specific to an application or task. Image fusion technology has good application prospects. It has been applied in many fields. In the medical field, we can reconstruct fused images by fusing multiple medical images from different sensors. The fused images are able to provide complementary information for medical analysis, enabling doctors to diagnose more quickly and accurately [2]. Moreover, surveillance is a typical application of image fusion [3]. Using an infrared–visible image fusion, a surveillance system can work effectively all day.

Image fusion methods based on transform domain can be divided into two types: multi-scale decomposition and sparse representation [4]. Multi-scale decomposition (MSD) is a very classical approach to performing image fusion. The basic process of MSD-based methods contains three steps. First, source images are converted into a specific transform domain, which typically decomposes the source images into multi-scale representations. Then, different fusion strategies are adopted for each scale level to acquire the multi-scale representation of the fused image in the transform domain. Finally, the corresponding inverse transformation is utilized to reconstruct the fused image. Traditional MSD methods can be divided into two categories. One includes the pyramid-based methods such as Laplacian pyramid (LP) [5] and gradient pyramid (GP) [6]. The other contains the wavelet transform-based methods, such as discrete wavelet transform (DWT) [7] and dual-tree complex wavelet transform (DTCWT) [8]. In addition, there are some new MSD methods, such as non-subsampled contourlet transform (NSCT) [9], shift-invariant shearlet transform (SIST) [10], non-subsampled shearlet transform (NSST) [11] and complex shearlet transform (CST) [12]. In recent years, the edge-preserving filter based-MSD has been a hot research direction. Edge-preserving smoothing filters, such as the guided filter (GF) [13], joint bilateral filter (JBF) [14] and rolling guidance filter (RGF) [15], can avoid ringing artifacts, since they do not blur strong edges in the decomposition process. Li et al. put forward a method of image fusion using GF [16]. Chen et al. proposed an infrared-visible image fusion method combining RGF and multi-directional decomposition [17]. Jian et al. combined RGF and JBF together for image fusion [18]. Although these methods achieve quite good performance on many types of images, they still have the following disadvantages: (1) the redundant information between source images leads to low information entropy of fused images; (2) contrast of fused images is likely to decrease.

Sparse representation (SR) is another classic image processing method. SR conforms to the physiological characteristics of human vision [19]. Besides, it has robustness to additive noise [20]. SR has been successfully applied in many image processing fields, such as image classification [21], image denoising [22], image communication [23] and image super-resolution [24]. Yang and Li [25] were the first ones to utilize SR to deal with image fusion problem. The general flow of SR-based methods contains three steps. First, an appropriate over-complete sparse dictionary is constructed through a mathematical model or example learning. Second, the sparse representations of the source images on the given sparse dictionary are obtained. The sparse representations are usually sets of sparse coefficient vectors. Finally, the corresponding vectors are fused to get the coefficient vector set of the fused image. The fused image is reconstructed with the same dictionary. Recently, a new SR-based method called joint sparse representation (JSR) was applied to image fusion successfully. The idea of JSR is to use sparse representation to divide two source images into a common part and two innovation parts. The common part contains the redundant information shared by the two source images, while the two innovation parts represent the complementary information of each source image. After that, different fusion rules are used to fuse the two kinds of part separately. JSR preserves the advantages of the SR method while eliminating the correlation between source images so that the fused image is less affected by the redundant information. Yu et al. proposed a JSR-based approach to carry out image denoising and fusion simultaneously [26]. Ma et al. combined JSR and optimum theory to address the multi-focus image fusion problem [27]. However, including JSR, all SR-based methods generally have the following disadvantages: (1) an over-complete dictionary may result in visual artifacts in the reconstructed image; (2) simple fusion strategy for sparse coefficient vectors leads to spatial inconsistency.

In order to make full use of the advantages and overcome the shortcomings of the above methods, we propose a new image fusion method combining JSR and MSD. Specifically, first, JSR is used to decompose two source images into a common image and two innovation images. Then, the innovation images are sent to RGF-based MSD fusion framework to obtain the fused innovation image. Finally, the fused innovation image and the common image are combined to obtain the fused image.

The main contributions of our work are as follows:

1. To improve the low information entropy caused by the redundant information between source images, only innovation images are performed edge-preserving MSD through RGF.
2. To suppress the artifacts that may be brought by JSR, weight maps are used to balance the contribution of innovation images.
3. To make fused images have high contrast, innovation images are used to guide the optimization of the weight maps.
4. To ensure the spatial consistency of fused images, the fusion of innovation images is performed according to optimized weight maps.

2. Related Work

2.1. Joint Sparse Representation

JSR is a novel method with which to perform image fusion. As with SR, JSR has two key issues: (1) sparse dictionary construction; (2) joint sparse coding to obtain coefficients. For JSR, constructing an over-complete sparse dictionary is exactly the same as in SR. In this study, we used K-SVD [28] to train the dictionary.

The biggest difference between JSR and SR is the method of sparse coding. In image processing, the objects for sparse coding are overlapping small image patches of source images. These patches are extracted by the sliding-window technique. The small patches of all the source images at the same position constitute a patch set. SR encodes each patch separately, while JSR simultaneously encodes all patches in the same set. JSR supposes that every patch consists of a common component and an innovation component. The common component is shared by all patches in the same set, while each patch has its own innovation component. There are two strategies for selecting the dictionaries for JSR. One strategy is to use one fixed dictionary for all the components for sparse encoding and reconstruction [26]. This strategy has low training cost, is easy to operate, and is suitable for the sparse representation of multiple types of images. The other strategy is to use a fixed dictionary for common components and an adaptive dictionary for innovation components for sparse coding and reconstruction [27]. This strategy may yield better results, but comes with additional computational costs. Our method uses the first strategy, which is to use a fixed dictionary for all components. Since there are always two source images for fusion, suppose each patch set has W patches. Given a flatten patch $v_i \in R^n$ ($1 \leq i \leq 2$) extracted from a source image and an over-complete dictionary $D \in R^{n \times m}$ ($n < m$), where n represents the length of the vector and m represents the number of atoms in the dictionary, respectively. The goal of JSR is to estimate a common sparse vector $x^C \in R^m$ and innovation sparse vectors $x_i^I \in R^m$ ($1 \leq i \leq 2$) with only a few nonzero entries, such that

$$v_i \approx Dx^C + Dx_i^I, \quad (1)$$

where Dx^C and Dx_i^I represent the common component and the innovation component of v_i , respectively.

When performing sparse coding, JSR needs to first concatenate the source image patches and the dictionary separately. The encoded sparse coefficient vectors are also concatenated. Let

$$V = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} D & D & 0 \\ D & 0 & D \end{bmatrix}, \quad X = \begin{bmatrix} x^C \\ x_1^I \\ x_2^I \end{bmatrix},$$

where V denotes the concatenated source patch, \tilde{D} denotes the concatenated dictionary and X denotes the concatenated sparse coefficient vector. The joint version of Equation (1) can be defined as follows:

$$V \approx \tilde{D}X. \quad (2)$$

The goal of JSR is to find an approximate optimal solution of X . This problem can be formulated as

$$\arg \min_X \|X\|_0 \quad s.t. \quad \|\tilde{D}X - V\|_2 \leq \varepsilon, \quad (3)$$

where ε is a sparse reconstruction error.

Our proposed method uses JSR as an image decomposition method to reduce correlations between source images and increase the information entropy of a fused image. First, the two source images are divided into small overlapping image patches through the sliding window technique. The dimensions of the patches depends on the specific dictionary. Assuming the dictionary $D \in R^{n \times m}$, then the size of each patch should be $\sqrt{n} \times \sqrt{n}$. Every patch is rearranged to a column vector. Second, two patches located at the same position compose a patch set. JSR is performed independently on each patch set. For every patch set, a common sparse vector belonging to the entire set and two innovative sparse vectors belonging to each patch in the set can be obtained. Then, a common patch and two innovation patches can be reconstructed by the corresponding sparse vector for each set. Finally, all the common patches and the corresponding innovation patches are averaged in the same order they were selected during the sliding window step. A common image and two innovation images are generated. The correlation between the two innovation images is lower than that between the two source images. We define the process of obtaining a common image and innovation images from source images as JSR decomposition.

2.2. Rolling Guidance Filter

The purpose of multi-scale decomposition is to obtain images of different blur levels to make full use of the information contained in the source images. Recently, edge-preserving filter-based MSD methods have become the mainstream of research. These methods are able to preserve high-contrast edges and obvious structures while blurring the image. The state-of-the-art edge-preserving MSD method was proposed by Jian et al. [18]. It uses a rolling guidance filter (RGF).

A RGF can be seen as an extension of joint bilateral filter (JBF). JBF is first proposed for image denoising by Petschnigg et al. [14]. JBF accepts an input image and a guidance image as input. The content of the output image is similar to the input image, while the structures and edges of it are similar to the guidance image. First, the Gaussian kernel g_δ is given as:

$$g_\delta(p, q) = \exp\left(-\frac{\|p - q\|^2}{2\delta^2}\right), \quad (4)$$

where p and q index pixel coordinates in the image, and δ is the standard deviation. With Equation (4), given a guidance image G and an input image I_{in} , the definition of JBF is as follows:

$$I_{out}(p) = \frac{1}{U} \sum_{q \in N(p)} g_{\delta_s}(p, q) g_{\delta_r}(G(p), G(q)) I_{in}(q), \quad (5)$$

where I_{out} is the output image, $U = \sum_{q \in N(p)} g_{\delta_s}(p, q) g_{\delta_r}(G(p), G(q))$ is the normalization factor, $N(p)$ is the set of neighboring pixels of p , δ_s is the spatial standard deviation and δ_r is the range standard deviation. The function g_{δ_s} sets the weight in the spatial domain based on the distance between the pixels, while the function g_{δ_r} sets the weight on the range based on intensity differences. δ_s and δ_r control the spatial and range weights, respectively. For the sake of simplicity, we abbreviate Equation (5) as follows:

$$I_{out} = JBF(I_{in}, G, \delta_s, \delta_r), \quad (6)$$

where $JBF(\cdot)$ denotes the JBF process.

