



## Article

# CBFM: Contrast Balance Infrared and Visible Image Fusion Based on Contrast-Preserving Guided Filter

Xilai Li , Xiaosong Li \* and Wuyang Liu

Guangdong-Hong Kong-Macao Joint Laboratory for Intelligent Micro-Nano Optoelectronic Technology, School of Physics and Optoelectronic Engineering, Foshan University, Foshan 528225, China; 20210300236@stu.fosu.edu.cn (X.L.); 2112203023@stu.fosu.edu.cn (W.L.)

\* Correspondence: lixiaosong@buaa.edu.cn

**Abstract:** Infrared (IR) and visible image fusion is an important data fusion and image processing technique that can accurately and comprehensively integrate the thermal radiation and texture details of source images. However, existing methods neglect the high-contrast fusion problem, leading to suboptimal fusion performance when thermal radiation target information in IR images is replaced by high-contrast information in visible images. To address this limitation, we propose a contrast-balanced framework for IR and visible image fusion. Specifically, a novel contrast balance strategy is proposed to process visible images and reduce energy while allowing for detailed compensation of overexposed areas. Moreover, a contrast-preserving guided filter is proposed to decompose the image into energy-detail layers to reduce high contrast and filter information. To effectively extract the active information in the detail layer and the brightness information in the energy layer, we proposed a new weighted energy-of-Laplacian operator and a Gaussian distribution of the image entropy scheme to fuse the detail and energy layers, respectively. The fused result was obtained by adding the detail and energy layers. Extensive experimental results demonstrate that the proposed method can effectively reduce the high contrast and highlighted target information in an image while simultaneously preserving details. In addition, the proposed method exhibited superior performance compared to the state-of-the-art methods in both qualitative and quantitative assessments.

**Keywords:** infrared and visible image fusion; contrast balance; guided curvature filter; Gaussian distribution of entropy; contrast-preserving guided filter



**Citation:** Li, X.; Li, X.; Liu, W. CBFM: Contrast Balance Infrared and Visible Image Fusion Based on Contrast-Preserving Guided Filter. *Remote Sens.* **2023**, *15*, 2969. <https://doi.org/10.3390/rs15122969>

Academic Editors: Riccardo Roncella and Mattia Previtali

Received: 24 April 2023

Revised: 2 June 2023

Accepted: 5 June 2023

Published: 7 June 2023

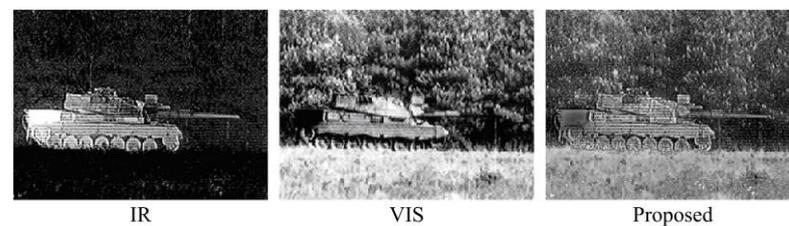


**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Numerous sensors are used for data collection in real-life scenarios. Unlike a single sensor, the use of multiple sensors to collect data from the same scene facilitates a comprehensive and accurate interpretation [1–3]. However, redundant data may be present when multiple sensors are used. Therefore, specific algorithms should be developed to extract critical pixel information. In infrared and visible image fusion (IVIF) [4–6], the target information of the thermal radiation is extracted from the infrared (IR) image, and the details and contrast information are obtained from the visible image. In IR imaging sensors, the physical properties of IR radiation are used to measure objects. Therefore, these sensors are not easily affected by complex external environments (i.e., haze and high-exposure areas). However, although IR images can approximately locate the area of a target based on the temperature or radiation difference, they cannot capture the detail and color information of the target scene, which results in low contrast. Therefore, using only a single IR image cannot provide comprehensive information about the scene. Visible imaging sensors capture the reflection of the target scene and can effectively preserve contrast and detail information in the scene. However, these sensors are easily affected by external environmental factors. Integrating useful detail information from a visible image and target information from an IR image into a single image can effectively compensate for

the innate defects of various sensors and facilitate human observation without the influence of complex environments. Therefore, IVIF technology has been developed rapidly and is applied in military reconnaissance [7] and target tracking [8], among other fields. Figure 1 shows the application of IVIF in the military domain, demonstrating its effectiveness in enhancing military reconnaissance tasks. The fusion results demonstrate the ability of the technology to detect and identify military facilities, such as tanks and armored vehicles, within a scene.



**Figure 1.** Examples of infrared (IR) and visible image fusion technology used in military applications.

IVIF methods can be categorized into deep learning (DL)-based and conventional methods. DL-based methods have been widely used in image fusion and achieve satisfactory results. For example, Xu et al. [9] proposed an unsupervised network for image fusion tasks that uses the same model and parameters to process various fusion problems. Zhang et al. [10] proposed a CNN-based image fusion framework that processes images in an end-to-end manner without postprocessing. To address the dependence of DL on the training dataset, Liu et al. [11] proposed a deep architecture for IVIF and achieved excellent fusion performance. However, DL-based algorithms rely on computer hardware and require intensive training. Furthermore, inadequate real-world datasets have limited the development of DL-based methods.

Multi-scale transform (MST)-based methods are commonly used in conventional IVIF methods, in which the source image is typically decomposed into multiple sub-bands and appropriate fusion rules are designed based on the features represented by the sub-bands [1]. For example, Chen et al. [12] proposed a target-enhanced MST decomposition model to extract details and energy information at various scales. However, multilevel and multidirectional decomposition may result in significant computational complexity of the algorithm. For IVIF tasks, Li et al. [13] proposed a total variation decomposition strategy to decompose an image into structure and texture layers, and construct appropriate weights for fusion in various layers. Nie et al. [14] proposed a total variation-based IVIF strategy to obtain fusion results using a variation model without designing fusion rules for various decomposition components such as in multi-scale decomposition methods. Furthermore, various filters are used in IVIF tasks. For example, to achieve the fine classification of IR and visible images, Mo et al. [15] proposed an attribute filter for IVIF that fully considers the attributes of objects in the source image and can effectively extract the target information in the IR image. In addition to the use of filters as a decomposition method, LatLRR [16] is widely used for IVIF. In this decomposition strategy, the source image is a combination of a low-rank component that represents the global structure and a prominent component that represents the local structure. Liu et al. [17] proposed a multi-decomposition LatLRR-based IVIF method that combined the advantages of MST and LatLRR-based methods to extract details from the source image. However, the inefficiency of this decomposition process limits its practical application.

In IVIF, target information extracted from the IR image and texture information obtained from the visible image are integrated into a single image. Typically, pixels at the same location in different source images do not affect each other, and the measurement values of the feature area pixels are higher than those in flat areas. However, because of the diversity and complexity of the shooting environments, many images can be affected to varying degrees. For example, when capturing outdoor photographs during the day, images may be overexposed and exhibit high contrast because of intense sunlight. When capturing

photographs outdoors at night, light from passing cars and streetlamps may cause unreasonable contrast and scattering. Under hazy weather, visible images may enhance gray color information, which may result in the loss of scene details and colors. These problems increase the difficulty of developing IVIF methods. However, in most IVIF algorithms, only the ideal situation for IVIF is considered, and the aforementioned problems are ignored. In complex scenes, the high contrast and haze grayscale information provided by visible image may cause the scene to lose useful details and textures. If these adverse visual phenomena are not suppressed, thermal radiation information or high-contrast features in the fusion results may be lost, resulting in unclear targets. Figure 2 displays the fusion results for the proposed algorithm and the comparison method, namely, relative total variation decomposition (RTVD) [13], for the two groups of IR and visible source images. When the source image exhibits overexposure and light-scattering effects, the RTVD method loses target pixel information and produces suboptimal fusion results. Although many image processing techniques, such as enhancement [18], dehazing [19,20], and denoising [21], have been proposed, if the image is preprocessed separately before fusion, this increases the complexity of the fusion task and is not conducive to practical applications.



**Figure 2.** Results for the proposed algorithm and the comparison method, relative total variation decomposition (RTVD), on two sets of IR and visible source images.

