Infrared and Visible Image Fusion under Different Illumination Conditions based on Illumination Effective Region Map

Ying Tong¹,², Jin Chen¹, Member, IEEE

¹Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
²Department of Communications Engineering, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan

Corresponding author: Ying Tong (e-mail: tonyg2334@163.com)

This work has been supported by National Natural Science Foundation of China (Grant No: 61701344) and Tianjin Higher Education Creative Team Funds Program in China.

ABSTRACT It is well known that the quality of visible images taken under different illumination conditions varies greatly, so the illumination factors will affect infrared and visible image fusion effects inevitably. This paper proposes an infrared and visible image fusion algorithm that satisfies poor illumination conditions. The algorithm is divided into three stages: image preprocessing, multiscale decomposition and image fusion. The purpose of image preprocessing is to improve the contrast of visible image and extract the visual salient regions of infrared image. In the multiscale analysis stage, infrared and visible images are decomposed into different scales by combining Gaussian transform and Rolling Guidance Filter, which effectively avoids the halo artifacts. In the image fusion stage, the fusion weights of base layer coefficients are determined based on the combination of the saliency map of infrared image and the Illumination Effective Region Map of visible image. The detail layer coefficients of images are fused by choose-max fusion rule based on the local variance of detail features of the original images. Experiments show that the fusion effects of the proposed algorithm are robust to illumination variations, and the fused images have good detail clarity and can preserve the effective information of the original images well in an unsatisfactory illumination condition.

INDEX TERMS Image fusion, Infrared imaging, Image sensors, Image decomposition

I. INTRODUCTION

The imaging principles of infrared sensors and visible sensors are quite different. Infrared sensors capture the infrared radiation emitted by the object in the scene. If the hardware factors of infrared sensors are not considered, their imaging effect mainly depends on the infrared radiation emitted by the object itself, but has little relationship with the surrounding illumination. Visible sensors receive the reflected light of the object. The sensors need to receive moderate amount of light to get a clear image, so they have greater dependence on environmental factors, that means visible imaging sensors cannot work normally under poor illumination conditions. Therefore, the effect of illumination factors on visible image will also be embodied directly in the fused image if it is not processed properly. In order to solve the above problems, this paper proposes a fusion algorithm of infrared and visible images, which is robust to poor illumination conditions. In order to achieve the ideal fusion effect, we adjust the brightness of the visible image taken under the conditions of darkness or over-exposure and extract the effective information of the visible image to the most, meanwhile highlight the salient areas of the infrared image. The algorithm can meet the needs of image fusion in unknown or harsh environments, as well as all-weather or multi-illumination conditions.

The main feature of the undesired performance of visible sensors caused by environment illumination factors is that the brightness of the image captured is too low, too high or uneven, resulting in poor contrast and low visibility. Some researchers have proposed the method of illumination estimation to fuse a group of differently exposed visible images in a High Dynamic Range (HDR) scene [1]. The recovery of image illumination is achieved by screening normal exposed pixels in the image and giving them a large
fusion weight. In the field of object recognition, illumination factors will seriously affect the recognition accuracy. L. Lu, et al [2] proposed a novel illumination invariant multispectral palmprint recognition method. X. Tan, et al [3] presented an image preprocessing chain that eliminates most of the effects of illumination variations while still preserves the essential details that are required for recognition. W. Hao, et al [4] proposed a Retinex-like algorithm based on Bilateral Filter to improve the image quality of visible images under poor visual conditions. In this paper, a brightness adjustment scheme based on Retinex and Illumination Effective Region Map is proposed in the preprocessing stage, which provides a guarantee for retaining the effective information of visible image.

Multiscale analysis is widely used in infrared and visible image fusion. By decomposing the two images to be fused into different scales, specific fusion rules can be formulated for components of different layers. In multiscale image analysis, the difference between two adjacent filtered images can represent the detail information of the current scale. The commonly used image multiscale analysis methods include pyramid transform, wavelet transform, edge-preserving filters, etc. This kind of multiscale analysis method is simple in principle and high in execution efficiency, but it is easy to cause halo effects at the edge of the image. Therefore, some researchers try to combine Gaussian Filter with other edge-holding filters to achieve the desired effect [5]. In [6], Gaussian Filter is combined with Bilateral Filter to decompose the image, and the fusion rules are formulated for the image components of different scales, and a good fusion effect is obtained. Rolling Guidance Filter is a Bilateral Filter based iterative filter proposed by Qi Zhang [7] et al. It can ensure the accuracy of large area object boundaries while removing and smoothing complex small areas in images. In this paper, the Gaussian Filter is combined with Rolling Guidance Filter to extract the base layer and a series of detail layers. Then the layers of different scales are fused according to their specific fusion rules. Experiments show that the proposed algorithm enhances the contrast of visible image effectively. The detail features are preserved well in the fused image. At the same time, the saliency feature of the infrared image is highlighted in the fused image. The contrast of the fused image is clear and consistent with human visual perception.