RGF was proposed by Zhang et al. [15]. The biggest feature of RGF is the ability to remove small structures while preserving the main content of the image. RGF is composed of two main steps: small

structure removal and edge recovery. As shown in Figure 1, RGF is an iterative process. Suppose J^t is the result in the t -th iteration and M is the number of total iterations, the iterative process of RGF is as follows:

$$J^t = JBF \left(I_{in}, J^{t-1}, \sigma_s, \sigma_r \right), \quad t = 1, 2, \dots, M, \tag{7}$$

where σ_s and σ_r denote the standard deviations of RGF in order to distinguish them from δ_s and δ_r of JBF.

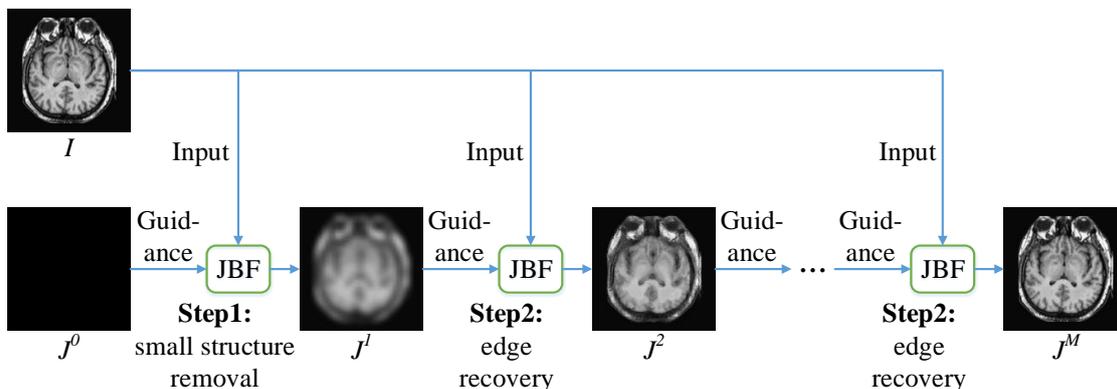


Figure 1. Rolling guidance filtering.

According to Equation (7), the iteration process of RGF is given as Algorithm 1.

Algorithm 1 : The iteration process of RGF.

Input: Input image I_{in} ; spatial standard deviation σ_s ; range standard deviation σ_r ; iteration number M .

- 1: Set J^0 as a constant image, i.e., $\forall p, J^0(p) = C$, where C is a constant value.
- 2: **for** $t = 1 : 1 : M$ **do**
- 3: $J^t = JBF(I_{in}, J^{t-1}, \sigma_s, \sigma_r)$.
- 4: **end for**

Output: Output image $I_{out} = J^M$.

Since J^0 is set as a constant image, the first iteration is equivalent to blurring I_{in} with a Gaussian filter with the standard deviation of σ_s to remove small structures. The remaining iterations are equivalent to continuous filtering of I_{in} by the joint bilateral filter. Specifically, in each iteration, JBF takes I_{in} as input and J^{t-1} as guidance, and σ_s and σ_r as standard deviations. Edges are recovered gradually during this process. After all the iterations are completed, J^M is the output I_{out} .

For simplicity, we denote the RGF filtering operation as follows:

$$I_{out} = RGF(I_{in}, \sigma_s, \sigma_r, M), \tag{8}$$

where $RGF(\cdot)$ denotes the RGF function.

3. Proposed Method

Our proposed method is shown in Figure 2, which contains four stages: (1) JSR decomposition; (2) weight map construction; (3) multi-scale decomposition; (4) fused image reconstruction.

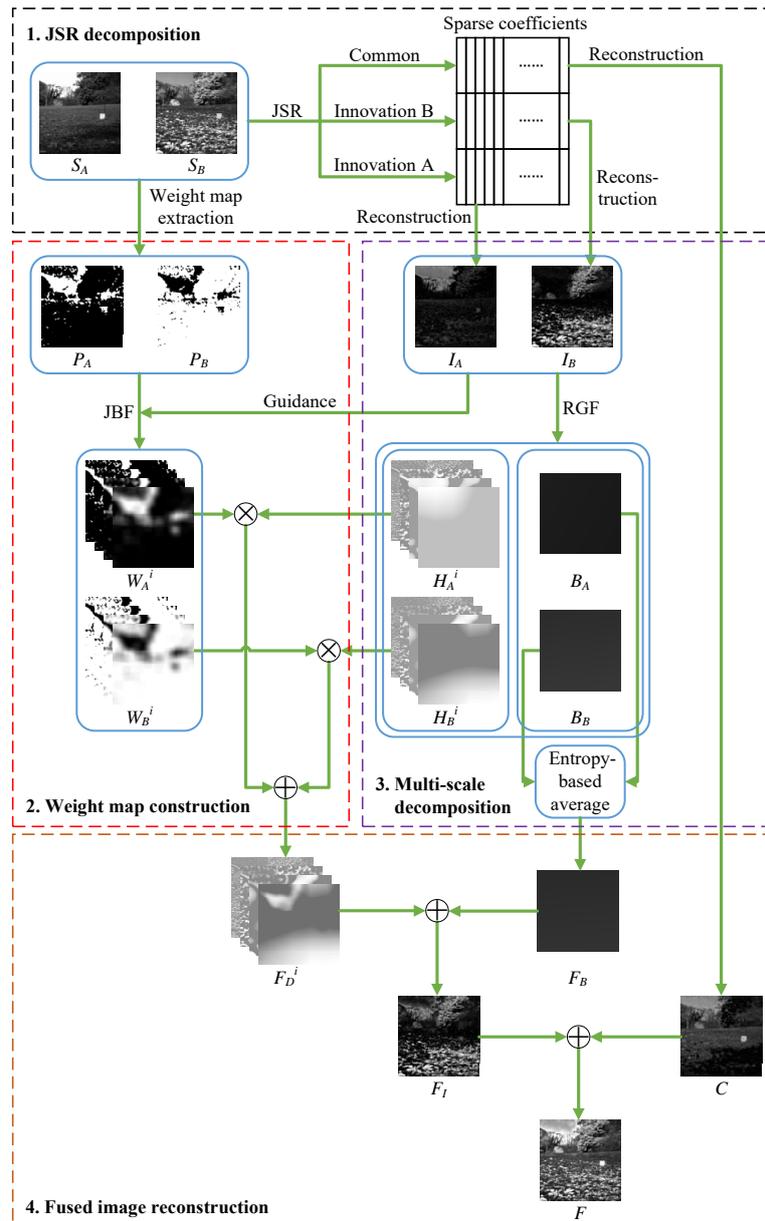


Figure 2. The schematic diagram of our proposed fusion method.

3.1. JSR Decomposition

We first use JSR decomposition to decompose two source images S_A and S_B into a common image C and two innovation images I_A and I_B . An example of JSR decomposition is given in Figure 3.

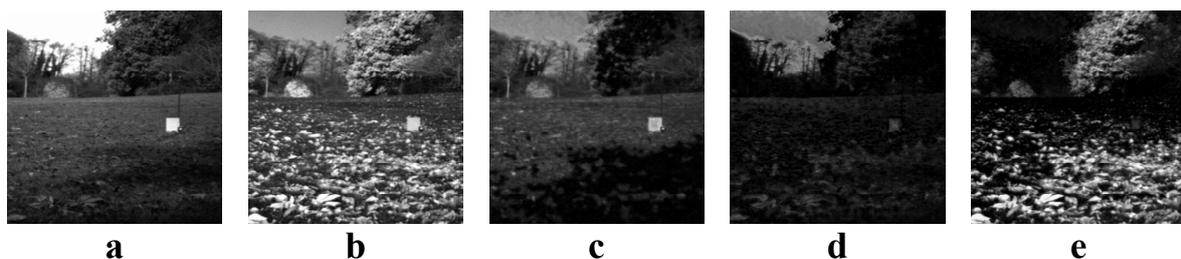


Figure 3. An example of JSR decomposition. (a,b) Source images; (c) The common image; (d) The innovation image of (a); (e) The innovation image of (b).

3.2. Weight Map Construction

The weight maps are references for fusing detail layers. In order to make the information accurate and complete, weight maps are generated from the source images. First, S_A and S_B are processed with Kirsch operator to obtain the saliency maps R_A and R_B . Next, the initial weight maps P_A and P_B are obtained through a pixel-by-pixel comparison of the saliency maps R_A and R_B defined as follows:

$$P_A(q) = \begin{cases} 0, & \text{if } R_A(q) \leq R_B(q) \\ 1, & \text{otherwise} \end{cases}, \quad (9)$$

$$P_B(q) = 1 - P_A(q),$$

where q denotes pixel coordinates. The order in which the two source images are considered does not affect the initial weight maps. At each pixel position q , the initial weight of the one with the larger saliency value is set to 1 and the other is set to 0. If $R_A(q) = R_B(q)$, then either of the two initial weights can be set to 1 and the other is set to 0, which does not affect the result. However, it is almost impossible for two saliency values to be exactly equal.

Finally, JBF is used to filter P_A and P_B , and the final weight maps are obtained. This step is as follows:

$$W_A^i = JBF(P_A, I_A, \delta_s^i, \delta_r^i), \quad i = 1, 2, \dots, K-1, \quad (10)$$

$$W_B^i = JBF(P_B, I_B, \delta_s^i, \delta_r^i), \quad i = 1, 2, \dots, K-1.$$

By adjusting δ_d^i and δ_r^i , weight maps for detail layers are optimized. The innovation images I_A and I_B are used as guidance images to enhance the difference between weight maps. This results in higher contrast in the fused image. Besides, this step makes the weight values same for the pixels with similar brightness and next to each other so that the problems caused by spatial consistency can be avoided [16].

3.3. Multi-Scale Decomposition

RGF is the key to performing this stage. When M and σ_r are set to constants, changes of σ_s can achieve different blur levels of rolling guidance filtering. As σ_s increases, the output of RGF becomes more blurred, which means it contains more low-frequency components. Therefore, the output of using the largest σ_s for RGF should be regarded as the base layer. Outputs using various smaller σ_s are sequentially differentiated and the difference values are regarded as the detail layers.