Therefore, we propose a contrast balance approach based on the IVIF method, CBFM, to adjust the overly high-contrast areas of visible source images through a contrast balance strategy. Thus, the pixels in these areas can be prevented from affecting the distribution of target information in the IR image. First, we propose a multilayer decomposition scheme based on guided curvature filtering to obtain the most distinctive energy layer in the visible image. The high-brightness pixel information in the energy layer causes an overly high contrast in the image; therefore, we reduce the weight of the energy layer when reconstructing the source image. To avoid the detail loss caused by pixel loss in the energy layer, we perform detail compensation in the detail layer; i.e., we increase the weight of the useful detail information in the detail layer. Using the contrast balance strategy, we obtain a preprocessed visible image. We then propose a novel weighted guided filter that constructs weights using a combination of contrast fidelity and sparse constraints to effectively blur overexposed areas and extract important detail information. Using this weighted guided filter, various source images can be decomposed into energy and detail layers. For energy layer fusion, we propose a weighted fusion rule, based on entropy and a Gaussian distribution, which adjusts the contrast in the energy layer via weighted averaging. For detail layer fusion, we propose a weighted energy-of-Laplacian (EOL) operator that can effectively detect pixels with high activity and extract clear structural and detail information. The source code of this work is publicly available at <https://github.com/ixilai/CBFM>.

The contributions of this study are as follows:

1. A novel IVIF algorithm based on contrast balance is proposed to effectively address the fusion challenge in complex environments, including maintaining reasonable contrast and detail fusion tasks affected by adverse phenomena such as overexposure, haze, and light diffusion.

2. A novel contrast balance strategy is proposed to reduce the adverse effects of overexposure in the visible light image by decreasing the weight of the energy layers in the source image and supplementing the details.
3. A contrast-preserving guided filter (CPGF) that constructs weights specifically for IVIF tasks is proposed. The IVIF outperforms the guided filter (GIF) and weighted GIF (WGF).

The remainder of this paper is structured as follows: Section 2 provides a detailed presentation of the proposed algorithm, Section 3 describes the experimental setup, and Section 4 presents the main conclusions of the study.

## 2. Materials and Methods

In this section, the specific steps of the proposed method are described in detail (Figure 3). First, to prevent high-contrast pixel information in visible images from affecting the distribution of thermal radiation target information, we proposed a contrast balance strategy for processing visible images. Additionally, we designed a new CPGF filter specifically for the IVIF task to effectively extract feature information of different dimensions from IR and visible images. This filter can decompose the source image into an energy layer and a detail layer. The energy layer contains most of the energy of the image and low-frequency information, which determines the overall structure and texture of the image. To extract pixels with the most information from the energy layer, we used a new weighted average fusion rule based on entropy Gaussian distribution. The detail layer represents image details or local features with high pixel activity. The weighted EOL operator can effectively extract significant edge details, making it suitable for the design of detail layer fusion rules. Finally, we added the fusion energy layer to the detail layer to obtain the final fusion result.

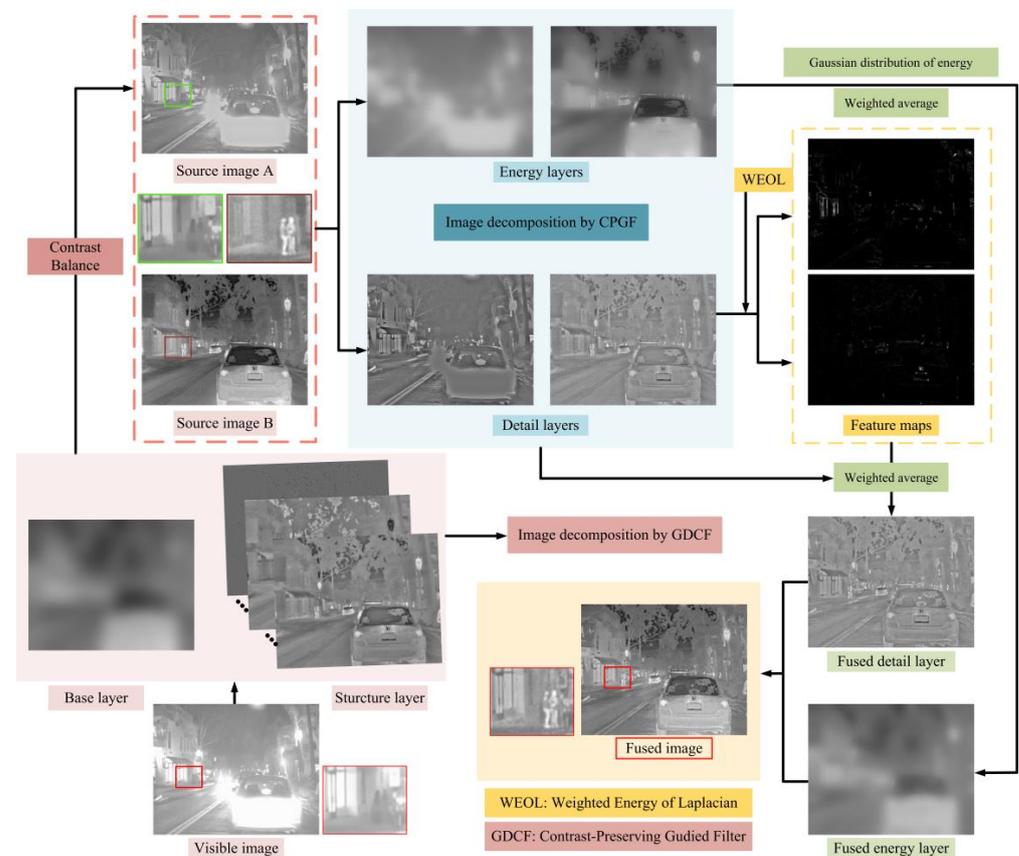


Figure 3. Framework of the proposed method.

### 2.1. Contrast Balance Strategy

Highly exposed visible images typically result in high contrast and loss of target thermal radiation information. Therefore, we proposed a strategy to balance the contrast in overexposed regions of visible images. We considered the visible image to be a combination of energy and detail layers. The energy layer generally carries considerable luminance and energy information in the source image. Therefore, the weights in the energy layer were reduced to adjust the high-contrast regions in the source image. However, detailed texture information may be lost when the weight of the energy layer is reduced. Therefore, we performed a detail enhancement process; i.e., the weight of the detail layer was increased in the source image so that the obtained preprocessed visible source image can effectively balance the contrast and enhance the distribution of useful details in the image.

To effectively extract detail information at various scales using the energy in the source image subject to [22], we proposed a novel guided curvature filtering (GDCF)-based image decomposition strategy. GDCF is based on GIF [23] and curvature filtering [24] and can decompose the source image into energy, fine structure (FS), and coarse structure (CS) layers.

#### 2.1.1. GDCF-based Multilayer Decomposition Strategy

GIF is used as a filter for image smoothing to effectively smooth out detail features of the image and preserve strong edges. Suppose the guide image is  $G$ , the input image is  $f$  and the output image is  $O$ . The local linear model can be expressed as follows:

$$O(p) = \alpha_{p'}G(p) + \beta_{p'}, \forall p \in \Omega_{\psi_1}(p') \quad (1)$$

where  $\Omega_{\psi_1}(p')$  represents a square window with radius  $\psi_1$  centered on pixel  $p'$ , and the values of  $\alpha_{p'}$  and  $\beta_{p'}$  are fixed in this window;  $\psi_1$  is set to 15.

The parameters of  $\alpha$ ,  $\beta$  are determined by minimizing the cost function, as follows:

$$E(\alpha_{p'}, \beta_{p'}) = \sum_{p \in \Omega_{\psi_1}(p')} \left( (\alpha_{p'}G(p) + \beta_{p'} - f(p))^2 + \hat{\lambda}\alpha_{p'}^2 \right) \quad (2)$$

where  $\hat{\lambda}$  is a regularization parameter that penalizes a large  $\alpha_{p'}$ , which was set to 0.3 in this study. Subsequently, the GIF process is represented as  $GIF(\cdot)$ .