II. ALGORITHM DESCRIPTION
The flow of the proposed algorithm can be divided into three parts, as shown in Fig. 1, they are image preprocessing, multiscale analysis and image fusion.

A. IMAGE PREPROCESSING
The purpose of this stage is to adjust the contrast of visible image by enhancement and denoising. The Illumination Effective Region Map of visible image is also formed in this step. As for infrared image, the visual saliency map is obtained during this process. It lays the foundation for setting the fusion weight of the base layer coefficients.

1) VISIBLE IMAGE PREPROCESSING
In this paper, Retinex algorithm [8] combined with Rolling Guidance Filter [7] is used to enhance and denoise visible images, so as to improve the quality of visible images under unfavorable illumination environments. Then, the brightness of image is estimated by transforming the processed image into YUV space, and the illumination effective areas are estimated by non-linear curve, in order to generate the Illumination Effective Region Map.

a. IMAGE CONTRAST ENHANCEMENT AND DENOISING
Retinex theory holds that the brightness of an object perceived by human eyes depends on the illumination of the

![FIGURE 1. Flow of the proposed algorithm.](image_url)
environment and the reflection of the incident light from the surface of the object [8].

\[ I(x, y) = I(x, y) \times R(x, y) \]  \hspace{1cm} (1)

In (1), \( I(x, y) \) represents the image signal received by human eyes or image sensors. It stands for the reflection property which is also the intrinsic property of the image. This part of information should be well preserved in the fused image. \( I(x, y) \) represents the illumination component of ambient light, which determines the dynamic range of image brightness. For images captured in unsatisfactory illumination environment, this part of information should be eliminated as clearly as possible. If the logarithm of the two sides of (1) is taken, the original appearance of the object \( R(x, y) \) can be obtained by subtracting the illumination component \( L(x, y) \) from \( I(x, y) \), as is shown in (2).

\[ \log[R(x, y)] = \log[I(x, y)] - \log[L(x, y)] \]  \hspace{1cm} (2)

Equation (2) shows that if we want to get \( R(x, y) \), only \( L(x, y) \) is required. Retinex approximates \( L(x, y) \) with the convolution of \( I(x, y) \) and a Gaussian kernel. So, \( R(x, y) \) can be expressed as follows:

\[ R(x, y) = \exp[\log(I(x, y)) - \log(I(x, y) \ast G(x, y))] \]  \hspace{1cm} (3)

where in (3), \( \ast \) represents convolution operation, \( G(x, y) \) represents a Gaussian kernel. After mapping \( R(x, y) \) to the range of \([0, 255] \), a contrast enhanced image can be obtained.

According to the characteristics of visible sensors, under-exposed or over-exposed images captured in an environment with insufficient or excessive illumination often contain much noise, which will not be eliminated after being processed by Retinex algorithm. The Retinex itself also produces halo artifacts when dealing with images that contain areas with high brightness. In order to reduce the influence of noise and halos, the proposed algorithm applies Rolling Guidance Filter on the enhanced images after Retinex enhancement.

Rolling Guidance Filter is an edge-preserving filter proposed by Qi Zhang et al. of the University of Hong Kong. It is an upgraded version of Bilateral Filter that preserves image edges while smoothing noise [7]. By combining the Retinex algorithm with Rolling Guidance Filter, the contrast-adjusted visible image can be denoised, which not only eliminates the original noise caused by the unsatisfactory exposure, but also smoothes the halo effects introduced by Retinex.
values on both sides will have less ideal imaging effects. After estimating the average brightness of the image, in order to determine the illumination effective regions in the image, the non-linear brightness adjustment curves for three types of images are designed, and the Illumination Effective Region Map is generated according to them. For normal exposure images with moderate brightness, a nonlinear curve as shown in Fig. 5(a) is designed. The values of the pixels with brightness between 64 and 191 are set to 1 in the Illumination Effective Region Map, and the values of the pixels out of this range are nonlinearly reduced. Fig. 5(b) is the brightness adjustment curve for over-exposed images. For most of the pixels in this type of image, the brightness is too high, while the amount of the low brightness pixels is small. So, the values of the pixels with low brightness (less than 128) are adjusted to 1 in the Illumination Effective Region Map, while the values of the pixels with brightness higher than 128 are suppressed nonlinearly. The higher the brightness, the greater the degree of suppression. On the contrary, for the images with insufficient brightness, we suppress the values of low brightness pixels and set the values of the pixels with brightness higher than 128 to 1. The curve is shown in Fig. 5(c).