Now we take I_A as an example to give the concrete multi-scale decomposition method. First, I_A is normalized to range $[0, 1]$ and blurred into $K-1$ levels through RGF; this process is described by Equation (11).

$$O_A^i = RGF(I_A, \sigma_s^i, \sigma_r, M), \quad i = 1, 2, \dots, K-1, \quad (11)$$

where O_A^i denotes the outputs of RGF using different σ_s^i . Let $\sigma_s^{i+1} > \sigma_s^i$; then, the smooth level is $O_A^{i+1} > O_A^i$. Furthermore, in order to give a unified equation form, let $O_A^0 = I_A$. Finally, base layers B_A and $K-1$ detail layers H_A^i ($i = 1, 2, \dots, K-1$) can be obtained by Equation (12)

$$B_A = O_A^{K-1}, \quad (12)$$

$$H_A^i = O_A^{i-1} - O_A^i, \quad i = 1, 2, \dots, K-1.$$

The same operation is performed to I_B to get B_B and H_B^i ($i = 1, 2, \dots, K-1$).

3.4. Fused Image Reconstruction

There are four steps to reconstruct the final fused image: (1) reconstruct the fused base layer F_B ; (2) reconstruct the fused detail layers F_H^i ; (3) reconstruct the fused innovation image F_I by combining F_B and F_H^i ; (4) reconstruct the final fused image F by combining F_I and the common image C .

First, the fused base layer F_B is reconstructed by entropy-based average. The base layer contains the low-frequency information of the image, which is equivalent to the average value of the image. Fusing base layers by the traditional simple averaging method may cause the fused image have low contrast and information entropy. Since the global variance of a base layer represents its overall contrast and amount of information, we regard global variance of a base layer as its entropy. Then, an entropy-based average method is used to fuse the base layers. Specifically, the entropies E_A and E_B of the two base layers B_A and B_B are calculated by:

$$\begin{aligned} E_A &= \text{var}(B_A), \\ E_B &= \text{var}(B_B), \end{aligned} \quad (13)$$

where $\text{var}(\cdot)$ denotes the global variance function. Second, the fused base layer F_B is reconstructed by weighted average using E_A and E_B as weights:

$$F_B = \frac{1}{E_A + E_B} (E_A B_A + E_B B_B). \quad (14)$$

Second, the fused detail layers F_H^i are reconstructed. The detail layers are simply multiplied by the corresponding weight maps and summed up to achieve the fusion. This process is described as follows:

$$F_H^i = H_A^i W_A^i + H_B^i W_B^i, \quad i = 1, 2, \dots, K-1. \quad (15)$$

Then, with the fused detail layers F_H^i and base layer F_B , the fused innovation image F_I can be obtained as follows:

$$F_I = F_B + \sum_{i=1}^{K-1} F_H^i. \quad (16)$$

Finally, the fused innovation image F_I and the common image C are merged together to obtain the ultimate fused image F :

$$F = F_I + C. \quad (17)$$

3.5. Workflow of Our Proposed Method

The workflow of our proposed method can be summarized as follows. Consider two source images S_A and S_B , dictionary D , the decomposition level K and the parameters of JBF and RGF. First, JSR is used to decompose the two source images into one common image C and two innovation images I_A and I_B . Second, the Kirsch operator is used to extract saliency maps P_A and P_B from source images. Regarding the innovation images as guidance, JBF is applied to the saliency maps to obtain weight maps W_A^i and W_B^i ($i = 1, 2, \dots, K-1$). Then, K-level multi-scale decomposition of the innovation images is performed by RGF to obtain the detail layers H_A^i and H_B^i ($i = 1, 2, \dots, K-1$) and the base layers B_A and B_B . Finally, the detail layers are fused according to the corresponding weight maps. The base layers are fused by an entropy-based fusion rule. By summing the fused detail layers F_H^i ($i = 1, 2, \dots, K-1$) and the fused base layer F_B , the fused innovation image F_I is obtained. The last step is to add F_I to the common image C , and the fused image is obtained.

When doing addition and multiplication, the problem of dynamic range of images needs to be addressed. In our workflow, multiplication and addition occur during the fusion of the detail and base layers. For the fusion of detail layers, two detail layers on the same decomposition level are respectively multiplied with their corresponding weight maps, and the products are added to obtain the fused detail layer. The two corresponding weight maps are complementary, that is, the sum of

two weights at the same pixel position is very close to 1. Therefore, the value range of the fused detail layers hardly changes. For the fusion of base layers, since the two weights are also complementary, the value range of the fused base layer does not change either. By adding all the fused detail layers and the fused base layer together, the fused image can be reconstructed. This final addition may cause a small number of pixels to be out of the reasonable range of $[0, 255]$. For these pixels, the values less than 0 are set to 0 and the values greater than 255 are set to 255. Finally, the value of each pixel is rounded into an integer.

The pseudo code of our proposed method is shown as Algorithm 2.

Algorithm 2 : Pseudo code of our proposed method.

Input: Source images S_A, S_B ; Dictionary D ; Decomposition level K ; JBF parameters δ_s^i, δ_r^i ; RGF

parameters σ_s^i, σ_r, M .

- 1: Use D to perform JSR decomposition on S_A, S_B to get C, I_A, I_B .
- 2: Process S_A, S_B with Kirsch operator to get R_A, R_B .
- 3: **for** q in pixel coordinate range of R_A **do**
- 4: **if** $R_A(q) \leq R_B(q)$ **then**
- 5: $P_A(q) = 0, P_B(q) = 1$.
- 6: **else**
- 7: $P_A(q) = 1, P_B(q) = 0$.
- 8: **end if**
- 9: **end for**
- 10: **for** $i = 1 : 1 : K - 1$ **do**
- 11: $W_A^i = JBF(P_A, I_A, \delta_s^i, \delta_r^i), W_B^i = JBF(P_B, I_B, \delta_s^i, \delta_r^i)$.
- 12: $O_A^i = RGF(I_A, \sigma_s^i, \sigma_r, M), O_B^i = RGF(I_B, \sigma_s^i, \sigma_r, M)$.
- 13: **end for**
- 14: $B_A = O_A^{K-1}, B_B = O_B^{K-1}$.
- 15: $O_A^0 = I_A, O_B^0 = I_B$.
- 16: **for** $i = 1 : 1 : K - 1$ **do**
- 17: $H_A^i = O_A^{i-1} - O_A^i, H_B^i = O_B^{i-1} - O_B^i$.
- 18: **end for**
- 19: $E_A = var(B_A), E_B = var(B_B)$.
- 20: $F_B = \frac{1}{E_A + E_B} (E_A B_A + E_B B_B)$.
- 21: $F_H = \sum_{i=1}^{K-1} H_A^i W_A^i + H_B^i W_B^i$.
- 22: $F_I = F_B + F_H$.
- 23: $F = F_I + C$.

Output: Fused image F .

4. Experimental Results and Analysis

4.1. Experimental Settings and Objective Evaluations

We tested all the methods using four categories of images, with four sets of source images in each category. Specifically, there are infrared-visible images shown in Figure 4, medical images shown in Figure 5, multi-focus images shown in Figure 6 and remote sensing images shown in Figure 7. The images we used for our experiments are downloadable from the following website: <https://sites.google.com/view/durgaprasadbavirisetti/datasets>.

The default parameters in our method are set according to [18]. Specifically, for the parameters of JBF, we set $\delta_s^i = \{1, 3, 10, 30\}$ ($i = 1, 2, 3, 4$) and $\delta_r^i = \{10, 30, 100, 300\}$ ($i = 1, 2, 3, 4$) as defaults. For the RGF part, we set $\sigma_s^i = \{3, 18, 108, 648\}$ ($i = 1, 2, 3, 4$), $\sigma_r = 0.2$ and $M = 5$ as a default. The other

parameters are discussed in the following experiments. All the experiment's programs were generated in Matlab R2019a (MathWorks, Natick, MA, USA) on an Intel(R) Core(TM)i5-6400CPU (Intel, Santa Clara, CA, USA) @ 2.70 GHz with 8.00 GB RAM.

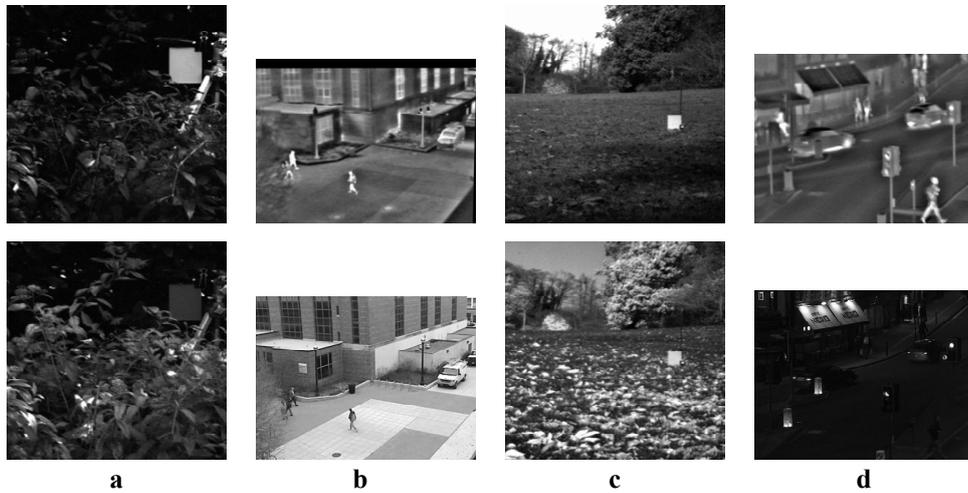


Figure 4. Infrared-visible image sets. (a–d) Four sets of infrared–visible source images.

