J. Meas. Sci. Instrum., 2024, 15(2): 176-184.
DOI: 10.62756/jmsi.1674-8042.2024018
http://jmsi.publish.founderss.cn
jmsi@nuc.edu.cn

Infrared and visible image fusion algorithm based on Retinex-enhanced multiscale decomposition

GUO Yin, DU Lixia*

School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

***Corresponding author:** DU Lixia (dlx1228@mail.lzjtu.cn)

Received: February 19, 2023 **Revised:** March 28, 2023 **Accepted:** May 11, 2023

Abstract: Aiming at the problems of poor contrast of fusion results, blurring of target margins and loss of background detail information under low illumination conditions of traditional infrared and visible image fusion algorithms, an infrared and visible image fusion algorithm based on multi-scale decomposition with Retinex enhancement was proposed. Firstly, the single-scale Retinex (SSR) algorithm information enhancement process was performed on the weak visible image using Retinex. Secondly, the source image was multi-scale decomposed using cross bilateral filtering to successively obtain the image information of the base layer image and the detail layer, and the fusion method combining the absolute value maximization strategy and guided filtering was used for the base layer image, and the fusion method of constructing weight map and significant map was used for the detail layer image. Finally, the processed base layer image and detail layer image were weighted to obtain the fused image. From the subjective analysis, the proposed method could effectively extract and fuse the important information in the source image, and obtain the image with high fusion quality and natural and clear visual effect. From the objective evaluation, the average accuracy of the proposed method on AG, SF, CE, and FMI was optimal when compared quantitatively with various fusion results.

Key words: image fusion; infrared and visible images; Retinex algorithm; multiscale decomposition; guided filtering

0 Introduction

Image fusion has been a hot research topic in the field of optical image processing. The purpose of the research is to fuse images of the same scene acquired by multiple sensors, so that the information can be complementary. Among them, infrared (IR) and visible sensors have received much attention. Infrared sensors have good detection performance and can acquire images of objects in low-light conditions or with obstacle occlusion. But the acquired infrared images lack texture detail information and have low contrast. While visible images acquired by visible sensors have rich texture detail information and high contrast, which is beneficial to the human eye's eyeperception of the scene. However, the images acquired in low-light or obstacle-obscured conditions are less effective. The effective fusion of infrared image and visible image can highlight the infrared target information and obtain the high-resolution detail information of visible image, so it is widely used in computer vision, aerospace, and military fields[1].

There are many methods for image fusion, and pixel-level image fusion in grayscale space is one of the more commonly used methods[2]. Several pixel-level image fusion methods are proposed for infrared and visible image fusion, and multi-scale fusion method is one of the most widely used methods[3-6]. The representative methods include anisotropic diffusion fusion (ADF)[7] and cross bilateral filter (CBF)[8]. ADF uses anisotropic diffusion process to filter each source image to extract the base and detail layers, but the fusion effect is poor. CBF uses cross bilateral filter to decompose the image and maintain the edge information of the image on the basis of extracting the detail image, but this algorithm does not consider the visibility of the target in the dark region, so it is less effective in fusing images with darker illumination.

With the development of deep learning, the applications in image fusion are becoming more and more widespread[9]. Prabahakar et al. proposed a simple CNN-based image fusion algorithm, but the depth features of this method are not well utilized[10]. To address these drawbacks, Liu et al. proposed a fusion framework based on pre-trained networks[11]. Firstly, the high-frequency components and low-frequency

© The Author(s) 2024. Published by Publishing Center of North University of China. The articles published in this open access journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

components of the source image are derived after image decomposition. The low-frequency components are fused using an averaging strategy, and the fused high-frequency components are obtained using a deep learning framework. The fused low-frequency components are combined with the fused high-frequency components, and the fused image is derived after reconstruction. However, the image decomposition method of this method is very simple, and the benefits of deep learning are not well represented.

Using deep learning methods suffers from the problem that the network is difficult to train when there is insufficient training data for IR and visible images. Liu et al.[12] proposed low rank representation (LRR). And then to solve the problem that this representation did not retain local structural information, potential low rank representation (LatLRR) was proposed[13]. Li et al.[14] proposed the use of LatLRR for IR and visible image fusion, and used a similar averaging fusion strategy for the low-rank and saliency parts. However, because the potential low-rank decomposition is still not complete, the conventional averaging fusion rules for the low-rank and saliency parts cannot fully utilize the information at different frequencies, which usually leads to decrease in the contrast of the fused images.

A multi-scale image fusion method based on Retinex algorithm was proposed in this paper. The fusion results were effectively improved with poor contrast, diminished target information, and unclear background details. In the paper, the Retinex algorithm was first used to enhance the visible light images. For those weak visible images, the single-scale Retinex (SSR) algorithm can well enhance the image detail information, improve the clarity, and retain the original information better. Secondly, the multi-scale decomposition of the source image is carried out by using cross-bilateral filtering to obtain the base image information and the detail layer image information in turn. The fusion method combining the absolute value maximization strategy and guided filtering is used to construct the fusion method. Finally, the fused image is obtained by weighting the base layer image and the detail layer image. The algorithm retains the details from the source image as much as possible and takes into account the spatial consistency of the extracted image salient information. The experimental results showed that the enhancement effect of the image was greatly improved, and the important information such as target, background details, and edges were effectively maintained.