b. ILLUMINATION EFFECTIVE REGION MAP

In order to preserve the essential detail information of the visible image captured in unfavorable illumination environment in the fused image, the illumination effective regions of the visible image are extracted in the image preprocessing stage. The so-called illumination effective region is the area whose average brightness is in the identifiable range of human eyes. First, the enhanced and denoised visible image is transformed into the YUV space. Y stands for the luma component (the brightness) and U and V are the chrominance (color) component of the image. So, the average brightness of the image can be estimated by Y component. Y ranges from 0 to 255. So, we divide 256 brightness levels into three sections (as shown in Fig. 4), which are I.0-63, II.64-191 and III.192-255. Among them, the images with average brightness in Dark section are set as the images with insufficient illumination, and the images with average brightness in Normal section are set as the images with normal exposure, while the images with average brightness in Bright section are set as over-exposed images.

![FIGURE 4. Brightness sections of visible images](image)

In the brightness range of 0-255, the pixels whose brightness values are located in the middle area are more suitable for human eye observation and are easier to be processed by computers, while the pixels corresponding to the brightness...
2) INFRARED IMAGE PREPROCESSING

Infrared images are the reflection of infrared radiation of the subjects, so they will not be affected by environment illumination. Contrast salient areas of infrared images are often of great importance for target recognition. Therefore, the proposed algorithm extracts the visual salient regions of the infrared image and generates the visual saliency map before image fusion, which is used as one of the basis for adjusting the image fusion weight of the base layer.

In this paper, the algorithm of literature [9] is used to extract the salient regions of infrared images. Its main idea is to extract the continuous region with large gray scale in the image by using the morphological principle. The effect of the algorithm is shown in Fig. 7.

FIGURE 7. Saliency map of infrared image (a) Infrared image (b) Visual saliency map

B. IMAGE MULTISCALE ANALYSIS

In this paper, a multiscale image decomposition method based on the combination of Gaussian Filter and Rolling Guidance Filter is proposed. Gaussian Filter is a linear smoothing filter commonly used in the field of image processing. It is effective for image denoising and smoothing. The steps of image decomposition are as follows:

1. Gaussian Filtering is performed on the original image $I$ to obtain a smoothed image $I_g^0$ which is a low-frequency image after removing some detail information from $I$, as shown in (4):

$$I_g^0 = \text{gaussian}(I, \sigma_g^0)$$

where $\sigma_g^0$ determines the smoothness of the image.

2. The original image $I$ is filtered by Rolling Guidance Filter to obtain image $I_r^0$. The details of $I$ are removed but strong edges are preserved in $I_r^0$, as shown in (5):

$$I_r^0 = \text{RollingGuidance}(I, \sigma_r^0, \sigma_r^0, N)$$

where $\sigma_r^0$ and $\sigma_r^0$ control the spatial and range weights respectively, $N$ is the decomposition scale.

3. Step 1 and 2 show that the difference between $I$ and $I_g^0$ is the detail feature of the image at the current scale, denoted as $D^0$, and the difference between $I_r^0$ and $I_g^0$ is the strong edge feature of the image at the current scale, which is recorded as $E^0$ in (6).

$$D^0 = I - I_g^0$$
$$E^0 = I_r^0 - I_g^0$$

4. Adjust the values of $\sigma_g^i$ and $\sigma_r^i$ to meet the needs of smoothing at different scales $i$, as shown in (7).

$$\sigma_g^i = \sigma_g^{i-1} \times 2$$
$$\sigma_r^i = \sigma_r^{i-1}/2$$

5. Gaussian Filter and Rolling Guidance Filter based on $I_g^{i-1}$ are used to calculate the detail layer of the $i$-th scale. Each scale contains the detail information layer $D^i$ and the strong edge information layer $E^i$, such as (8).

$$I_g^i = \text{gaussian}(I_g^{i-1}, \sigma_g^i)$$
$$I_r^i = \text{RollingGuidance}(I_g^{i-1}, \sigma_r^i, \sigma_r^i, N) (i > 0)$$
$$D^i = I_g^i - I_r^i$$
$$E^i = I_r^i - I_g^i$$