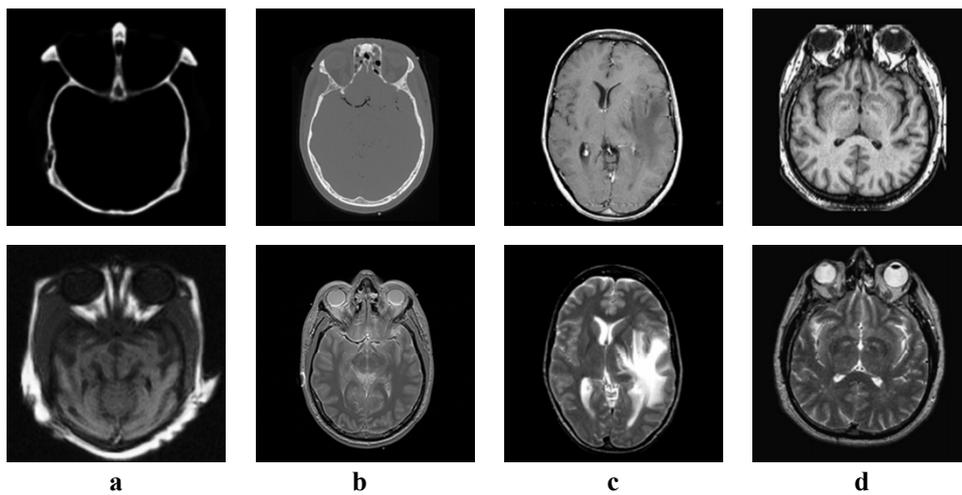


Figure 5. Medical image sets. (a–d) Four sets of medical source images.



Figure 6. Multi-focus image sets. (a–d) Four sets of multi-focus source images.

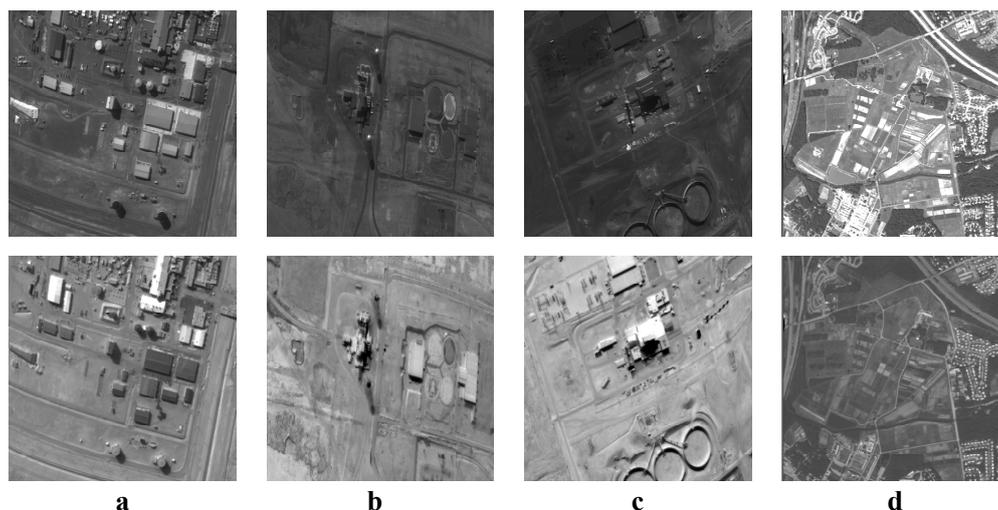


Figure 7. Remote sensing image sets. (a–d) Four sets of remote sensing source images.

The objective evaluation has a certain reference value for judging the quality of image fusion. In our experiments, four metrics were used to measure the information entropy and visual quality of the results of different methods:

1. Mutual information (MI) [29] based on Shannon entropy and relative entropy. It measures the correlation between the source image and the fused image to indicate how much information is retained.
2. Feature mutual information (FMI) [30] indicates the entropy of features in fused image. It measures the amount of information in image features carried from the source images to the fused image. Besides, it is a non-reference image fusion metric.
3. The normalized weighted edge preservation value ($Q^{AB/F}$) [31] measures the visual information quality of the fusion, and more edge information can lead to higher values for this metric.
4. Nonlinear correlation information entropy ($NCIE$) [32] is based on nonlinear joint entropy. It measures the general correlation between the source images and the fused image.

Higher values of the above four metrics demonstrate a better fusion effect. The codes of the metrics are provided by Qu et al. [33].

4.2. Discussions about Parameters

4.2.1. Size of JSR Dictionary

This experiment tested the effect of different JSR dictionary sizes on the fusion performance. Suppose the dictionary $D \in R^{n \times m}$ ($n < m$); then n affects the size of image patches while m affects the completeness of the dictionary. A larger n indicates a larger size of image patches, which leads to an enlarged window to perform joint sparse representation, while a larger m increases the completeness of the dictionary. It is important to find a suitable dictionary size. n and m values which are too small cause large reconstruction errors and loss of details, while overly large n and m easily cause artifacts and take more time.

In this experiment, the decomposition level K was set to 5 while other parameters were set to default values. All the dictionaries used in this experiment were trained with 100,000 natural image patches for 180 iterations by the K-SVD algorithm, as suggested in [34]. All 16 sets of source images were tested. The average values of the objective metrics and the average time cost were compared. First, m was fixed to 512 and n was set to $\{36, 64, 100\}$, respectively. Then, n was fixed to 64 and m was set to $\{128, 256, 512\}$, respectively. The results are shown in Tables 1 and 2.

Table 1. Objective evaluation of different n values for the JSR dictionary. The best and second best results of each metric are marked in red and bold, respectively.

Metric	36×512	64×512	100×512
<i>MI</i>	5.3075	5.3377	5.3179
<i>FMI</i>	0.5650	0.5600	0.5556
$Q^{AB/F}$	0.6971	0.6978	0.6974
<i>NCIE</i>	0.8224	0.8226	0.8225
Time	163.74	398.33	1182.71

Table 2. Objective evaluation of different m values for the JSR dictionary. The best and second best results of each metric are marked in red and bold, respectively.

Metric	64×128	64×256	64×512
<i>MI</i>	5.2805	5.3309	5.3377
<i>FMI</i>	0.5538	0.5638	0.5600
$Q^{AB/F}$	0.6936	0.6973	0.6978
<i>NCIE</i>	0.8223	0.8225	0.8226
Time	176.07	251.15	398.33

It can be seen that the size of 64×512 achieved the highest *MI*, $Q^{AB/F}$ and *NCIE* in both experiments, and its *FMI* values were both the second highest. At the same time, its time cost was not too high compared to others. Therefore, it is reasonable to choose the size of the JSR dictionary as 64×512 .

4.2.2. Number of Decomposition Level

This experiment tested the effect of different decomposition levels K on the fusion performance. For the innovation image of each source image, RGF decomposition generates $K - 1$ detail layers and one base layer. The base layer can be considered as the last detail layer with the highest degree of blur. At the same time, $K - 1$ weight maps are generated by JBF to guide the fusion of detail layers. To ensure that at least one detail layer exists, the minimum value of K should be 2.

In this experiment, K was set to $\{2, 3, 4, 5\}$, respectively. The other parameters were set to default. The 64×512 dictionary was used for JSR as mentioned before. All 16 sets of source images were tested. The average values of the objective metrics were compared. The results are shown in Figure 8.

As K increases, *MI* and *NCIE* also increase. The highest values of *MI* and *NCIE* were at $K = 5$, but both the growth rates from $K = 4$ to $K = 5$ were low. Both the highest growth rates were from $K = 3$ to $K = 4$. *FMI* reached the highest value at $K = 2$. After reaching the highest value, the value dropped. *NCIE* reached the highest value at $K = 3$. Except from $K = 2$ to $K = 3$, the growth rate of *NCIE* was approximately equal to zero. Since two metrics reached the maximum value at $K = 5$, and the growth rates of all four metrics tended to be zero at $K = 5$, it is reasonable to choose $K = 5$ as the default decomposition level.

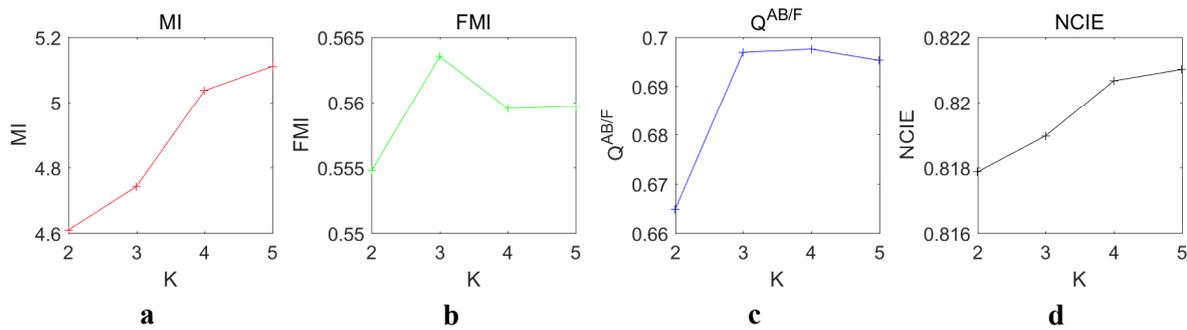


Figure 8. Objective evaluation of different decomposition level K . (a–d) The values of MI , FMI , $Q^{AB/F}$ and $NCIE$, respectively.

We also compared the calculation times required for different decomposition levels. The comparison results are shown in Table 3.

Table 3. Time cost of different decomposition levels.

K	2	3	4	5
Time	374.70	376.97	381.58	398.33

As K increases, the time cost increases at a very slow rate. Compared with the time cost of JSR decomposition, the time cost of MSD using RGF is very small. The 5-level decomposition is only about 6% slower than the 2-level decomposition. It is worthwhile to trade these time costs for performance improvements. Overall, it is reasonable to choose $K = 5$ as the default decomposition level.