Regarding the GDCF-based multilayer decomposition strategy, given that  $f_g^{(i)}$  ( $i = 1, 2, \dots, n$ ) is the filtered result of the  $i$ th GIF;  $f_c^{(i)}$  ( $i = 1, 2, \dots, n$ ) is the filtered result of the  $i$ th CF; and  $n$  is a constant that represents the number of decomposition layers (and is set to 3 in this study). Furthermore,  $f$  can be expressed as follows:

$$f = \sum_{i=1}^n (FS^{(i)} + CS^{(i)}) + EG \quad (3)$$

where  $EG$ ,  $FS^{(i)}$ , and  $CS^{(i)}$  represent the base layer,  $i$ th fine structure, and  $i$ th coarse structure, respectively, which can be expressed as follows:

$$EG = f_g^{(n)} \quad (4)$$

$$FS^{(i)} = \begin{cases} f - f_c^{(1)}, & i = 1 \\ f_g^{(i-1)} - f_c^{(i)}, & i = 2, 3, \dots, n \end{cases} \quad (5)$$

$$CS^{(i)} = f_c^{(i)} - f_g^{(i)}, \quad (i = 1, 2, \dots, n) \quad (6)$$

where  $f_g^{(i)}$  and  $f_c^{(i)}$  are computed as follows:

$$f_g^{(i)} = \begin{cases} GIF(f, f, \hat{\lambda}), i = 1 \\ GIF(f_g^{(i-1)}, f_g^{(i-1)}, \hat{\lambda}), i = 2, 3, \dots, n \end{cases} \quad (7)$$

$$f_c^{(i)} = \begin{cases} CF(f, v), i = 1 \\ CF(f_c^{(i-1)}, v), i = 2, 3, \dots, n \end{cases} \quad (8)$$

where  $CF(\cdot)$  represents the CF operation [22] and  $v$  represents the iteration number, which was set to five in this study.

### 2.1.2. Contrast Balance

After obtaining the fine structure, rough structure, and base layers, we adjusted the contrast in the source image using the following weighted fusion rules to enhance the detail information in the figure:

$$P = 0.8 \times EG + K \times FS^{(1)} + (K - 1) \times FS^{(2)} + (K - 1.5) \times FS^{(3)} + K \times CS^{(1)} + (K - 0.5)CS^{(2)} + (K - 1) \times CS^{(3)} \quad (9)$$

where  $K$  was set to 1.8. In Equation (9), the weights assigned to the first 3 layers of  $FS$  are  $K$ ,  $K - 1$ , and  $K - 1.5$ , whereas the weights for the first 3 layers of  $CS$  are  $K$ ,  $K - 0.5$ , and  $K - 1$ . This weighting scheme is based on the characteristics of  $FS$  and  $CS$  in representing different types of information in the source image.  $FS$  captures fine structure details, whereas  $CS$  represents large-scale edge structure information. The weight values gradually decrease as we move from the first layer to subsequent layers, because the amount of detailed information tends to diminish. This weighting scheme ensures a balanced contribution from each layer. In regions with excessive contrast, almost all the detailed information is lost, and only a small amount of large-scale edge information exists. Hence, a larger weight must be assigned to the  $CS$  overall during detailed compensation. In addition, because the  $FS$  layer contains more detailed information, if its weight is considerably large, it may lead to information redundancy in the preprocessed image, which can affect the visual result. In addition, because undesirable visual phenomena, such as overexposure, generally occur only in visible images, the contrast-balancing strategy applies only to the visible images. In the subsequent fusion process, we denoted the IR source image as  $f_1$  and preprocessed the visible image as  $f_2$ .

## 2.2. Image Decomposition by CPGF

### 2.2.1. Proposed CPGF

A previous study [25] revealed that because the parameter  $\lambda$  is fixed in conventional GIF, blurring is concentrated near the edges and halos are introduced. To solve this problem, Li et al. [25] added edge-aware weights to GIF to form the WGF, as follows:

$$E(\alpha_{p'}, \beta_{p'}) = \sum_{p \in \Omega_{\psi_1}(p')} \left( (\alpha_{p'} G(p) + \beta_{p'} - f(p))^2 + \frac{\lambda}{\Gamma_G(p')} \alpha_{p'}^2 \right) \quad (10)$$

where  $\Gamma_G(p')$  denotes the edge perception weight and  $\lambda$  is a regularization parameter.

Compared with GIF, WGF can avoid halo artifacts associated with the filtering process and produce a higher-quality filtered image. However, for IR and visible image fusion tasks, image filtering is not performed for edge-preserving smoothing, but to efficiently extract target and detail information in the scene without the interference of high-contrast areas and to smooth overly bright pixel information. Therefore, edge-aware weights in [25] are not necessary for IVIF; thus, the assignment of appropriate weighting factors for specific fusion tasks is critical. Subject to [26], we introduced a weight that can improve the excessive

contrast in the source image by constructing new weights using the specification of contrast fidelity and sparsity constraints, as follows:

Contrast fidelity  $\zeta_1$  is defined by the following expression:

$$\zeta_1 = \|\omega - f\|_2^2 \tag{11}$$

where  $\omega$  denotes the weight map, and Equation (11) balances the contrast in the input image by using the L2 norm to maintain a reasonable light-to-dark ratio of the target to background information, without requiring the exact pixel intensity as the input image.

Sparse constraint  $\zeta_2$  is defined as follows:

$$\zeta_2 = \|\nabla\omega - \nabla f\|_1 \tag{12}$$

where  $\nabla\omega$  and  $\nabla f$  represent the gradients of  $\omega$  and  $f$ , respectively.

The weighting problem can be formulated by minimizing the following objective function according to the contrast fidelity and sparsity norms:

$$\xi(\omega) = \zeta_1 + \gamma\zeta_2 = \|\omega - f\|_2^2 + \gamma\|\nabla\omega - \nabla f\|_1 \tag{13}$$

where  $\gamma$  is the balance parameter control between  $\zeta_1$  and  $\zeta_2$ , which was set to 15 in this study. In addition,  $\|\cdot\|_2^2$  refers to the squared L2 norm and  $\|\cdot\|_1$  represents the L1 norm, which calculates the sum of the absolute values of the elements of a vector or matrix. According to [26], the above Equation (13) can be solved to obtain the contrast weights  $\omega$ .  $\omega$  is embedded in Equation (10) to form a new CPGF, as follows:

$$E(\alpha_{p'}, \beta_{p'}) = \sum_{p \in \Omega_{\psi_1}(p')} \left( (\alpha_{p'}G(p) + \beta_{p'} - f(p))^2 + \frac{\lambda}{\omega(p')} \alpha_{p'}^2 \right) \tag{14}$$

where  $\alpha_{p'}$  and  $\beta_{p'}$  are obtained as:

$$\alpha_{p'} = \frac{\mu_{G \odot f, \psi_1}(p') - \mu_{G, \psi_1}(p')\mu_{f, \psi_1}(p')}{\sigma_{G, \psi_1}^2(p') + \frac{\lambda}{\omega(p')}} \tag{15}$$

$$b_{p'} = \mu_{f, \psi_1}(p') - \alpha_{p'}\mu_{G, \psi_1}(p') \tag{16}$$

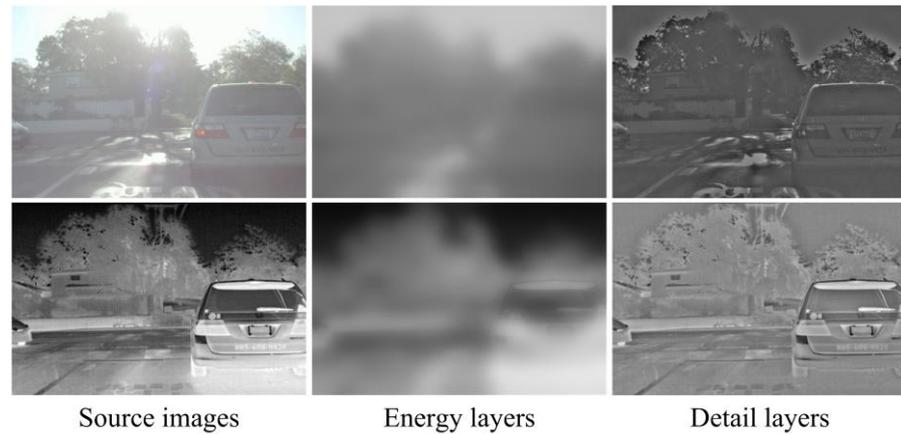
where  $\odot$  represents the element-by-element product of two matrices;  $\mu_{G \odot f, \psi_1}(p')$ ,  $\mu_{G, \psi_1}(p')$ , and  $\mu_{f, \psi_1}(p')$  are the mean values of  $G \odot f$ ,  $G$ , and  $f$ , respectively. Furthermore,  $\sigma_{G, \psi_1}^2(p')$  is the variance in a local window with a radius of  $3 \times 3$  centered on pixel  $p'$ .