6. The basic layer $B$ of the image is the Gaussian Filtered image of layer $N$, which does not include details and strong edge information. It is the smoothing part of the image, as shown in (9).

$$B = I_g^N$$

7. The original image can be reconstructed by (10).

$$I = B + \sum_{i=0}^{N} (D^i + E^i)$$

Fig. 5 shows the multiscale decomposition effect of the proposed algorithm.
The decomposition results of Fig. 8 show that the original image can be decomposed into different scales, and the image coefficients at each scale can be decomposed into detail features and strong edge features. By comparing with the original image, it can be proved that the image reconstruction effect of the algorithm is ideal, the texture and edge information is complete and there are no halo artifacts.

C. IMAGE FUSION RULES

After multiscale decomposition, the original image is decomposed into base layer \( B \), \( D \) component and \( E \) component at different scales, where \( D \) and \( E \) correspond to image detail textures and strong edge features, and their fusion effect determines the details of the fused image. The base layer \( B \) reflects the smooth background of the image. The fusion effect of layer \( B \) determines the overall brightness and contrast of the image.

1) THE FUSION RULE OF BASE LAYER

We have extracted the illumination effective regions of the visible image, and generated the Illumination Effective Region Map in the previous step. The higher the gray value of the pixels in the map, the more weight the corresponding region should be given in the fused image, and the visual salient regions extracted from the infrared image have the same significance as well. For this reason, a weighted average fusion rule based on Illumination Effective Region Map and visual saliency map is proposed for the base layer, as shown in (11).

\[
B_F(i,j) = w \times B_v(i,j) + (1 - w) \times B_{IR}(i,j) \]

Where \( B_v(i,j) \) is the visible image, \( B_{IR}(i,j) \) is the infrared image, \( w = 0.5 + \left( \frac{\text{Map}_v(i,j)}{\text{Map}_v(i,j) - T} \right) \times w_1 \times \text{Map}_{IR}(i,j) \times w_2 \)

where \( \text{Map}_v(i,j) \) represents the value corresponding to the pixel located at \((i,j)\) in the Illumination Effective Region Map of the visible image. The larger the value of \( \text{Map}_v(i,j) \), the more ideal the local brightness corresponding to the pixel is, so larger fusion weight should be assigned to the pixel. Therefore, for the visible image illumination Effective Region Map, a threshold \( T \) is set to adjust the influence of the illumination factors on the visible image fusion weight. When the pixel value in the map is higher than the threshold, the fusion weight should be larger, and vice versa. \( \text{Map}_{IR}(i,j) \) stands for the value corresponding to the pixel located at \((i,j)\) in the visual saliency map of the infrared image. In the proposed algorithm, based on the brightness adjustment curve shown in Fig. 5, the threshold \( T \) is set to 0.5. \( w_1 \) and \( w_2 \) are used to adjust the influence of the illumination and infrared image salient regions on the fusion weight respectively. The values of \( w_1 \) and \( w_2 \) in the algorithm are set to 0.4 and 0.2.

2) THE FUSION RULE OF DETAIL LAYERS

The key point of the detail layers fusion is to ensure the textures and the edges are clear. This paper applies the choose-max scheme on the basis of the absolute value of the local variance to detail layers fusion. In the image multiscale decomposition process, the original images have been decomposed into detail feature \( D \) and strong edge feature \( E \) at different scales. We select the detail feature image \( D \) at each scale and calculate the absolute value of the local variance of each pixel, since it reflects the richness of the image texture at the current scale. It is used as a basis to set the fusion weights of the detail layers. The fusion rule of detail layers is shown as (12):

\[
D_F(i,j) = \begin{cases} 
D_v(i,j) & \text{if } \text{abs}(\text{Var}(D_v(i,j))) \geq \text{abs}(\text{Var}(D_{IR}(i,j))) \\
D_{IR}(i,j) & \text{otherwise}
\end{cases}
\]

\[
E_F(i,j) = \begin{cases} 
E_v(i,j) & \text{if } \text{abs}(\text{Var}(E_v(i,j))) \geq \text{abs}(\text{Var}(E_{IR}(i,j))) \\
E_{IR}(i,j) & \text{otherwise}
\end{cases}
\]

where, \( \text{abs}(\text{Var}(D_v(i,j))) \) indicates the absolute value of the local variance of the visible image \( D \) layer at a certain scale, \( \text{abs}(\text{Var}(D_{IR}(i,j))) \) denotes the same value of the infrared image. In order to ensure the uniformity of detailed features and strong edge features fusion at the same scale, the fusion rules of \( D \) layers and \( E \) layers are the same.