1 Related theories

1.1 Retinex theory

Due to the influence of illumination and other factors which lead to low contrast and poor visibility in the dark region part of the visible image, the Retinex algorithm is proposed to restore the essential information of the image to enhance the image contrast[15]. Retinex consists of two phrases: Retina and Cortex. Retinex theory, as a model of luminance and color perception of human vision, reveals that the object color changes are not affected by changes in the illuminated environment, but are caused by reflections from the object surface with consistency. Therefore, the observed figure $F(x,y)$ can be decomposed into two different images: the reflective image $R(x,y)$ and the illuminated image $L(x,y)$. $L(x,y)$ can be considered as an inhomogeneity of illumination, which is the main factor affecting the image, while $R(x,y)$ is considered as the original image that is not affected by external influences, and its mathematical model is expressed as

$$F(x,y)=R(x,y)L(x,y). \tag{1}$$

From the mathematical model of Retinex theory[16], the illuminated image $L(x,y)$ can be regarded as the well-behaved noise attached to the ideal image $R(x,y)$. The purpose of image enhancement is to remove the influence of $L(x,y)$ from the observed image $F(x,y)$ and restore the essential information of the scene, i. e., $R(x,y)$, for the purpose of image enhancement. And the SSR image enhancement algorithm translates Eq. (1) into the logarithmic domain to calculate the incident image $R(x,y)$. SSR can be expressed as

$$\begin{aligned}\log R(x,y)&=\log F(x,y)-\log L(x,y)=\\ &\log F(x,y)-\log\left[G(x,y)*F(x,y)\right],\end{aligned} \tag{2}$$

where * is the convolution operator; $F(x,y)$ is the standard surround function; and $\iint G(x,y)\mathrm{d}x\mathrm{d}y=1$. The observed image $F(x,y)$ is used to estimate $L(x,y)$. The estimate of illumination is obtained by processing the ratio between the pixels in $F(x,y)$ and their surround-weighted mean values. The Gaussian function used in SSR is

$$G(x,y)=\omega\exp\left(-\frac{x^2+y^2}{\sigma^2}\right), \tag{3}$$

where σ is the scale factor and plays an important role in the processing of SSR algorithms. When σ is small, SSR

has a higher dynamic range compression capability and better image detail enhancement, but leads to a certain degree of tonal lightening. When σ is large, SSR reduces the dynamic range compression ability, but maintains the color fidelity of the output image well, and the image details are not complete. Therefore, the parameter σ should be chosen with respect to the size of the original image to balance the computational complexity and the quality of the fused image. ω is the normalization factor. $R(x,y)$ can be obtained by[17]

$$R(x,y)=\exp\left\{\log F(x,y)-\log\left[G(x,y)*F(x,y)\right]\right\}. \tag{4}$$

The computational flow of SSR is shown in Fig.1.

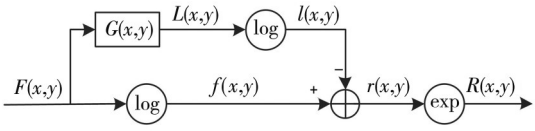

Fig. 1 SSR operation flow chart

1.2 Cross bilateral filtering theory

Cross bilateral filtering (CBF) not only has the advantage of smoothing the image, but also preserves the edge information of the image. CBF is a non-linear, non-iterative local technique that combines an edge stop function with a low-pass filter, resulting in a kernel of the function decay filter when the intensity difference between pixels becomes large. In summary, cross bilateral filtering was chosen for multi-scale decomposition of infrared images with visible images.

The CBF shaping filter kernel uses the geometric proximity in the image and the similarity of the grayscale of neighboring pixels while filtering the image B. The CBF output of the image at the pixel is

$$B_{\mathrm{CBF}}(p)=\frac{1}{W}\sum_{q\in S}G_{\sigma_s}\left(\|p-q\|\right)G_{\sigma_r}\left(A(p)-A(q)\right)B(q), \tag{5}$$

where σ_s is the size of the spatial neighborhood used to filter a pixel, and σ_r controls how much an adjacent pixel is downweighted. The Euclidean distance between pixel p and q is $\|p-q\|$, $W=\sum_{q\in S}G_{\sigma_s}\left(\|p-q\|\right)\cdot G_{\sigma_r}\left(A(p)-A(q)\right)$ is a normalization constant, the range of the adjacent region of pixel p is S, the edge stopping function is $G_{\sigma_r}\left(A(p)-A(q)\right)=\mathrm{e}^{\frac{|A(p)-A(q)|^2}{2\sigma_r^2}}$, and geometric proximity is $G_{\sigma_s}\left(\|p-q\|\right)=\mathrm{e}^{\frac{\|p-q\|^2}{2\sigma_s^2}}$.

The standard deviation of noise in the image is σ_n.

Setting parameters σ_s and σ_r affects the performance of the bilateral filter varies with σ_n. The assignment of σ_s is usually taken in the range (1.5, 2) and its values are relatively independent. The value of σ_r/σ_n is usually in the interval (1,3).