4.3. Validity of Our Combination Strategy

In this experiment, our proposed method was compared to original JSR-based [26] and RGF-based [18] image fusion methods to validate the effectiveness of our combination strategy. Our method adopted the aforementioned default parameters, and the JSR-based and RGF-based methods adopted the default parameters from their papers. All four categories of source images were tested. The average values of the objective metrics of each category were compared separately. Some examples of the fused images are shown in Figure 9. The average values of the objective metrics are shown in Table 4.

According to Figure 9, RGF and our method perform similarly in terms of subjective effects. But the subjective effect of JSR is not as good as the other two. For fused images of JSR, the details in some images are smoothed, and some images have local artifacts. The objective metrics in Table 4 show that for images other than medical ones, our method is the best. This means that the fused images of our method have higher information entropy and visual quality. In the results of medical images, our method achieves best MI and $Q^{AB/F}$, while JSR achieves best FMI and $NCIE$. However, according to the medical image fusion example in Figure 9, our method generates clearer details than the JSR method.

Our method combines JSR and RGF for better image fusion performance, while JBF can be regarded as a special form of RGF. First, JSR decomposition is used to obtain two innovation images and a common image. The two innovation images contain the complementary information of the two source images, which have more information entropy and less redundancy. The complementary information is what really needs to be fused. Since the common image contains the redundant information shared by the two source images, it should be directly included in the fused image without modification. Second, the combination of JBF and RGF for image fusion was proven to be very effective. RGF is used for multi-scale edge-preserving decomposition, which makes the most use of details and textures. JBF is used to obtain the corresponding weight maps, which take spatial consistency well into account and reduces local artifacts [18]. To further improve the quality of fusion, the innovation images

generated by JSR were used as the input images of RGF and the guidance images of JBF. The MSD using RGF can extract complementary details of the innovation images at different levels. Meanwhile, using the innovation images as the guidance images of JBF can make the weight maps balance the contribution of the innovation images well. Then, different strategies are used to fuse the detail layers and the base layer to make the fused innovation image have high contrast and more details. Finally, adding the common image directly to the fused innovation image ensures that common information can be preserved.

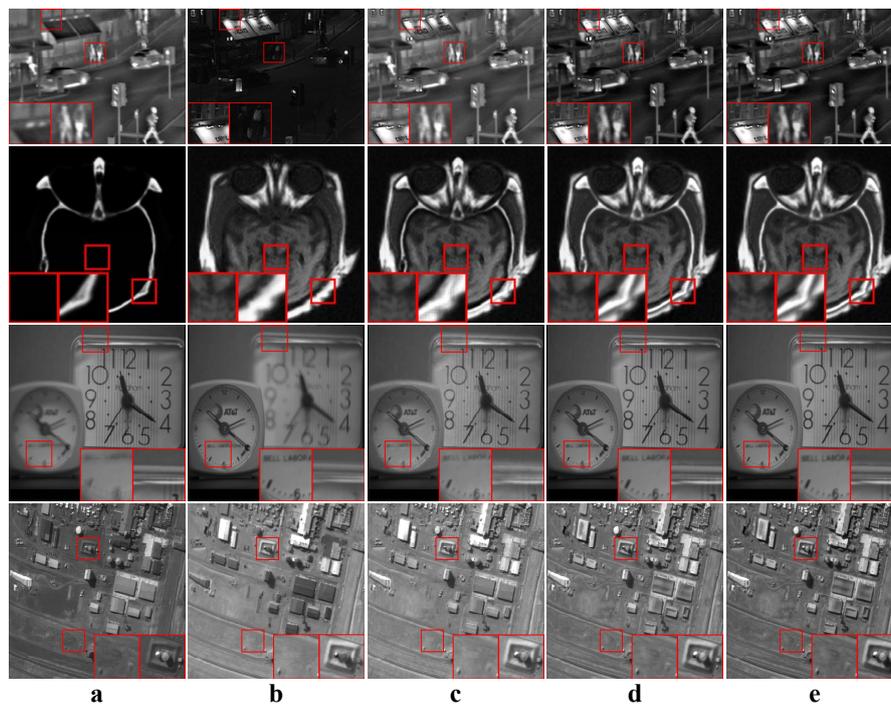


Figure 9. Some fused images of JSR, RGF and our proposed method. (a,b) Source images; (c) The fused results of JSR; (d) The fused results of RGF; (e) The fused results of our proposed method.

Table 4. Objective evaluation of JSR, RGF and our method. The best and second best results of each metric are marked in red and bold, respectively.

Category	Metric	JSR	RGF	OURS
Infrared-visible	MI	3.6597	3.7839	4.2105
	FMI	0.4946	0.5456	0.5477
	$Q^{AB/F}$	0.6130	0.6644	0.6652
	NCIE	0.8106	0.8113	0.8143
Medical	MI	4.1862	4.0194	4.2164
	FMI	0.5439	0.5278	0.5228
	$Q^{AB/F}$	0.6177	0.6768	0.6800
	NCIE	0.8133	0.8119	0.8130
Multi-focus	MI	6.9542	8.8911	8.9213
	FMI	0.5475	0.6316	0.6324
	$Q^{AB/F}$	0.7449	0.7890	0.7891
	NCIE	0.8316	0.8465	0.8467
Remote sensing	MI	2.9600	3.7494	4.0035
	FMI	0.4555	0.5337	0.5370
	$Q^{AB/F}$	0.5846	0.6508	0.6567
	NCIE	0.8082	0.8145	0.8165

Overall, our method is superior to the single JSR-based and RGF-based methods. This experiment proves the validity of our proposed combination strategy.

4.4. Comparison with Other Methods

In order to prove the superiority of the proposed method, we compare it with 13 other image fusion methods here. The comparative methods include the adaptive sparse representation (ASR) [35], the convolutional sparse representation (CSR) [36], curvelet transform (CVT) [37], dual-tree complex wavelet transform (DTCWT) [8], the gradient transfer fusion (GTF) [38], the hybrid multi-scale decomposition (H-MSD) [39], the convolutional neural network (CNN) [40], Laplacian pyramid (LP) [5], the general framework based on multi-scale transform and sparse representation (MSSR) [34], the multi-resolution singular value decomposition (MSVD) [41], nonsubsampling contourlet transform (NSCT) [9], the visual saliency map and weighted least square optimization (WLS) [42] and the fast filtering image fusion (FFIF) [43]. All these methods were given default parameters from their related papers. For our proposed method, the dictionary size was 64×512 , the decomposition level K was set to 5 and the other parameters were set to defaults.

In this experiment, all four categories of source images were tested. The average values of the objective metrics of each category were compared separately. Some examples of the fused images are shown in Figures 10–13. The average values of the objective metrics are shown in Tables 5–8.

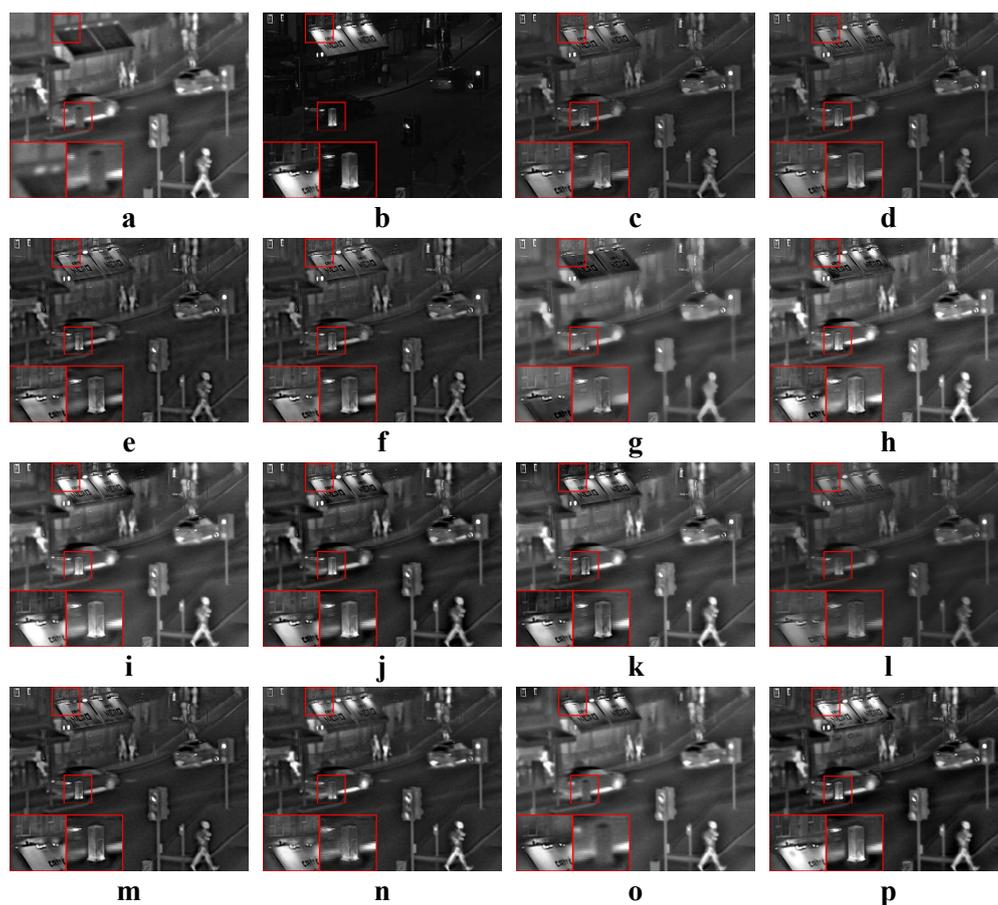


Figure 10. Examples of the fusion results of infrared-visible images. (a,b) Source images; (c–p) The fused results of ASR, CSR, CVT, DTCWT, GTF, H-MSD, CNN, LP, MSSR, MSVD, NSCT, WLS, FFIF, and our proposed method, respectively.

Table 5. Objective evaluation of infrared-visible image fusion. The best and second best results of each metric are marked in red and bold, respectively.