III. Experiment and Simulation

In order to test the effectiveness of the algorithm, the experiment fused five sets of infrared and visible images captured in natural illumination environment of the same scene at different time of the day. Meanwhile, we chose several algorithms which have better fusion effect proposed in recent years to fuse the above images and compare their fusion effects. The algorithms being compared are image fusion through infrared feature extraction and visual information preservation (FEVIP) [10], image fusion based on visual saliency map and weighted least square optimization (VSM&WLS) [5] and image fusion using latent low-rank representation (LLRR) [11]. The fusion effects are shown in Fig.9.

From the fusion effects of Fig. 9, the images fused by the proposed algorithm have clear contrast. The overall brightness of the image is consistent with human visual perception. The details are clear, and the salient regions of the infrared image are completely preserved. Compared with other algorithms, the fusion effects of our algorithm are less affected by poor illumination conditions of visible images.
and Entropy, to compare the effects of several fusion algorithms. The larger the values of the evaluation parameters are, the better the performance of the algorithm is. The data in Table 1 proves that the proposed algorithm can preserve the original image details and salient features well, and the images fused by the proposed algorithm have larger spatial resolution and richer detail information than that of the algorithms being compared. FEVIP can get a clear fusion effect under normal illumination condition but it cannot show an ideal fusion performance while fusing the over-exposed visible images. VSM&WLS and LLRR can highlight the visual salient regions of the infrared images but the details of their fused images are not clear and sharp especially in poor visibility conditions. By comparison, it can be found that the proposed algorithm has better adaptability to different illumination conditions, and the fusion effect is more stable.

### Table 1

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameters</th>
<th>FEVIP</th>
<th>VSM&amp;WLS</th>
<th>LLRR</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spatial Frequency</td>
<td>0.0273</td>
<td>0.0225</td>
<td>0.0166</td>
<td>0.0349</td>
</tr>
<tr>
<td></td>
<td>Entropy</td>
<td>6.2965</td>
<td>5.5505</td>
<td>5.5218</td>
<td>6.3263</td>
</tr>
<tr>
<td>2</td>
<td>Spatial Frequency</td>
<td>0.0324</td>
<td>0.0385</td>
<td>0.0292</td>
<td>0.1022</td>
</tr>
<tr>
<td></td>
<td>Entropy</td>
<td>6.3328</td>
<td>5.8152</td>
<td>5.8384</td>
<td>6.8169</td>
</tr>
<tr>
<td>3</td>
<td>Spatial Frequency</td>
<td>0.1498</td>
<td>0.1959</td>
<td>0.1020</td>
<td>0.1995</td>
</tr>
<tr>
<td></td>
<td>Entropy</td>
<td>7.0943</td>
<td>7.3915</td>
<td>7.0439</td>
<td>7.6624</td>
</tr>
<tr>
<td>4</td>
<td>Spatial Frequency</td>
<td>0.1070</td>
<td>0.0828</td>
<td>0.0477</td>
<td>0.1237</td>
</tr>
<tr>
<td></td>
<td>Entropy</td>
<td>2.8867</td>
<td>6.4184</td>
<td>6.3767</td>
<td>6.6672</td>
</tr>
<tr>
<td>5</td>
<td>Spatial Frequency</td>
<td>0.0773</td>
<td>0.0580</td>
<td>0.0329</td>
<td>0.0821</td>
</tr>
<tr>
<td></td>
<td>Entropy</td>
<td>1.5685</td>
<td>5.8600</td>
<td>5.9163</td>
<td>6.2632</td>
</tr>
</tbody>
</table>

### IV. CONCLUSION

Aiming at the problem that the fusion effect of infrared and visible image is not ideal due to the influence of illumination on visible images, a fusion algorithm of infrared and visible image adapted to all-weather and poor illumination environments is proposed. The main work of this paper includes: preprocessing of visible and infrared image to adjust the contrast of visible image and extract the salient regions of infrared image, calculating and generating the Illumination Effective Region Map of visible image and using it as the basis for setting image fusion weights, designing a multiscale decomposition method which combines Rolling Guidance Filter and Gaussian Filter to suppress the edge halo phenomenon. Fusion rules are designed for the decomposed base layer and detail layers respectively, ensuring the clarity and integrity of the detail information and the salient features, while adjusting the overall brightness of the fused image.

### REFERENCES