1.3 Bootstrap filtering theory

The bootstrap filter is a linear filter with good edge-preserving properties. The core idea of the filter is to use the guide map to guide the filter to perform certain range of smoothing and certain edge boundary region preservation operations, so as to complete the image edge preservation and frame retention functions, thus it can be found that the filter uses the guide map to guide the whole filtering process. Nowadays, guided filtering has a good application in the field of image processing. The specific implementation of bootstrap filtering is as follows.

Let I be the input image, J be the bootstrap map, and the bootstrap filtering process can be expressed as

$$G(I,J,r,\varepsilon)=\bigcup_{i\in I}\left\{a_kJ_i+b_k\middle|\, i\in w_k\right\}, \tag{6}$$

where the coefficients are a_k and b_k, and the filter window is w_k (a square with side r), which can be calculated by

$$a_k=\frac{\frac{1}{|w|}\sum_{i\in w_k}J_iI_i-\bar{J}_k\bar{I}_k}{\sigma_k^2+\varepsilon}, \tag{7}$$

$$b_k=\bar{I}_k-a_k\bar{J}_k, \tag{8}$$

where the regularization parameter is ε; the variance is σ_k^2, and the mean values of images I and J within w_k are denoted as $\bar{I}_k$ and $\bar{J}_k$.

2 Improved infrared and visible light image fusion method

Starting from the characteristic of how infrared images are imaged, the SSR algorithm is performed on the weak visible image, and the image information is enhanced by the processing method, and Retinex is not used directly. The advantage of this approach is that it is better to retain the detail information in the original image while improving the clarity of the image. Secondly, multi-scale decomposition is performed on the image to be fused to get the detail layer and the base layer images, and different fusion rules are used for the different layers. Finally, the images of the detail layer and the base layer are superimposed to get the final fused image result. The convergence framework is shown in Fig.2.

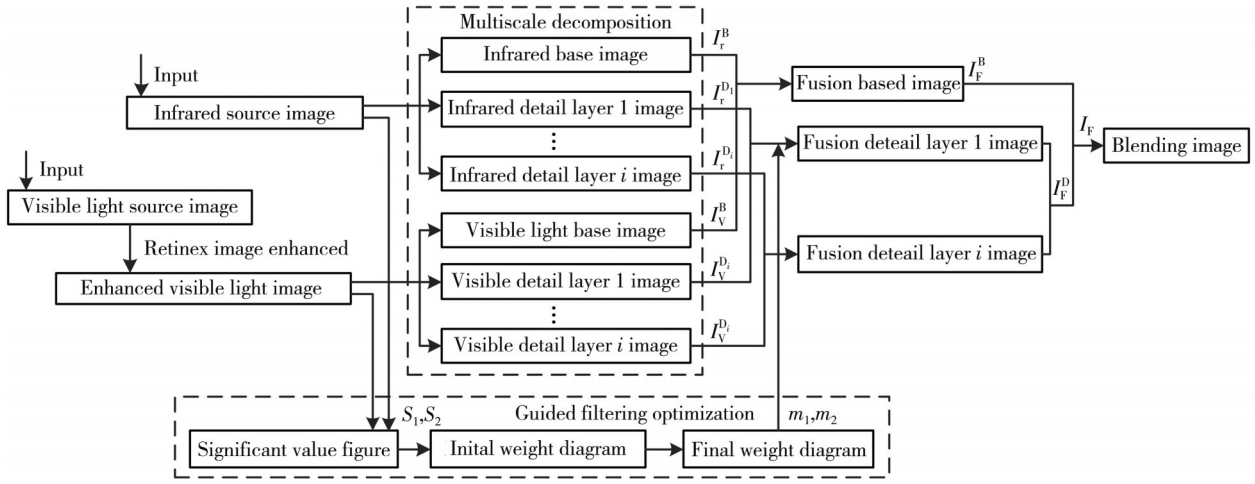

Fig. 2 Flow diagram of image fusion method

Step 1: Select visible and infrared images of the same size and resolution of $m \times n$ (m, n are multiples of 4).

Step 2: SSR processing is performed on the selected visible image, and in order to allow the quality of the image to be further enhanced, a suitable σ value needs to be selected according to the characteristics of the image.

Step 3: Multi-scale decomposition of infrared images and enhanced visible images is carried out by CBF. Let I_r be the infrared image, I_v be the enhanced visible image, and $S_{CBF}(I_1,I_2,\sigma_s,\sigma_r)$ be the CBF operation symbol, where the input images to the CBF are I_1 and I_2. The base layer of the infrared image in the first layer is I_r^{B}, the detail layer of the infrared image in the ith layer is $I_r^{D_i}$, the base layer of the visible image in the first layer is I_v^{B}, and the detail layer of the visible image in the ith layer is $I_v^{D_i}$. Then multi-layer base and detail layers can be obtained by CBF decomposition operation, and Eq. (9) is the first layer of base and detail layers.

$$\begin{cases} I_r^{B_1} = S_{CBF}(I_r,I_v,\sigma_s,\sigma_r), & I_r^{D_1} = 1 - I_r^{B_1}, \\ I_v^{B_1} = S_{CBF}(I_v,I_r,\sigma_s,\sigma_r), & I_v^{D_1} = 1 - I_v^{B_1}. \end{cases} \tag{9}$$

Using the same principle, the base and detail layers of successive layer i are

$$\begin{cases} I_r^{B_i} = S_{CBF}(I_r^{B_{i-1}},I_v^{B_{i-1}},\sigma_s,\sigma_r), & I_r^{D_i} = I_r^{B_{i-1}} - I_r^{B_i}, \\ I_v^{B_i} = S_{CBF}(I_v^{B_{i-1}},I_r^{B_i},\sigma_s,\sigma_r), & I_v^{D_i} = I_v^{B_{i-1}} - I_v^{B_i}, \end{cases}$$

$$i = 1,2,\cdots,n. \tag{10}$$

Step 4: The images of the base layer are fused by taking the maximum value criterion for the images, along with the optimization of the guided filtering.