Metric	ASR	CSR	CVT	DTCWT	GTF	H-MSD	CNN
<i>MI</i>	2.7134	2.7878	2.2697	2.3902	2.5465	2.6970	2.9490
<i>FMI</i>	0.5202	0.4510	0.4596	0.4892	0.4874	0.4324	0.4818
$Q^{AB/F}$	0.5986	0.5890	0.5512	0.5796	0.4994	0.5686	0.6290
<i>NCIE</i>	0.8064	0.8066	0.8052	0.8055	0.8061	0.8064	0.8072
Metric	LP	MSSR	MSVD	NSCT	WLS	FFIF	OURS
<i>MI</i>	2.6575	3.4726	2.9739	2.4802	2.7887	4.9717	4.2105
<i>FMI</i>	0.5003	0.5044	0.3972	0.4988	0.4339	0.5775	0.5477
$Q^{AB/F}$	0.6366	0.6065	0.4123	0.6144	0.5574	0.6405	0.6652
<i>NCIE</i>	0.8062	0.8107	0.8072	0.8057	0.8065	0.8226	0.8143

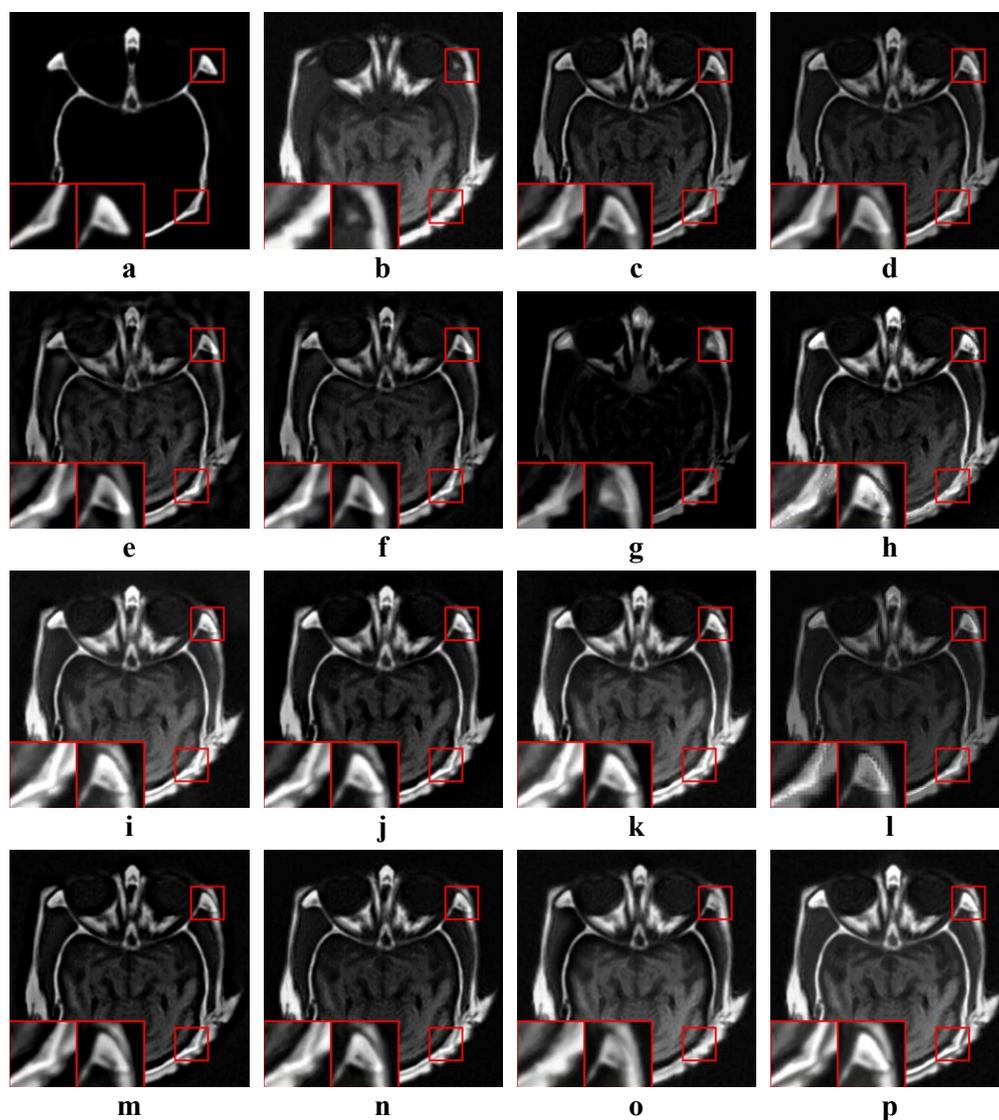


Figure 11. Examples of the fusion results of medical images. (a,b) Source images; (c–p) The fused results of ASR, CSR, CVT, DTCWT, GTF, H-MSD, CNN, LP, MSSR, MSVD, NSCT, WLS, FFIF, and our proposed method, respectively.

Table 6. Objective evaluation of medical image fusion. The best and second best results of each metric are marked in red and bold, respectively.

Metric	ASR	CSR	CVT	DTCWT	GTF	H-MSD	CNN
<i>MI</i>	3.4473	3.3705	2.6794	2.9084	2.9051	3.2624	3.5522
<i>FMI</i>	0.5638	0.5087	0.3534	0.4478	0.5125	0.4738	0.5152
$Q^{AB/F}$	0.6037	0.5976	0.5170	0.5488	0.4288	0.5639	0.6416
<i>NCIE</i>	0.8092	0.8090	0.8069	0.8075	0.8076	0.8086	0.8097
Metric	LP	MSSR	MSVD	NSCT	WLS	FFIF	OURS
<i>MI</i>	3.2668	3.6737	3.5279	3.1658	3.5519	4.6729	4.2164
<i>FMI</i>	0.5243	0.5406	0.4731	0.5063	0.4907	0.6086	0.5228
$Q^{AB/F}$	0.6384	0.6422	0.4713	0.6220	0.5914	0.6535	0.6800
<i>NCIE</i>	0.8085	0.8102	0.8097	0.8082	0.8097	0.8151	0.8130

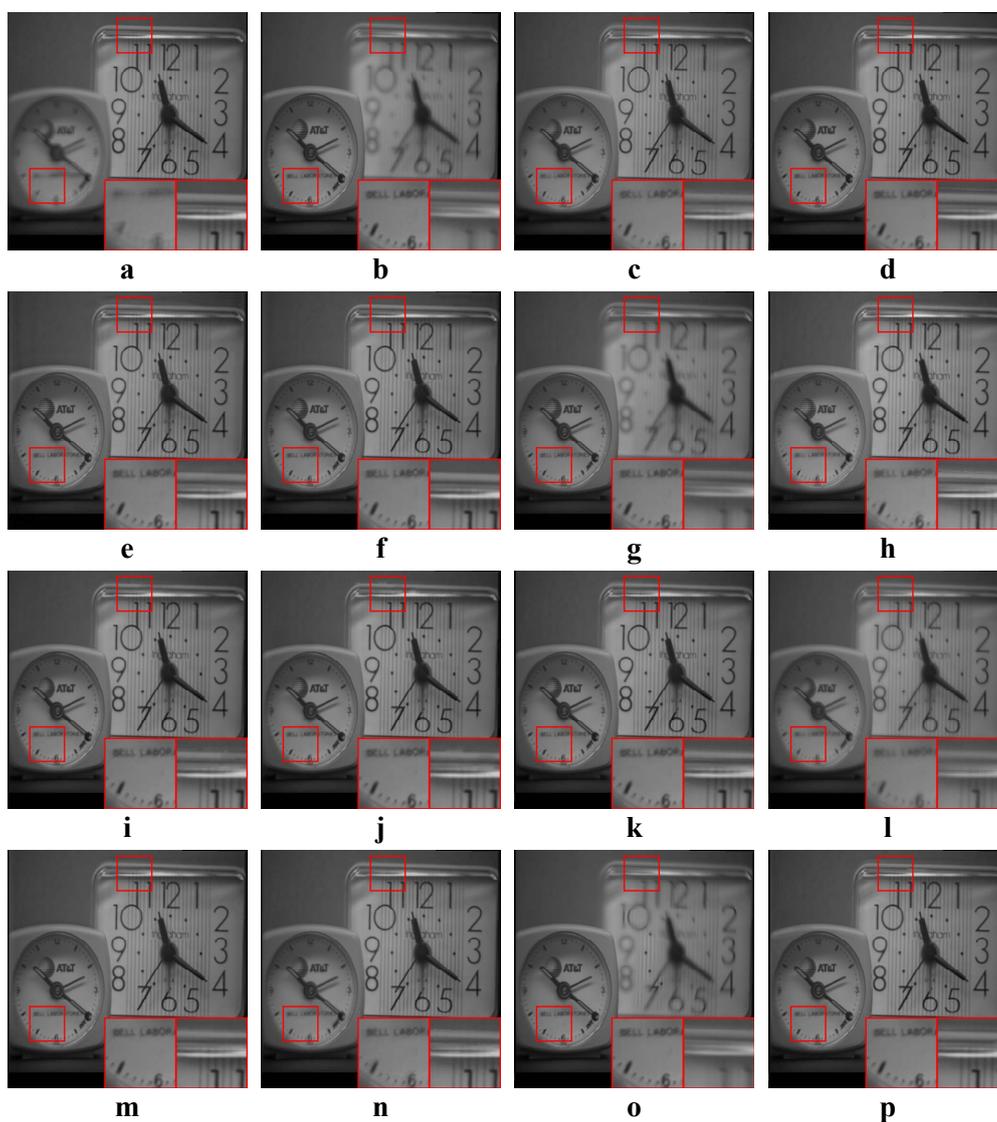


Figure 12. Examples of the fusion results of multi-focus images. (a,b) Source images; (c–p) The fused results of ASR, CSR, CVT, DTCWT, GTF, H-MSD, CNN, LP, MSSR, MSVD, NSCT, WLS, FFIF, and our proposed method, respectively.

Table 7. Objective evaluation of multi-focus image fusion. The best and second best results of each metric are marked in red and bold, respectively.