### 2.2.2. Image Decomposition

First, we filtered the IR and visible source images using the proposed WGF and obtained the energy layer  $E_t$  of the source image  $f_t$ , where  $t$  denotes the number of source images, which was set to two in this study. Next, we obtained the detail layer  $D_t$  using the following equation:

$$D_t = f_t - E_t \tag{17}$$

Figure 4 displays an example of image decomposition using CPGF. The effect of excessive contrast on detail extraction can be effectively avoided using the proposed model, and the performance of the algorithm is improved.



**Figure 4.** Example of the proposed contrast-preserving guided filter-based image decomposition.

### 2.3. Energy Layer Fusion

Although excessive contrast information in the energy layer can be effectively reduced using the proposed WGF, the energy layer still contains the most luminance and target information in the source image. Entropy represents the aggregation of the luminance distribution in the source image, and the extreme probability density function of entropy can be used to determine the degree of image blurring [27]. In the energy layer of IR and visible source images, clear pixels typically contain more energy information, which indicates that they have a higher entropy value compared to blurred pixels. Therefore, in this study, we used the Gaussian distribution of image entropy [28] to construct fusion weights to effectively retain the most representative energy information in various source images. The energy layer  $E_t$  and the entropy of  $e_t$  can be expressed as follows:

$$e_t = - \sum_j P_j(f_t) \log_2 P_j(f_t) \quad (18)$$

where  $P_j(f_t)$  is the probability of intensity value  $j$  in the image  $f_t$ . Therefore, the probability can be calculated as follows:

$$P(e_t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(e_t - \varepsilon)^2}{2\sigma^2}\right) \quad (19)$$

As in [28],  $\varepsilon$  and  $\sigma$  were set to 7.4600 and 0.8732, respectively. According to  $P(e_t)$ , we can construct the fusion weights to obtain the fusion results of the energy layer as follows:

$$FE = \frac{P(e_1)}{P(e_1) + P(e_2)} \times E_1 + \frac{P(e_2)}{P(e_1) + P(e_2)} \times E_2 \quad (20)$$

where  $E_1$  and  $E_2$  represent the energy layers of the source images  $f_1$  and  $f_2$ , respectively.

### 2.4. Detail Layer Fusion

The detail layer represents rich texture information in the source image, and the ideal fusion result should contain clear details in both the visible and IR images. The energy of Laplacian of the image (EOL) [29] can be used by the Laplace operator to analyze the high spatial frequencies associated with the sharpness of the image boundaries, and the model can effectively detect the information of pixels with greater sharpness in the image, i.e., the detail corresponding to the source image. Therefore, we proposed a novel weighted EOL (WEOL) method to construct a weight map of the detail layer. The input image is  $D_t$ , and EOL is expressed as follows:

$$EOL(f_t) = \sum_x \sum_y (D_{xx} + D_{yy})^2 \quad (21)$$

according to [29],  $D_{xx} + D_{yy}$  can be defined as:

$$D_{xx} + D_{yy} = -D(x-1, y-1) - 4D(x-1, y) - D(x-1, y+1) - 4D(x, y-1) + 20D(x, y) - 4D(x, y+1) - D(x+1, y-1) - 4D(x+1, y) - D(x+1, y+1) \quad (22)$$

In the proposed algorithm, the WEOL is denoted as follows:

$$WEOL(D_t) = \sum_x \sum_y W' \cdot (D_{xx} + D_{yy})^2 \quad (23)$$

where the weight matrix  $W'$  is defined as follows:

$$W' = \frac{1}{15} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix} \quad (24)$$

The detail layer fusion result is obtained based on  $WEOL(f_t)$  to construct the detail layer fusion weights to obtain the detail layer fusion results.

$$FD = \frac{WEOL(D_1)}{WEOL(D_1) + WEOL(D_2)} \times D_1 + \frac{WEOL(D_2)}{WEOL(D_1) + WEOL(D_2)} \times D_2 \quad (25)$$

where  $D_1$  and  $D_2$  represent the detail layers of source images  $f_1$  and  $f_2$ , respectively.

### 2.5. Fusion Result Construction

Finally, the fused energy layer and detail layer are added to develop fusion result  $F$  as follows:

$$F = FE + FD \quad (26)$$

## 3. Experiments

### 3.1. Experimental Setup

We compared the proposed algorithm with nine state-of-the-art algorithms and conducted extensive experiments on two publicly available IR and visible image datasets. The first dataset was a road dataset [9] that contained 221 sets of IR and visible images captured under complex road conditions. The primary challenge was to effectively suppress the impacts of high contrast from strong light sources during the day and night. The second dataset was from [30] and contained 48 sets of IR and visible images captured in outdoor scenes. The proposed algorithm is abbreviated as CBFM in our experiments. LatLRR [16], RTVD [13], TEMF [12], MFEIF [11], U2Fusion [9], DIVFusion [31], GAN-McC [32], SDDGAN [33], and UMFusion [34] are the nine state-of-the-art algorithms that were compared with our proposed algorithm. The source code with default parameters provided by the original authors was used. In addition, CBFM and three traditional methods were tested using Matlab 2021b on a PC with an AMD Ryzen 5 4600H with a Radeon Graphics processor and NVIDIA GeForce GTX 1650 graphics card. The results of the fusion of six deep learning-based comparison methods were obtained using Python 3.8 and a discrete PC with an Intel Core i7-7700HQ processor and NVIDIA GeForce GTX-1070 graphics card.

Six objective evaluation metrics, namely, gradient-based fusion performance ( $Q_G$ ) [35], image fusion metric based on a multi-scale scheme ( $Q_M$ ) [36], Piella's metric ( $Q_S$ ) [36], entropy (EN) [37], average gradient (AG) [38], and spatial frequency (SF) [39], were used to evaluate the fusion performance of all the methods. Specifically,  $Q_G$  evaluates the amount of edge information;  $Q_M$  is a fusion image quality metric based on a multi-scale scheme; AG reflects the gradient information in the fusion result (i.e., the richer the gradient information, the higher the image quality); EN is based on information theory and measures the amount of information contained in the image; and SF reflects the sharpness of the

fused image, wherein a higher value represents a higher-quality fused image and better algorithm performance. The six metrics can be used to comprehensively evaluate the fusion performance of various methods, with higher values representing better performance.

### 3.2. Parameter Analysis

For the regularization parameter in Equation (10),  $\lambda$  considerably affects the performance of the proposed algorithm. Therefore, we performed parametric analysis to determine the optimal parameter values to ensure excellent fusion performance. First, we fixed the values of other parameters of the model and set  $\lambda \in \{0.05, 0.07, \dots, 0.4\}$ . We then randomly selected 30 sets of images from the road dataset to add to the parameter analysis experiment. The scores for each index when different values were obtained are presented in Table 1; the maximum values are in bold. When  $\lambda$  increases,  $Q_M$ , EN, AG, and SF gradually increase, indicating that the parameter  $\lambda$  can affect the ability of the model to extract features. However, the trend in the values of the  $Q_G$  and  $Q_S$  indicators does not conform to the aforementioned rule. When the value of  $\lambda$  was set to 0.4, both the  $Q_G$  and  $Q_S$  metrics achieved the lowest scores among all the comparison experiments. This outcome suggests that an excessively large  $\lambda$  value can result in redundant detail information and structural distortion in the fusion results. Such distortions are unfavorable for producing visually useful images that are consistent with the human visual system. Therefore, to achieve visual balance and optimal fusion performance, we set the value of  $\lambda$  to 0.3. This value met the requirements for the proposed algorithm, resulting in excellent scores for all the metrics.