The initial weight values of the base layer images are calculated by

$$w_1^{B} = \begin{cases} 1, & |I_r^{B_n}| > |I_v^{B_n}|, \\ 0, & \text{else}, \end{cases} \tag{11}$$

$$w_2^{B} = \begin{cases} 1, & |I_v^{B_n}| > |I_r^{B_n}|, \\ 0, & \text{else}. \end{cases} \tag{12}$$

Among them, the weight maps of the base layers of the infrared and visible images are w_1^{B} and w_2^{B}. In this paper, for the existence of some problems of poor edge consistency, a guided filter was added after the pre-weight map with the purpose of smoothing their images while performing edge preservation operations. Guided filtering was used to check the consistency of the initial weight map to generate the weight map.

$$m_1 = G_{r,\varepsilon}(w_1^{B},I_r^{B_n}), \tag{13}$$

$$m_2 = G_{r,\varepsilon}(w_2^{B},I_v^{B_n}), \tag{14}$$

where the final weight maps of the detail layer images are m_1 and m_2, resulting in the final fused image I_F^{B} of the base layer image is

$$I_F^{B} = m_1 I_r^{B_n} + m_2 I_v^{B_n}. \tag{15}$$

Step 5: The detail layer image design weighted average fusion rule is used to construct the weights using the saliency map, which preserves the salient features in the source image. The saliency maps S_1^i and S_2^i of the infrared and visible image layer i are $S_1^i = |I_v^{D_i}|$ and $S_2^i = |I_v^{D_i}|$, respectively, and the calculated weights can be obtained according to their saliency maps.

$$w_1^{D_i} = \frac{S_1^i}{S_1^i + S_2^i}, \quad i = 1,2,\cdots,n, \tag{16}$$

$$w_2^{D_i} = \frac{S_2^i}{S_1^i + S_2^i}, \quad i = 1,2,\cdots,n, \tag{17}$$

where the weights of the infrared image and the visible image layer i are plotted as $w_1^{D_i}$ and $w_2^{D_i}$, which leads to the final detail layer fusion image I_F^{D} as

$$I_F^{D} = \sum_{i=1}^{n} w_1^{D_i} I_r^{D_i} + w_2^{D_i} I_v^{D_i}. \tag{18}$$

Step 6: The fused base layer image and detail layer

image are superimposed to generate the final fused image.

$$I_F = I_F^B + I_F^D. \tag{19}$$

3 Results and discussion

3.1 Experimental results

The experimental platform was MATLAB R2016a, and the computer configuration was Intel(R) processor with CPU main frequency 2.0 GHz and 8 GB memory. In this paper, two different sets of experimental images were selected from the dataset widely used for image fusion and the open source dataset TNO. The first set of experimental images were seven pairs of images from Zhang et al.[21], where five pairs of sample images of tree, camp, jeep, captein and park are shown in Fig.3(a), (b), (c), (d), and (e). The second group of experiments were the 2 pairs of color images parking lot in the method of Zhang et al.[21], as shown in Fig.3(f) and (g).

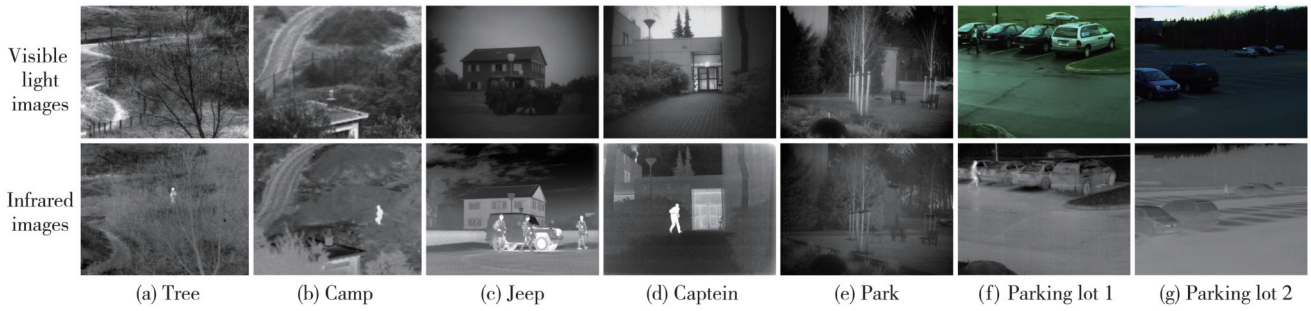

Fig. 3 Five pairs of grayscale image sample source((a) Tree; (b) Camp; (c) Jeep; (d) Captain; (e) Park); and two pairs of color image sample source((f) Parking lot 1; (g) Parking lot 2)

The key to a good fusion method is an effective image information extraction and an appropriate fusion principle that allows the extraction and integration of useful information from the source image without introducing any artifacts in the fused image processing. In order to verify the effectiveness and superiority of the method in this paper, comparison experiments were performed on two sets of grayscale images. The comparison was made with ADF[7], CBF[8], CNN[11], and LatLRR[14] methods on the 1st set of experimental images.