Metric	ASR	CSR	CVT	DTCWT	GTF	H-MSD	CNN
<i>MI</i>	7.5714	7.6874	7.2197	7.4468	7.6793	7.5452	8.5647
<i>FMI</i>	0.6022	0.4586	0.5643	0.5942	0.5963	0.5534	0.6065
$Q^{AB/F}$	0.7746	0.7570	0.7571	0.7710	0.6210	0.7477	0.7835
<i>NCIE</i>	0.8367	0.8363	0.8343	0.8360	0.8397	0.8363	0.8441
Metric	LP	MSSR	MSVD	NSCT	WLS	FFIF	OURS
<i>MI</i>	7.9235	7.8695	6.5958	7.5796	7.3741	9.2663	8.9213
<i>FMI</i>	0.6057	0.5987	0.4236	0.5963	0.5680	0.6558	0.6324
$Q^{AB/F}$	0.7829	0.7807	0.6212	0.7795	0.7647	0.7403	0.7891
<i>NCIE</i>	0.8390	0.8385	0.8299	0.8367	0.8349	0.8533	0.8467

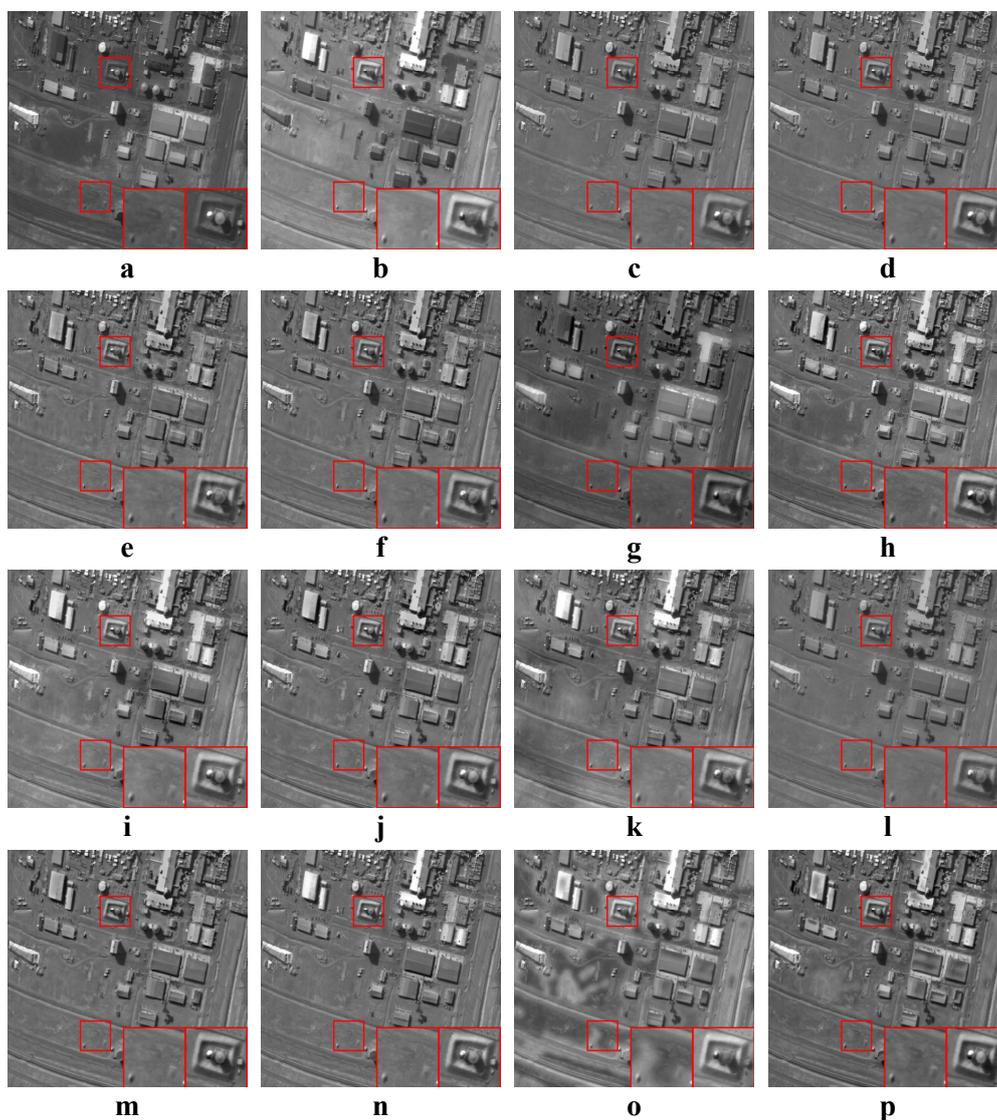


Figure 13. Examples of the fusion results of remote sensing images. (a,b) Source images; (c–p) The fused results of ASR, CSR, CVT, DTCWT, GTF, H-MSD, CNN, LP, MSSR, MSVD, NSCT, WLS, FFIF, and our proposed method, respectively.

Table 8. Objective evaluation of remote sensing image fusion. The best and second best results of each metric are marked in red and bold, respectively.

Metric	ASR	CSR	CVT	DTCWT	GTF	H-MSD	CNN
<i>MI</i>	2.0875	2.2285	1.8935	1.9541	1.7620	2.1677	3.6505
<i>FMI</i>	0.5057	0.4519	0.4305	0.4590	0.4967	0.4141	0.4741
$Q^{AB/F}$	0.5375	0.5908	0.5526	0.5802	0.4661	0.5428	0.6194
<i>NCIE</i>	0.8047	0.8052	0.8044	0.8045	0.8037	0.8055	0.8143
Metric	LP	MSSR	MSVD	NSCT	WLS	FFIF	OURS
<i>MI</i>	2.1795	3.1134	2.0510	2.0314	2.1937	4.5241	4.0035
<i>FMI</i>	0.4784	0.4646	0.3010	0.4757	0.4199	0.5473	0.5370
$Q^{AB/F}$	0.6173	0.5941	0.4055	0.6119	0.5389	0.6298	0.6567
<i>NCIE</i>	0.8051	0.8110	0.8044	0.8046	0.8050	0.8209	0.8165

4.4.1. Analysis of Infrared–Visible Results

According to Figure 10, our method produces clear structures of the windows and the barrel. GTF produces the sharpest silhouettes and structures. However, in the GTF result, the light under the windows in the visible image is not fused, which reduces the local contrast and makes the letters unclear. CNN performs similarly to our method. The results of other methods either have unclear structures or have low contrast.

According to Table 5, our method is the best on $Q^{AB/F}$ and the second best on the other three metrics. FFIF is the best on *MI*, *FMI* and *NCIE*, while it is the second best on $Q^{AB/F}$. However, the visual quality of FFIF is not so good. Some information in the visible image is not fused at all. Some details from the visible image are missing in the fused image. Although the metrics show that the results of FFIF have high information entropy, this is in exchange for the decline in visual quality. In contrast, our method achieves a balance between visual quality and information entropy.

4.4.2. Analysis of Medical Results

According to Figure 11, our method produces both high contrast and clear structures. With the exception of CNN, MSSR, FFIF and our method, all other methods make the texture from the middle part of (b) unclear due to low contrast. However, the structures produced by CNN, MSSR and FFIF are not as clear as our method.

According to Table 6, our method is the best on $Q^{AB/F}$ and the second best on *MI* and *NCIE*. FFIF is the best on *MI*, *FMI* and *NCIE*, while it is the second best on $Q^{AB/F}$. ASR is the second best on *FMI*. However, the edges and structure of the FFIF result are unclear. The structures of the two source images are mixed together, and it looks confusing. The result of ASR has low contrast and the middle part of the fused image is not clear. As pointed out by $Q^{AB/F}$, our method has the best visual quality. At the same time, our method also has high information entropy.

4.4.3. Analysis of Multi-Focus Results

According to Figure 12, our method preserves the clear letters and edges well. Letters in ASR, H-MSD, MSVD and WLS results are not very clear. The edges of the right clock in MSSR and NSCT are not well preserved. In the CSR result, there is an obvious artifact above the clock edge. The fusion results of GTF and FFIF are very poor, the clock on the right is very fuzzy, and the information in the source image (a) is hardly fused. CVT, DTCWT, CNN, LP perform similarly to our method.

According to Table 7, our method is the best on $Q^{AB/F}$ and the second best on the other three metrics. FFIF is the best on *MI*, *FMI* and *NCIE*. CNN is the second best on $Q^{AB/F}$. However, as mentioned earlier, the fusion quality of FFIF for some multi-focus images is very poor. In this case, high values of the metrics are not so convincing. Our method can achieve high values of metrics while ensuring visual quality.

4.4.4. Analysis of Remote Sensing Results

According to Figure 13, our method preserves the textures on the ground and the structures of the buildings well. ASR and MSVD fail to preserve the straight line structures in the right box. Except our method, all other methods do not preserve the textures on the ground well.

According to Table 8, our method is the best on $Q^{AB/F}$ and the second best on the other three metrics. FFIF is the best on MI , FMI and $NCIE$, while it is the second best on $Q^{AB/F}$. However, the results of FFIF suffer from spatial inconsistencies, and there are many discontinuous black patches on the ground. In contrast, our method has the best visual quality while having high information entropy.

4.4.5. Summary of the Analysis

According to Figures 10–13 and the $Q^{AB/F}$ metric, our method is the best in terms of visual quality compared with the other 13 methods. The MI , FMI and $NCIE$ metrics of our method are slightly lower than those of FFIF. Although the metrics show that the fused images of FFIF have higher information entropy, FFIF does not make full use of the information in each source image. This makes the visual quality of the FFIF method not so good. In contrast, our method achieves a good balance between visual quality and information entropy.

Overall, this experiment demonstrates that our method is comparable to or better than state-of-the-art methods both in visual and objective evaluations.