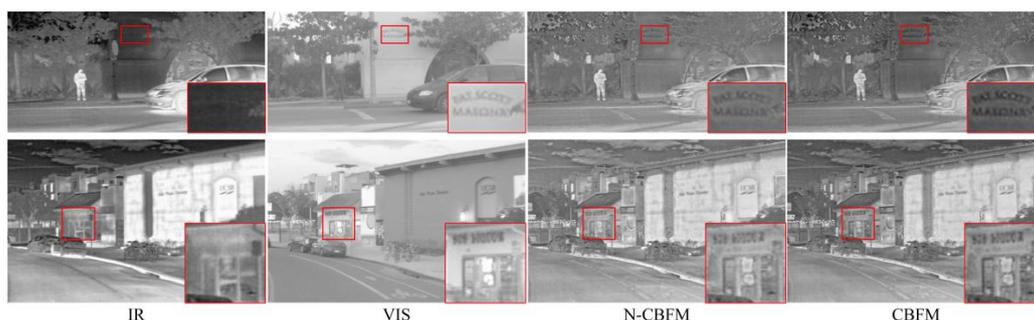
**Table 1.** Quantitatively comparing the effect of different values of  $\lambda$ .

$\lambda$	$Q_G$	$Q_M$	$Q_S$	EN	AG	SF
0.05	0.5238	0.6027	<b>0.7914</b>	7.0539	4.9628	13.1062
0.07	<b>0.5240</b>	0.6123	0.7901	7.0659	5.0875	13.4770
0.1	<b>0.5240</b>	0.6203	0.7886	7.0770	5.1983	13.8106
0.2	0.5234	0.6296	0.7862	7.0930	5.3534	14.2851
0.3	0.5229	0.6338	0.7851	7.0994	5.4133	14.4711
0.4	0.5227	<b>0.6345</b>	0.7846	<b>7.1028</b>	<b>5.4450</b>	<b>14.5696</b>

### 3.3. Ablation Analysis

#### 3.3.1. Ablation Analysis of Contrast Balance Strategy

To effectively reduce the effects of high contrast and diffusion from overexposed regions in visible images, a novel contrast-balancing strategy was proposed. An ablation analysis was performed to determine whether this strategy could improve the fusion performance of the proposed model. The algorithm without contrast-balancing treatment was labeled N-CBFM, as shown in Figure 5; the three sets of fusion results for CBFM and N-CBFM are also provided. The contrast-balancing strategy effectively improved the fusion performance of the model to better retain reasonable contrast and present the detail information in source images.



**Figure 5.** Example of qualitative comparison of contrast-balanced strategy ablation experiments.

### 3.3.2. Ablation Analysis of the Proposed Filter

For the specificity of the IVIF task, we designed a novel weighted guided filter whose weights consisted of contrast fidelity with sparsity constraints. To evaluate the effectiveness of the strategy, we performed an ablation analysis. First, we selected 20 sets of source images from the road dataset as the data source. Second, we compared GIF and WGF to the proposed method as the filters for image decomposition in the proposed algorithm. For comparison, we named the method of image decomposition by GIF as A-CBFM and the method of image decomposition by WGF as B-CBFM. The same parameters were used for all three filters to ensure an unbiased comparison. The experimental results for the methods are shown in Table 2. The maximum values are in bold. Table 2 shows that CBFM obtained the highest scores for the four metrics, and although the  $Q_G$  and  $Q_S$  metrics did not score as well as B-CBFM, they were generally better than the compared algorithms. Thus, the decomposition of images using the proposed CPGF yields better fusion performance than GIF and WGF methods.

**Table 2.** Quantitative comparison of different comparison algorithms.

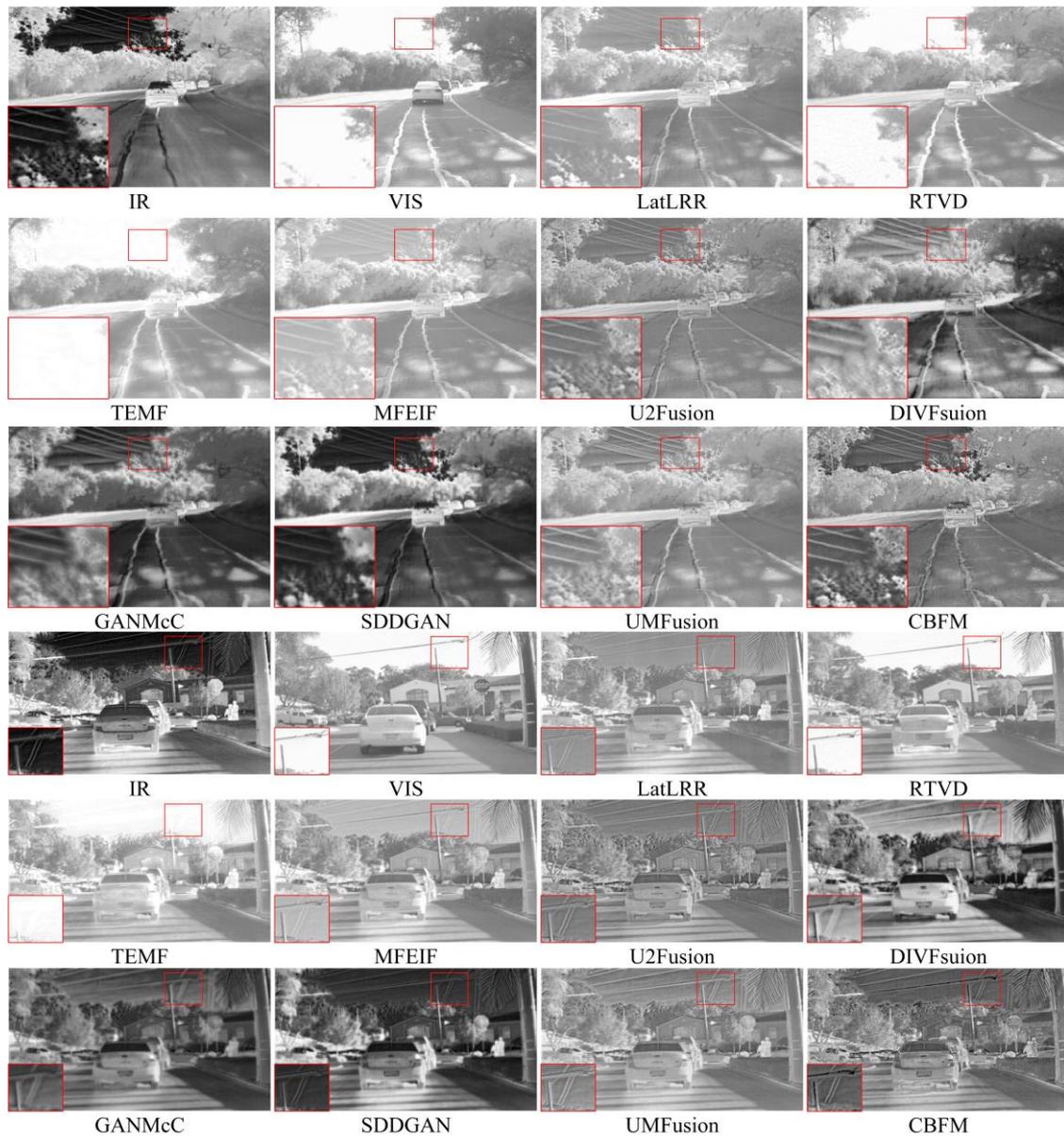
Methods	$Q_G$	$Q_M$	$Q_S$	EN	AG	SF
A-CBFM	0.5333	0.5789	0.7680	7.1411	5.9602	15.6960
B-CBFM	<b>0.5375</b>	0.4986	<b>0.7811</b>	6.9772	4.7292	11.8979
CBFM	0.5327	<b>0.5915</b>	0.7665	<b>7.1500</b>	<b>6.0415</b>	<b>15.9530</b>