The 2nd set of experiments were to compare 2 color images with other traditional and advanced algorithms. Both subjective analysis and objective evaluation were compared. The visual effect of the fusion result was compared subjectively. Objectively, average gradient (AG), edge intensity(EIN), spatial frequency(SF), cross entropy(CE), feature mutual information(FMI), and peak signal to noise ratio(PSNR) were chosen as evaluation indexes to evaluate the fused images. AG, EIN, SF, and PSNR were used to directly evaluate the fused images, and the higher value of these four fusion evaluation indexes indicated the higher quality of the fused images. The higher the value of FMI and the lower the value of CE, the higher the quality of fusion.

3.2 Subjective analysis

Fig.4 shows the fusion results of ADF, CBF, CNN, LatLRR fusion methods, and the proposed methods on tree, camp, jeep, captein, and park images. The first and second rows showed the fusion results obtained by the four fusion methods with camp and dune as the input images, respectively. In particular, the road, grass, and woods in row 1 and the hills, roads, and house roofs in row 2 were clear, but these were not clear in the fusion results of ADF, CBF, and LatLRR, which introduced a lot of artificial artifacts to the infrared images and was more in line with the human visual observation. The fusion results of ADF and CBF were not clear for people, street lights, fences and around the tree canopy, and more information about the infrared target was maintained. In contrast, the fusion results obtained by the method in this paper were naturally clear, avoiding the target information being diminished under low-light conditions, and were closer to the ideal fused image. For example, the details of flowers and trees in ADF and CBF were lost, there were a lot of artifacts, and the targets of people were not prominent, so the overall clarity of these methods was not high. The CNN and LatLRR methods were not obvious enough, and the road marks were not clear, which was not conducive to long time observation by human eyes. In contrast, the fusion results of this paper were natural, and the parts of people, houses, road marks, and tree branches bordering the sky were clearly discernible, almost without artifacts, which indicated that the proposed method could reasonably allocate infrared and visible spectral information. The fifth row showed the fusion results of park image pairs. It could be seen that the method in this paper showed better edge preservation

ability compared with other methods. The fusion result of ADF, CBF and CNN was low contrast and the overall visual effect was dark. The main reason for this problem was that these three algorithms did not enhance the dark area before fusion, resulting in poor effect. The fusion effect of the LatLRR algorithm was good. For example, the algorithm could enhance the dark objects such as feces in the figures, but the enhancement effect of the algorithm on the ground weeds and other information was poor. According to the results of Fig. 5, it could be seen that the proposed method could comprehensively and accurately extract information from infrared and visible light source images and transfer it to the fused image, effectively solving problems such as reduced target information, unclear background details, and blurred edges in the image fusion process.