5. Conclusions

In this paper, a JSR and RGF based image fusion method is proposed. Our method uses JSR for image decomposition, reduces the correlation and highlights the complementary information between source images. This improves the low information entropy caused by the redundant information between source images. Multi-scale decomposition using RGF can remove small structures while preserving obvious edges in the innovation images. It can extract complementary details of the innovation images at different levels. Using weight maps to balance the contribution of the innovation images can suppress the artifacts that may be brought by JSR. The innovation images are used to guide the optimization of the weight maps so that the fused image can have high contrast. The fusion of innovation images is performed according to optimized weight maps to ensure the spatial consistency of the fused innovation image. Finally, adding the common image directly to the fused innovation image without processing ensures that the common information contained in the two source images can be well retained. Experimental results demonstrate that our main contributions have been achieved, and our method can achieve better performance than state-of-the-art methods.

Author Contributions: Conceptualization, Y.L. and X.Y.; data curation, Y.L.; formal analysis, T.C. and G.J.; funding acquisition, X.Y. and G.J.; investigation, Y.L. and R.Z.; methodology, M.K.A. and T.C.; project administration, X.Y.; resources, X.Y. and M.K.A.; software, Y.L.; supervision, X.Y. and R.Z.; validation, Y.L. and R.Z.; visualization, Y.L., M.K.A. and T.C.; writing—original draft, Y.L.; writing—review and editing, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The research in this paper was partially sponsored by National Natural Science Foundation of China (61701327, 61711540303), the Science Foundation of Sichuan Science and Technology Department (2018GZ0178), the Open research fund of State Key Laboratory (614250304010517), The Applied Basic Research Programs of Science and Technology Department of Sichuan Province (2019YJ0110) and The Science and Technology Service Industry Demonstration Programs of Sichuan Province (2019GFW167).

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

References

1. Goshtasby, A.A.; Nikolov, S. Image fusion: Advances in the state of the art. *Inf. Fusion* **2007**, *8*, 114–118.
2. James, A.P.; Dasarthy, B.V. Medical image fusion: A survey of the state of the art. *Inf. Fusion* **2014**, *19*, 4–19. [[CrossRef](#)]

3. Ma, J.; Yong, M.; Chang, L. Infrared and visible image fusion methods and applications: A survey. *Inf. Fusion* **2019**, *45*, 153–178.
4. Li, S.; Kang, X.; Fang, L.; Hu, J.; Yin, H. Pixel-level image fusion: A survey of the state of the art. *Inf. Fusion* **2017**, *33*, 100–112. [[CrossRef](#)]
5. Sahu, A.; Bhateja, V.; Krishn, A.; Himanshi. Medical image fusion with Laplacian Pyramids. In Proceedings of the 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom), Greater Noida, India, 7–8 November 2014; pp. 448–453.
6. Petrović, V.S.; Xydeas, C.S. Gradient-Based Multiresolution Image Fusion. *IEEE Trans. Image Process.* **2004**, *13*, 228–237. [[CrossRef](#)]
7. Li, H.; Manjunath, B.; Mitra, S. Multisensor Image Fusion Using the Wavelet Transform. *Graph. Model. Image Process.* **1995**, *57*, 235–245. [[CrossRef](#)]
8. Lewis, J.J.; O’Callaghan, R.J.; Nikolov, S.G.; Bull, D.R.; Canagarajah, N. Pixel- and region-based image fusion with complex wavelets. *Inf. Fusion* **2007**, *8*, 119–130. [[CrossRef](#)]
9. Zhang, Q.; Guo, B.L. Multifocus image fusion using the nonsubsampling contourlet transform. *Signal Process.* **2009**, *89*, 1334–1346. [[CrossRef](#)]
10. Lei, W.; Li, B.; Tian, L.F. Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients. *Inf. Fusion* **2014**, *19*, 20–28.
11. Liu, S.; Wang, J.; Lu, Y.; Li, H.; Zhao, J.; Zhu, Z. Multi-Focus Image Fusion Based on Adaptive Dual-Channel Spiking Cortical Model in Non-Subsampled Shearlet Domain. *IEEE Access* **2019**, *7*, 56367–56388. [[CrossRef](#)]
12. Liu, S.; Shi, M.; Zhu, Z.; Zhao, J. Image fusion based on complex-shearlet domain with guided filtering. *Multidimens. Syst. Signal Process.* **2017**, *28*, 207–224. [[CrossRef](#)]
13. He, K.; Sun, J.; Tang, X. Guided Image Filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1397–1409. [[CrossRef](#)] [[PubMed](#)]
14. Petschnigg, G.; Szeliski, R.; Agrawala, M.; Cohen, M.; Hoppe, H.; Toyama, K. Digital photography with flash and no-flash image pairs. *ACM Trans. Graph.* **2004**, *23*, 664–672. [[CrossRef](#)]
15. Qi, Z.; Shen, X.; Li, X.; Jia, J. Rolling Guidance Filter. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014.
16. Shutao, L.; Xudong, K.; Jianwen, H. Image fusion with guided filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875. [[CrossRef](#)] [[PubMed](#)]
17. Chen, L.; Yang, X.; Lu, L.; Liu, K.; Jeon, G.; Wu, W. An image fusion algorithm of infrared and visible imaging sensors for cyber-physical systems. *J. Intell. Fuzzy Syst.* **2019**, *36*, 4277–4291. [[CrossRef](#)]
18. Jian, L.; Yang, X.; Zhou, Z.; Zhou, K.; Liu, K. Multi-scale image fusion through rolling guidance filter. *Future Gener. Comput. Syst.* **2018**, *83*, 310–325. [[CrossRef](#)]
19. Olshausen, B.A.; Field, D.J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **1996**, *381*, 607–609. [[CrossRef](#)]
20. Yue, D.; Dong, L.; Xie, X.; Lam, K.M.; Dai, Q. Partially occluded face completion and recognition. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 4145–4148.
21. Tu, B.; Zhang, X.; Kang, X.; Zhang, G.; Wang, J.; Wu, J. Hyperspectral Image Classification via Fusing Correlation Coefficient and Joint Sparse Representation. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 340–344. [[CrossRef](#)]
22. Liu, S.; Liu, M.; Li, P.; Zhao, J.; Zhu, Z.; Wang, X. SAR Image Denoising via Sparse Representation in Shearlet Domain Based on Continuous Cycle Spinning. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2985–2992. [[CrossRef](#)]
23. Qin, Z.; Fan, J.; Liu, Y.; Gao, Y.; Li, G.Y. Sparse Representation for Wireless Communications: A Compressive Sensing Approach. *IEEE Signal Process. Mag.* **2018**, *35*, 40–58. [[CrossRef](#)]
24. Fang, L.; Zhuo, H.; Li, S. Super-resolution of hyperspectral image via superpixel-based sparse representation. *Neurocomputing* **2018**, *273*, 171–177. [[CrossRef](#)]
25. Yang, B.; Li, S. Multifocus Image Fusion and Restoration with Sparse Representation. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 884–892. [[CrossRef](#)]
26. Yu, N.; Qiu, T.; Feng, B.; Wang, A. Image Features Extraction and Fusion Based on Joint Sparse Representation. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 1074–1082. [[CrossRef](#)]

27. Ma, X.; Hu, S.; Liu, S.; Fang, J.; Xu, S. Multi-focus image fusion based on joint sparse representation and optimum theory. *Signal Process. Image Commun.* **2019**, *78*, 125–134. [[CrossRef](#)]
28. 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](#)]
29. Qu, G.; Zhang, D.; Yan, P. Information measure for performance of image fusion. *Electron. Lett.* **2002**, *38*, 313–315. [[CrossRef](#)]
30. Haghghat, M.B.A.; Aghagolzadeh, A.; Seyedarabi, H. A non-reference image fusion metric based on mutual information of image features. *Comput. Electr. Eng.* **2011**, *37*, 744–756. [[CrossRef](#)]
31. Xydeas, C.S.; Petrovic, V. Objective image fusion performance measure. *Electron. Lett.* **2000**, *36*, 308–309. [[CrossRef](#)]
32. Qiang, W.; Yi, S. Performances evaluation of image fusion techniques based on nonlinear correlation measurement. In Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Como, Italy, 18–20 May 2004; Volume 1, pp. 472–475.
33. Qu, X.B.; Yan, J.W.; Xiao, H.Z.; Zhu, Z.Q. Image Fusion Algorithm Based on Spatial Frequency-Motivated Pulse Coupled Neural Networks in Nonsampled Contourlet Transform Domain. *Acta Autom. Sin.* **2008**, *34*, 1508–1514. [[CrossRef](#)]
34. Liu, Y.; Liu, S.; Wang, Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf. Fusion* **2015**, *24*, 147–164. [[CrossRef](#)]
35. Liu, Y.; Wang, Z. Simultaneous image fusion and denoising with adaptive sparse representation. *IET Image Process.* **2015**, *9*, 347–357. [[CrossRef](#)]
36. Liu, Y.; Chen, X.; Ward, R.K.; Jane Wang, Z. Image Fusion With Convolutional Sparse Representation. *IEEE Signal Process. Lett.* **2016**, *23*, 1882–1886. [[CrossRef](#)]
37. Nencini, F.; Garzelli, A.; Baronti, S.; Alparone, L. Remote sensing image fusion using the curvelet transform. *Inf. Fusion* **2007**, *8*, 143–156. [[CrossRef](#)]
38. 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](#)]
39. Zhou, Z.; Wang, B.; Li, S.; Dong, M. Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters. *Inf. Fusion* **2016**, *30*, 15–26. [[CrossRef](#)]
40. Liu, Y.; Chen, X.; Cheng, J.; Peng, H.; Wang, Z. Infrared and visible image fusion with convolutional neural networks. *Int. J. Wavelets Multiresolution Inf. Process.* **2018**, *16*, 1850018. [[CrossRef](#)]
41. Naidu, V.P.S. Image Fusion Technique using Multi-resolution Singular Value Decomposition. *Def. Sci. J.* **2011**, *61*, 479–484. [[CrossRef](#)]
42. 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](#)]
43. Zhan, K.; Xie, Y.; Wang, H.; Min, Y. Fast filtering image fusion. *J. Electron. Imaging* **2017**, *26*, 1–18. [[CrossRef](#)]