### 3.4. Subjective Evaluation

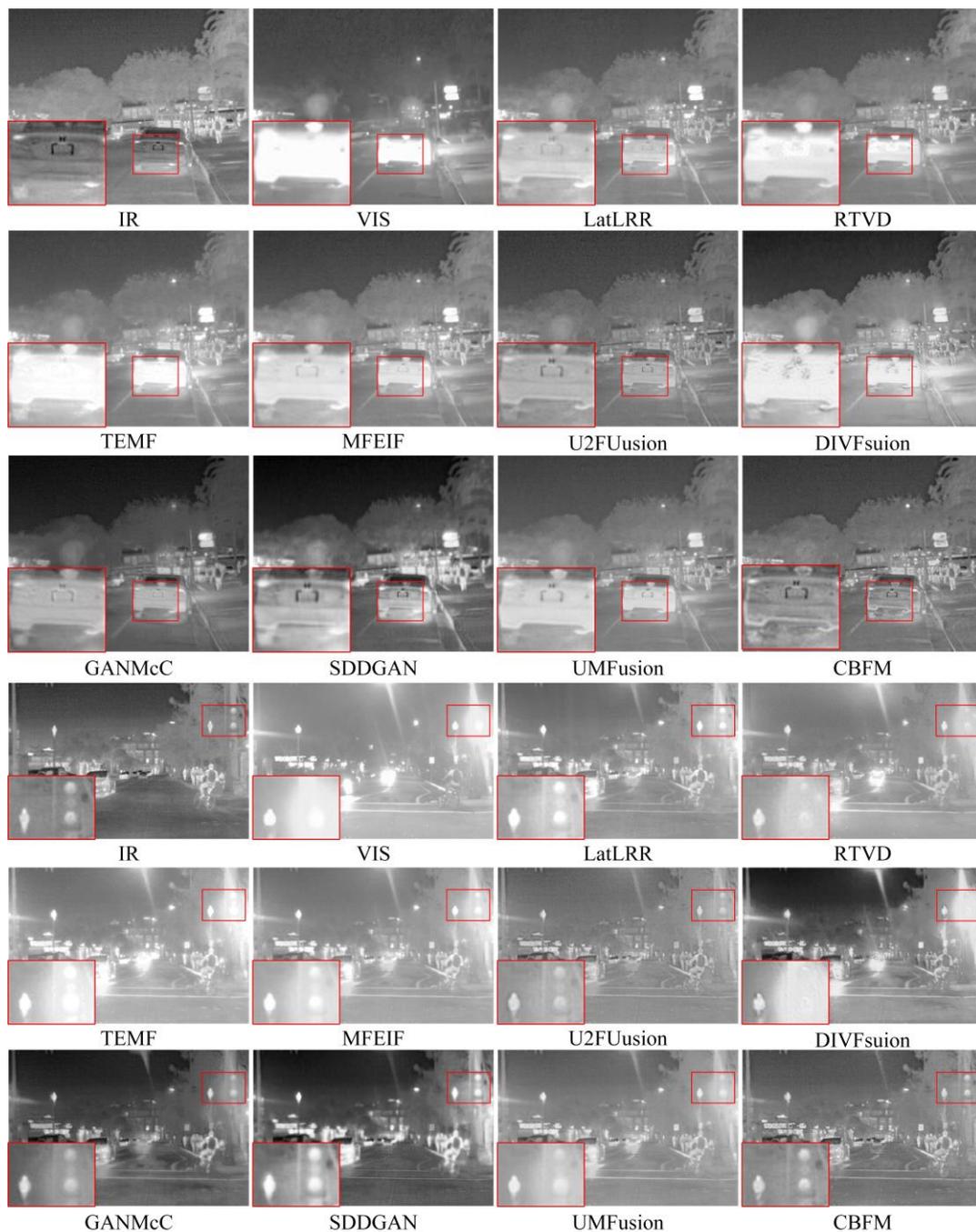
Figure 6 displays the fusion results for the proposed algorithm with nine state-of-the-art comparison methods for two classical image pairs from the road dataset. For an intuitive comparison, we zoomed in on a local region of the fusion results located in the high-contrast region of the visible image. The IR image provides rich target information and provides a clear view of the distribution of cars on the road, whereas the visible image contains texture information. RTVD and TEMF are easily affected by the high-contrast region in the visible image and cannot identify the target information in the IR image. In contrast, the four methods LatLRR, MFEIF, U2Fusion, and UMFusion can retain some target information such as power lines in the sky and leaves on trees. However, they cannot maintain reasonable contrast to some extent, which can reduce the visual effect of the fusion results, and do not provide a comprehensive interpretation of the scene. Three methods, DIVFusion, GANMcC, and SDDGAN, appear to produce blurred images and lose the target and detail information in the scene. The CBFM exhibits superior ability to extract target information from the IR image and the texture information from the visible image, and can maintain a reasonable contrast in the fusion result. This result could be attributed to the filtering of most of the high-contrast pixel information in the proposed WGF to ensure that the luminance target information in the IR image can be effectively extracted. If contrast balance is not applied to the highlighted region, this phenomenon may lead to a situation similar to that of the RTVD and TEMF methods, in which the overly bright pixel information in the visible image can mask the target luminance information in the IR image. The proposed algorithm can achieve better fusion performance than the five state-of-the-art contrast methods.

Figure 7 displays the fusion results of the proposed algorithm for nine state-of-the-art comparison methods for two classical image pairs from the road dataset. In this study, we selected images captured at night, and both headlights and streetlights on the road had light-diffusion effects, which could produce undesirable visual effects in the fusion results if they are not suppressed. All five algorithms—RTVD, TEMF, MFEIF, DIVFusion, and UMFusion—were affected by the light-diffusion effect, which resulted in the loss of some of the target information in the IR images. Although LatLRR and U2Fusion effectively mitigated this situation, they still lost part of the image contrast and exhibited insufficient feature extraction capability. Three methods, namely, DIVFusion, GANMcC,

and SDDGAN, appeared to blur the target information and generate suboptimal fusion results. The proposed algorithm can maintain excellent fusion performance despite the light diffusion effect at night, effectively balancing the contrast between various source images and providing excellent visual effects for fusion results. The proposed algorithm handled the images of various situations and exhibited a superior fusion performance than the five state-of-the-art algorithms.



**Figure 6.** Qualitative comparison of CBFM compared with nine state-of-the-art algorithms on road dataset.



**Figure 7.** Qualitative comparison of CBFM with nine state-of-the-art methods with headlights and streetlights on the road of the road dataset.

Figure 8 shows the fusion results of the two sets of IR and visible images captured from outdoor scenes. As displayed in the figure, in the first fusion task, the main challenge arises from the sky region of the visible image, which has high contrast and energy information because of the intense light source of the sun, which may lead to the loss of target information in the IR image if it not suppressed. Moreover, the visible image is rich in detail information; therefore, the preservation of details should be considered when balancing the contrast. In the second fusion task, the primary challenge originates from the foggy sky in the visible image, which may lead to the loss of texture information and inconspicuousness of the target information in the fusion result. Therefore, a comparison of these two fusion tasks can effectively reflect the performance of the various fusion algorithms. Both the

TEMF and MFEIF methods were disturbed by high-contrast information, which led to inconspicuous target information of clouds in the sky. The DIVFusion method contained incorrect target information, resulting in poor visual quality fusion results. Under the influence of a foggy sky, all four methods, LatLRR, TEMF, SDDGAN, and UMFusion, failed to retain the texture information of tree leaves in visible images and produced suboptimal fusion results. As observed in the red magnified area, all five methods lost some useful pixel information during the fusion process, which resulted in a lower contrast with a loss of texture. Unlike the comparison methods, the proposed algorithm, CBFM, exhibited excellent performance in both sets of fusion tasks, proving that the contrast-balancing strategy can effectively solve the fusion problem in complex situations without losing detail information while achieving reasonable contrast.