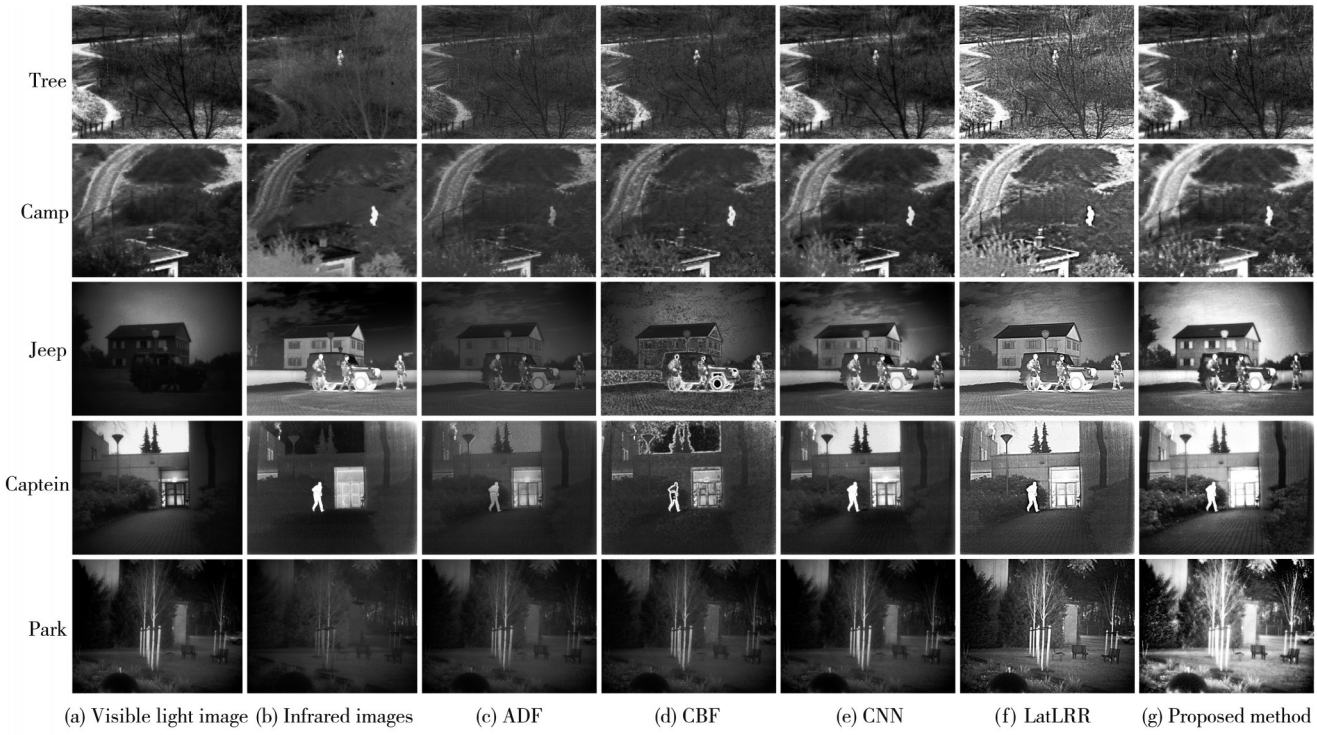

Fig. 4 Experiment 1: Fusion results of gray image by 5 kinds of algorithms

Fig. 5 shows the fusion results of the two pairs of parking lot color sample images in the second group of experimental images. It could be seen that the ADF algorithm had the worst visual effect in the fusion result of image 1 and image 2, because the contrast was low due to the influence of daylight and darkness, so the fusion result had poor contrast and the details of pedestrians and vehicles were difficult to distinguish.

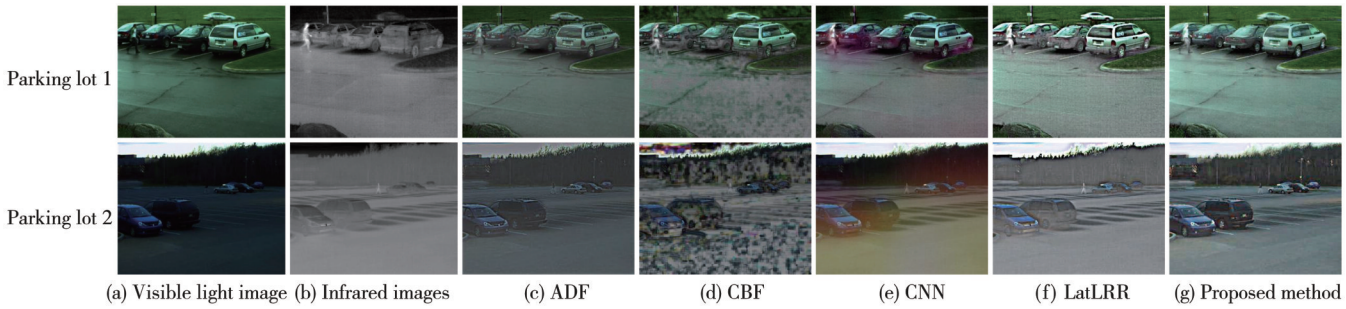

Fig. 5 Experiment 2: Fusion results of color image by 5 kinds of algorithms

The CBF algorithm uses the idea of weighted averaging to fuse the base and detail layer images, which is easy to introduce other interference information other than image information, making many noise spots in the fusion result, especially the ground and vehicle spots, and even causing confusion and difficulty in distinguishing each object target, and the overall visual effect was deviated. The contrast between cars and roads in the CNN algorithm in the two groups of images was not obvious, and it was difficult to distinguish the main targets of the images when the light was not bright. And in the second group of images the overall tone was yellowish, and the distant jungle and the group of cars were easy to cause confusion, and the results were

relatively defective. For LatLRR algorithm, the resolution of the fused image was higher than that of other algorithms, but the contrast of the two structural images observed after integration was not obvious, and the resultant color was more similar. If the observation time was slightly longer, it would cause visual fatigue.

Comprehensive evaluation of the proposed algorithm showed a better visual effect and effectively solved the problems of diminished target information, unclear background details, blurred edges, and inconspicuous contrast in image fusion.

3.3 Objective analysis

The values of the six evaluation indexes for the different results of the above five fusion methods in seven pairs of images are shown in Fig.6. In order to accurately conduct comparative analysis among the data and avoid extreme data interference, the objective evaluation values of each fusion method were obtained by means of average value calculation. In addition, Mean values of each index of different fusion methods and the running time of 7 pairs of images are compared in Table 1.

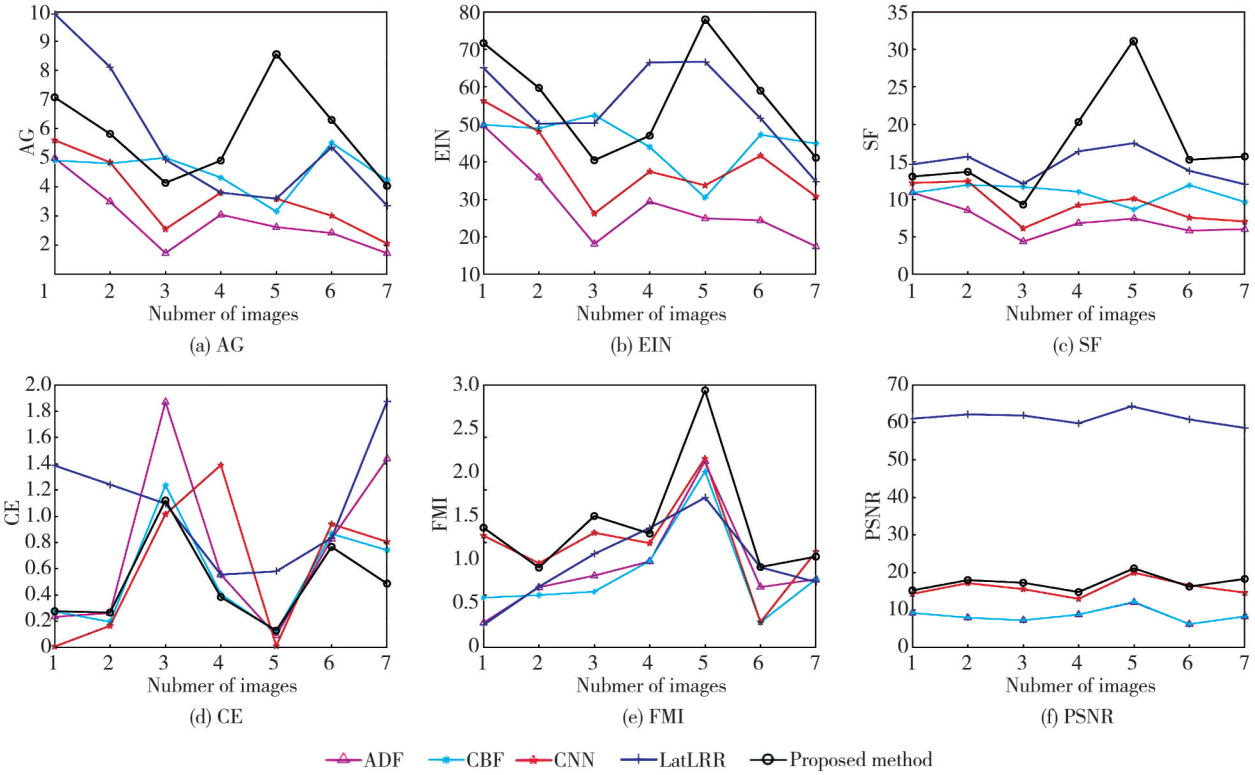

Fig. 6 Objective evaluation on various fusion methods of 7 pairs of infrared and visible images

Table 1 shows that the overall advantage of the proposed method is outstanding, except for a few index values that are not optimal. The mean accuracy of AG, SF, CE, and FMI were optimal, while the mean values of EIN and PSNR were not optimal after the LatLRR method. The main reason was that the LatLRR algorithm performed multi-layer feature extraction through a deep learning framework, had an overall constraint on the coefficient matrix, and the noise part was removed under the constraint, and had a better ability to capture information on both the local part and the global part. The LatLRR method had a clearer overall contour of the fused image, retained more local information, and thus had better edge strength and peak signal-to-noise ratio indexes.

Overall, the proposed method had obvious advantages over the existing multi-scale methods in six evaluation indexes.

Table 1 Mean values of each index and running time of different fusion methods

Method	AG	EIN	SF	CE	FMI	PSNR	Time/s
ADF	2.860 4	28.550 0	7.137 3	0.753 8	1.259 0	8.547 8	0.471 0
CBF	4.560 6	45.401 8	10.819 0	0.548 3	1.199 8	8.563 4	14.064 5
CNN	3.638 2	39.160 5	9.248 8	0.619 3	1.486 1	15.882 2	47.562 5
LatLRR	5.583 9	**56.694 8**	15.271 7	1.080 9	1.305 3	**61.193 5**	138.847 0
Proposed method	**5.829 3**	55.029 8	**16.931 6**	**0.490 2**	**1.687 9**	17.272 9	12.076 5

The running time of the proposed method was about 4 times faster than CNN and 11 times faster than LatLRR, respectively. However, it was slower than ADF method, might be due to the use of bootstrap filtering, which required iterative operations during the computation. It was worthwhile to increase the running time in order to improve the image fusion effect and increase the application value of the method. Thus, it could be shown that the proposed method can greatly improve the fusion efficiency while maintaining the good fusion effect.

4 Conclusions

Infrared and visible image information in the same scene are usually complementary, and their effective fusion is more suitable for human perception and computer processing. In this paper, a fusion method based on Retinex-enhanced multi-scale decomposition was proposed by combining the idea of multi-scale decomposition with the fusion weight based on significance. The effectiveness and superiority of the method were verified by experiments.

1) For those weak visible images, the SSR algorithm could enhance the detail information of the image, improve the clarity and retain the original information better.

2) The information of infrared image and visible image on large scale is very different. The multi-scale decomposition of the source image by CBF was not only beneficial to the subsequent fusion processing, but also reduced the algorithm complexity of the existing methods.

3) Optimization of the base layer by absolute value enlargement and bootstrap filtering improved the shortcomings of the fusion process, such as artifacts and boundary misalignment. Extracting the significant information of the source image to construct a weight map for image fusion of the detail layer could solve the phenomenon of blurred details in the fused image.