**Figure 8.** Qualitative comparison of CBFM with nine state-of-the-art methods on IVIF dataset.

### 3.5. Objective Evaluation

In addition to subjective visual comparisons, objective evaluations based on the entire road dataset were performed in the comparison experiments. Table 3 presents the scores of the proposed algorithm and the comparison methods on the road dataset for various metrics. The maximum values are highlighted in bold, the second-highest scores are shown in red, and the third-highest scores are shown in blue. Table 3 shows that the proposed algorithm obtained the best scores for all five metrics, proving that it performed well in extracting features and retaining information in the source images. The U2Fusion method obtained the second highest score for the two metrics, similar to the subjective evaluation, and achieved the best visual results among all the compared methods. However, U2Fusion obtained the lowest score in the EN metric because it lost most of the contrast information during fusion.

**Table 3.** Quantitative comparison of the proposed algorithm and different comparison methods on the road scene dataset. Bold is the best, red is second, and blue is third.

Methods	$Q_G$	$Q_M$	$Q_S$	EN	AG	SF
LatLRR	0.3990	0.4393	0.7786	6.9161	3.7787	10.1186
RTVD	0.4555	<b>0.5538</b>	0.7415	7.0207	4.1072	10.9419
TEMF	0.3657	0.4118	0.7335	6.9794	3.6752	9.8647
MFEIF	0.4312	0.4558	<b>0.7832</b>	7.0488	3.7644	9.5561
U2Fusion	<b>0.4884</b>	0.4514	<b>0.8133</b>	6.8021	<b>4.6377</b>	<b>11.4237</b>
DIVFusion	0.2883	0.3309	0.6213	<b>7.5318</b>	<b>4.8050</b>	<b>11.6477</b>
GANMcC	0.3606	0.4034	0.7126	<b>7.2366</b>	3.7788	9.0192
SDDGAN	0.3655	0.3946	0.7377	<b>7.5261</b>	4.3825	10.4413
UMFusion	<b>0.4785</b>	<b>0.4771</b>	<b>0.8133</b>	7.0474	4.0954	10.5501
CBFM	<b>0.5489</b>	<b>0.7217</b>	<b>0.8186</b>	7.1383	<b>5.8674</b>	<b>15.2249</b>

Table 4 shows the performance of the methods on the dataset [9]. The maximum values are highlighted in bold, the second-highest scores are shown in red, and the third-highest scores are shown in blue. The results of the quantitative comparison show that the proposed algorithm achieved the best scores in three metrics:  $Q_G$ , AG, and SF. The proposed algorithm achieved fusion with high-resolution and rich feature information. Given that the purpose of this algorithm was to maintain a reasonable contrast of the image, some of the pixel energy information may be missing. Thus, the data does not facilitate comparison with some of the methods for the EN metrics. RTVD and U2Fusion achieved excellent performance on this dataset. However, a qualitative evaluation revealed that RTVD exhibited a partial loss of texture information during fusion. Therefore, its score on the SF index was lower compared to the scores for U2Fusion and CBFM. Thus, consistent with the conclusions obtained for the subjective evaluation, the proposed algorithm outperformed the other nine methods in fusion. This result proved the superiority of the contrast-balancing strategy.

### 3.6. Computational Time

Table 5 presents the computational time comparison results for all 10 methods. Notably, the TEMF method exhibits the lowest computational cost among the compared methods. However, this can be attributed to its decomposition of fewer layers, potentially leading to a compromised performance in handling overexposed regions. In contrast, the LatLRR method, which relies on learning the projection matrix, demonstrates relatively higher time consumption. Although the proposed algorithm may not achieve optimal computational efficiency owing to the longer runtime of CPGF, we acknowledge this limitation and plan to explore the impact of different parameters on the computational load in future research. A direct comparison of the six DL-based methods may seem unfair owing to the various platforms used to run the proposed algorithm. Nonetheless, DL-based algorithms facilitate higher computational efficiency, which is advantageous.

**Table 4.** Quantitative comparison of the proposed algorithm and different comparison methods on IVIF dataset [30]. Bold is the best, red is second, and blue is third.

Methods	Q <sub>G</sub>	Q <sub>M</sub>	Q <sub>S</sub>	EN	AG	SF
LatLRR	0.4176	0.3169	0.8251	7.1010	4.8726	14.0734
RTVD	<b>0.4914</b>	<b>0.5246</b>	0.7986	7.1284	5.1918	14.3747
TEMF	0.4268	<b>0.3709</b>	0.8085	6.8659	4.9588	14.0858
MFEIF	0.4642	0.3622	<b>0.8330</b>	7.1174	4.8914	12.8592
U2Fusion	<b>0.4830</b>	0.3106	<b>0.8442</b>	7.0989	<b>6.3451</b>	<b>16.0479</b>
DIVFusion	0.3029	0.2252	0.7218	<b>7.5289</b>	4.9692	12.0900
GANMcC	0.3847	0.2971	0.7791	7.1071	4.7454	12.6711
SDDGAN	0.3573	0.2654	0.7675	<b>7.4608</b>	4.3714	11.1505
UMFusion	0.4582	0.3566	<b>0.8438</b>	<b>7.1387</b>	<b>5.1984</b>	<b>14.4103</b>
CBFM	<b>0.5246</b>	<b>0.4097</b>	0.8277	7.1164	<b>7.0740</b>	<b>20.2707</b>

**Table 5.** Comparison of the average computational time of 10 methods on the road scene dataset.

Methods	Time	Methods	Time
LatLRR	29.1462	DIVFusion	2.64
RTVD	0.5451	GANMcC	1.103
TEMF	0.01	SDDGAN	0.166
MFEIF	0.093	UMFusion	0.7692
U2Fusion	0.861	CBFM	8.5802

#### 4. Conclusions

In this study, a novel IVIF algorithm based on a contrast-balancing strategy was proposed to effectively mitigate the impact of high-contrast regions, such as highly exposed areas, and address the issue of light diffusion effects present in the source image. Furthermore, we proposed the CPGF to categorize the source image into detail and energy layers. To fuse the detail layers, we proposed a weighted EOL operator to obtain the highest fusion weights. To fuse the energy layers, we designed a weighted fusion rule based on the Gaussian distribution of entropy. The final fusion results were obtained by summing the fused energy and detail layers. Experimental results revealed that the fusion performance of the proposed algorithm was superior compared to other state-of-the-art comparison methods and effectively suppressed undesirable visual effects in the source image. In addition, excellent fusion results that provide a comprehensive interpretation of the scene were obtained. The proposed contrast balance strategy may not achieve significant advantages in terms of computational efficiency because a multi-scale decomposition approach is used to extract the sub-bands of various features. Therefore, in the future, we will explore methods to optimize the strategy to reduce computational complexity.

**Author Contributions:** Conceptualization and methodology, X.L. (Xilai Li) and X.L. (Xiaosong Li); software, X.L. (Xilai Li) and W.L.; writing—original draft preparation, X.L. (Xilai Li); writing—review and editing, X.L. (Xiaosong Li); supervision, X.L. (Xiaosong Li); visualization, X.L. (Xilai Li) and W.L.; funding acquisition, X.L. (Xiaosong Li). All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (No. 62201149) and the Ji Hua Laboratory (No. X200051UZ200) of Guangdong province, China.

**Data Availability Statement:** The source code of is publicly available at <https://github.com/ixilai/CBFM>.

**Acknowledgments:** The authors would like to thank the editors and anonymous reviewers for their constructive and valuable comments and suggestions.