Compared with other existing fusion methods, the proposed method could greatly improve the fusion rate while ensuring high fusion quality, and offered the possibility of real-time fusion of infrared and visible images.

Acknowledgement

This work was supported by National Natural Science Foundation of China (No. 62063014).

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] QI H S, RONG C Z, XIAO L M, et al. Infrared and visible image fusion algorithm based on dual-tree complex wavelet transform and guided filtering. Communications Technology, 2019, 52(2): 330-336.

[2] HUO X, ZOU Y, CHEN Y, et al. Dual-scale decomposition and saliency analysis based infrared and visible image fusion. Journal of Image and Graphics, 2021, 26(12): 2813-2825.

[3] XU S P, LIN Z Y, ZHANG G Z, et al. Low-illumination image enhancement algorithm using a hybrid implementation strategy of deep learning and image fusion. Journal of Electronics, 2021, 49(1): 72.

[4] LI Z, HU H M, ZHANG W, et al. Spectrum characteristics preserved visible and near-infrared image fusion algorithm. IEEE Transactions on Multimedia, 2021, 23: 306-319.

[5] ZHAO F, ZHAO W, YAO L, et al. Self-supervised feature adaption for infrared and visible image fusion. Information Fusion, 2021, 76: 189-203.

[6] GAO X Q, LIU G, XIAO G, et al. Infrared and visible image fusion algorithm based on FPDE. Journal of Automation, 2020, 46(4): 186-194.

[7] BAVIRISETTI D P, DHULI R. Fusion of infrared and visible sensor images based on anisotropic diffusion and karhunen-loeve transform. IEEE Sensors Journal, 2016, 16(1): 203-209.

[8] SHREYAMSHA KUMAR B K. Image fusion based on pixel significance using cross bilateral filter. Signal, Image and Video Processing, 2015, 9(5): 1193-1204.

[9] MA Q, ZHU B, ZHANG H W. A dual-band image fusion method bbased on VGG network. Laser and Infrared, 2019, 49(11): 1374-1380.

[10] PRABHAKAR K R, SRIKAR V S, BABU R V. Deep fuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs//IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 4724-4732

[11] LIU Y, CHEN X, CHENG J, et al. Infrared and visible image fusion with convolutional neural networks. International Journal of Wavelets, Multiresolution and Information Processing, 2018, 16(3): 1850018.

[12] LIU G, LIN Z, YU Y. Robust subspace segmentation by low-rank representation//27th International Conference on Machine Learning, Haifa, Israel, New York: ACM, 2010: 663-670.

[13] LIU G, YAN S. Latent low-rank representation for subspace segmentation and feature extraction//2011 International

Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE, 2011: 1615-1622.
[14] LI H, WU X J. Infrared and visible image fusion using latent low-rank representation. arXiv Preprint, 2018, arXiv:1804.08992.
[15] LI J, LI S J, DUAN X H, et al. Infrared image enhancement based on retinex and probability nonlocal means filtering. Acta Photonica Sinica, 2020, 49(4): 410003
[16] KONG L J, ZHANG M M. Fusion algorithm of low visible light and infrared image based on retinex. Packaging Engineering, 2020, 41(19): 237-244.
[17] CUI Z Y, ZHANG S H. Image enhancement algorithm based on multi-scale retinex and bilateral filter. Laser Journal, 2015, 36(4): 90-93.
[18] QU H C, WANG Y P, GAO J K, et al. Mode adaptive infrared and visible image fusion. Infrared Technology, 2022, 44(3): 268-276.
[19] LIU F. Research of image dehazing and hardware implementation. Hefei: University of Science and Technology of China, 2014.
[20] CHEN G Q. Research on multi-sensor image fusion technology based on multi-scale analysis. Jilin: Jilin University, 2015.
[21] ZHANG X, YE P, XIAO G. VIFB: a visible and infrared image fusion benchmark//IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 468-478.

基于Retinex增强的多尺度分解的红外与可见光图像融合算法

郭 印, 杜丽霞*

兰州交通大学 电子与信息工程学院, 甘肃 兰州 730070

摘 要: 针对传统红外与可见光图像融合算法在低照度条件下融合结果对比度差、目标边际模糊、背景细节信息丢失等问题, 提出一种基于Retinex增强的多尺度分解的红外与可见光图像融合算法。首先, 利用Retinex对弱可见光图像进行单尺度Retinex(SSR)算法信息增强处理; 其次, 利用交叉双边滤波对源图像进行多尺度分解, 先后得到图像的基层图像信息和细节层图像信息, 对基础层图像采用绝对值取大策略和引导滤波相结合的融合方法, 对细节层图像采用构建权重图和显著图的融合方法; 最后, 将处理后的基础层图像和细节层图像加权得到融合图像。从主观分析来看, 本文方法可以有效提取和融合源图像中的重要信息, 得到融合质量高、视觉效果自然清晰的图像。从客观评价来看, 多组图像融合结果表明, 本文所提方法的AG、SF、CE和FMI指标均较好。

关键词: 图像融合; 红外与可见光图像; Retinex算法; 多尺度分解; 引导滤波

引用格式: GUO Yin, DU Lixia. Infrared and visible image fusion algorithm based on Retinex-enhanced multiscale decomposition. Journal of Measurement Science and Instrumentation, 2024, 15(2): 176-184.