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

## References

1. Li, X.; Zhou, F.; Tan, H. Joint image fusion and denoising via three-layer decomposition and sparse representation. *Knowl.-Based Syst.* **2021**, *224*, 107087. [[CrossRef](#)]
2. Li, X.; Wang, X.; Cheng, X.; Tan, H.; Li, X. Multi-Focus Image Fusion Based on Hessian Matrix Decomposition and Salient Difference Focus Detection. *Entropy* **2022**, *24*, 1527. [[CrossRef](#)] [[PubMed](#)]
3. Liu, X.; Gao, H.; Miao, Q.; Xi, Y.; Ai, Y.; Gao, D. MFST: Multi-Modal Feature Self-Adaptive Transformer for Infrared and Visible Image Fusion. *Remote Sens.* **2022**, *14*, 3233. [[CrossRef](#)]
4. Li, H.; Cen, Y.; Liu, Y.; Chen, X.; Yu, Z. Different Input Resolutions and Arbitrary Output Resolution: A Meta Learning-Based Deep Framework for Infrared and Visible Image Fusion. *IEEE Trans. Image Process.* **2021**, *30*, 4070–4083. [[CrossRef](#)]
5. Li, H.; Zhao, J.; Li, J.; Yu, Z.; Lu, G. Feature dynamic alignment and refinement for infrared–visible image fusion: Translation robust fusion. *Inf. Fusion* **2023**, *95*, 26–41. [[CrossRef](#)]
6. Qi, B.; Jin, L.; Li, G.; Zhang, Y.; Li, Q.; Bi, G.; Wang, W. Infrared and visible image fusion based on co-occurrence analysis shearlet transform. *Remote Sens.* **2022**, *14*, 283. [[CrossRef](#)]
7. Zhou, H.; Ma, J.; Yang, C.; Sun, S.; Liu, R.; Zhao, J. Nonrigid feature matching for remote sensing images via probabilistic inference with global and local regularizations. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 374–378. [[CrossRef](#)]
8. Li, H.; Wu, X.-J.; Durrani, T. NestFuse: An Infrared and Visible Image Fusion Architecture Based on Nest Connection and Spatial/Channel Attention Models. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9645–9656. [[CrossRef](#)]
9. Xu, H.; Ma, J.; Jiang, J.; Guo, X.; Ling, H. U2Fusion: A Unified Unsupervised Image Fusion Network. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 502–518. [[CrossRef](#)]
10. Zhang, Y.; Liu, Y.; Sun, P.; Yan, H.; Zhao, X.; Zhang, L. IFCNN: A general image fusion framework based on convolutional neural network. *Inf. Fusion* **2020**, *54*, 99–118. [[CrossRef](#)]
11. Liu, J.; Fan, X.; Jiang, J.; Liu, R.; Luo, Z. Learning a Deep Multi-Scale Feature Ensemble and an Edge-Attention Guidance for Image Fusion. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 105–119. [[CrossRef](#)]
12. Chen, J.; Li, X.; Luo, L.; Mei, X.; Ma, J. Infrared and visible image fusion based on target-enhanced multiscale transform decomposition. *Inf. Sci.* **2020**, *508*, 64–78. [[CrossRef](#)]
13. Chen, J.; Li, X.; Wu, K. Infrared and visible image fusion based on relative total variation decomposition. *Infrared Phys. Technol.* **2022**, *123*, 104112. [[CrossRef](#)]
14. Nie, R.; Ma, C.; Cao, J.; Ding, H.; Zhou, D. A Total Variation With Joint Norms For Infrared and Visible Image Fusion. *IEEE Trans. Multimed.* **2022**, *24*, 1460–1472. [[CrossRef](#)]
15. Mo, Y.; Kang, X.; Duan, P.; Sun, B.; Li, S. Attribute filter based infrared and visible image fusion. *Inf. Fusion* **2021**, *75*, 41–54. [[CrossRef](#)]
16. Li, H.; Wu, X.-J. Infrared and visible image fusion using latent low-rank representation. *arXiv* **2018**, arXiv:1804.08992.
17. Liu, X.; Wang, L. Infrared polarization and intensity image fusion method based on multi-decomposition LatLRR. *Infrared Phys. Technol.* **2022**, *123*, 104129. [[CrossRef](#)]
18. Chen, Q.; Zhang, Z.; Li, G. Underwater Image Enhancement Based on Color Balance and Multi-Scale Fusion. *IEEE Photonics J.* **2022**, *14*, 3963010. [[CrossRef](#)]
19. Liu, X.; Li, H.; Zhu, C. Joint Contrast Enhancement and Exposure Fusion for Real-World Image Dehazing. *IEEE Trans. Multimed.* **2022**, *24*, 3934–3946. [[CrossRef](#)]
20. Raikwar, S.C.; Tapaswi, S. Lower Bound on Transmission Using Non-Linear Bounding Function in Single Image Dehazing. *IEEE Trans. Image Process.* **2020**, *29*, 4832–4847. [[CrossRef](#)]
21. Li, H.; He, X.; Tao, D.; Tang, Y.; Wang, R. Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning. *Pattern Recognit.* **2018**, *79*, 130–146. [[CrossRef](#)]
22. Tan, W.; Zhou, H.; Song, J.; Li, H.; Yu, Y.; Du, J. Infrared and visible image perceptive fusion through multi-level Gaussian curvature filtering image decomposition. *Appl. Opt.* **2019**, *58*, 3064–3073. [[CrossRef](#)] [[PubMed](#)]
23. Li, S.; Kang, X.; Hu, J. Image Fusion with Guided Filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875. [[CrossRef](#)] [[PubMed](#)]
24. Gong, Y.; Sbalzarini, I.F. Curvature Filters Efficiently Reduce Certain Variational Energies. *IEEE Trans. Image Process.* **2017**, *26*, 1786–1798. [[CrossRef](#)] [[PubMed](#)]
25. Li, Z.; Zheng, J.; Zhu, Z.; Yao, W.; Wu, S. Weighted Guided Image Filtering. *IEEE Trans. Image Process.* **2015**, *24*, 120–129. [[CrossRef](#)] [[PubMed](#)]
26. Li, G.; Lin, Y.; Qu, X. An infrared and visible image fusion method based on multi-scale transformation and norm optimization. *Inf. Fusion* **2021**, *71*, 109–129. [[CrossRef](#)]
27. Fang, Y.; Ma, K.; Wang, Z.; Lin, W.; Fang, Z.; Zhai, G. No-Reference Quality Assessment of Contrast-Distorted Images Based on Natural Scene Statistics. *IEEE Signal Process. Lett.* **2015**, *22*, 838–842. [[CrossRef](#)]
28. Ou, F.-Z.; Wang, Y.-G.; Zhu, G. A novel blind image quality assessment method based on refined natural scene statistics. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 1004–1008.
29. Huang, W.; Jing, Z. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognit. Lett.* **2007**, *28*, 493–500. [[CrossRef](#)]
30. Fredembach, C.; Süsstrunk, S. Colouring the near infrared. In Proceedings of the IS&T/SID 16th Color Imaging Conference, Portland, OH, USA, 10–15 November 2008.

31. Tang, L.; Xiang, X.; Zhang, H.; Gong, M.; Ma, J. DIVFusion: Darkness-free infrared and visible image fusion. *Inf. Fusion* **2023**, *91*, 477–493. [[CrossRef](#)]
32. Ma, J.; Zhang, H.; Shao, Z.; Liang, P.; Xu, H. GANMcC: A Generative Adversarial Network With Multiclassification Constraints for Infrared and Visible Image Fusion. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 5005014. [[CrossRef](#)]
33. Zhou, H.; Wu, W.; Zhang, Y.; Ma, J.; Ling, H. Semantic-supervised Infrared and Visible Image Fusion via a Dual-discriminator Generative Adversarial Network. *IEEE Trans. Multimed.* **2021**, *25*, 635–648. [[CrossRef](#)]
34. Wang, D.; Liu, J.; Fan, X.; Liu, R. Unsupervised misaligned infrared and visible image fusion via cross-modality image generation and registration. *arXiv* **2022**, arXiv:2205.11876.
35. Xydeas, C.S.; Petrovic, V. Objective image fusion performance measure. *Electron. Lett.* **2000**, *36*, 308–309. [[CrossRef](#)]
36. Liu, Z.; Blasch, E.; Xue, Z.; Zhao, J.; Laganieri, R.; Wu, W. Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 94–109. [[CrossRef](#)]
37. Estevez, P.A.; Tesmer, M.; Perez, C.A.; Zurada, J.A. Normalized Mutual Information Feature Selection. *IEEE Trans. Neural Netw.* **2009**, *20*, 189–201. [[CrossRef](#)]
38. Cui, G.; Feng, H.; Xu, Z.; Li, Q.; Chen, Y. Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition. *Opt. Commun.* **2015**, *341*, 199–209. [[CrossRef](#)]
39. Zheng, Y.; Essock, E.A.; Hansen, B.C.; Haun, A.M. A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. *Inf. Fusion* **2007**, *8*, 177–192. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.