

# Feature-Level Fusion of Infrared Image and Visible Light Image Based on Two-Dimensional Discrete Wavelet Transform and High-Pass Filtering

Zhigang Ren (✉ [16837619@qq.com](mailto:16837619@qq.com))

Peoples Liberation Army Engineering University - Shijiazhuang <https://orcid.org/0000-0001-7265-6496>

Guoquan Ren

Peoples Liberation Army Engineering University - Shijiazhuang

---

## Research Article

**Keywords:** Visible light images, Effective fusion, discrete cosine wavelet transform, infrared images.

**Posted Date:** November 22nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1021012/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Feature-level fusion of infrared image and visible light image based on two-dimensional discrete wavelet transform and high-pass filtering

Zhigang Ren<sup>a</sup>, Guoquan Ren<sup>a, \*</sup>

<sup>a</sup> Department of Vehicle and Electrical Engineering, Shijiazhuang Branch, Army Engineering University of PLA, Shijiazhuang 050003, China

<sup>b</sup> Corresponding author: Guoquan Ren (e-mail: rrrggqqq@163.com)

---

## ABSTRACT

Visible light images and infrared images can effectively complement each other in visual information. Effective fusion of infrared images and visible light images can not only provide better visual effects, but also provide more visual information. In this paper, the discrete cosine wavelet transform is used to obtain the basic information of the image, and the high-pass filter is used to obtain the characteristic information of the image, and the basic information and the characteristic information are effectively fused. In this paper, two-level discrete cosine wavelet transform is performed on the image, which can effectively extract features without distorting the image; secondly, use image quantization to encode, compress and decode the image, reduce the amount of image data, increase the computing speed, and be effective Perform fusion operations on color images. A representative image evaluation method is used to evaluate the fusion image, and the effectiveness of the proposed method is discussed. Experiments have proved that this method is more effective than other methods, and the effect is more obvious in the fusion of color images and infrared images.

---

## 1. Introduction

Any object above absolute zero ( $-273.15\text{ }^{\circ}\text{C} / 0\text{ K}$ ) in nature is constantly emitting infrared rays of different intensities and wavelengths. For this characteristic, we can use infrared sensors for image acquisition. Although the infrared sensor has the characteristics of all-weather and strong environmental adaptability, the application of infrared sensor is limited because of its weak perception of scene brightness and low image resolution. The use of visible light sensor image acquisition of high resolution and strong perception of the brightness of the environment, can effectively make up for the shortcomings of the infrared sensor. Therefore, the subject information and details of the target can be well obtained by using the fusion of visible light and infrared sensors. The fusion methods of visible images (VI) and infrared images (IR) can be divided into two categories: region-based fusion and pixel-based fusion. The region-based fusion method is through image segmentation or extracting the salient regions and regions of interest of the image. at present, the commonly used fusion method is the use of neural network; the pixel-based fusion method is usually easy to implement, generally based on the transform domain, such as image pyramid, wavelet transform and contour transform [1-5]. The key of the region-based fusion method is the local salient region information of the IR, which is easy to ignore the detailed information of the image; the method based on pixel fusion pays attention to extracting the detailed information of the image, and the expression of the salient region information of the image is not obvious. Based on the above information, this paper proposes a pixel fusion method based on wavelet transform to express the basic information of IR and VI. High-pass filtering method is used to extract the details of IR and VI. Considering that the VI contains more visual information data, image quantization compression and decoding are used to remove data redundancy to improve the computing speed.

---

\* Corresponding author. Tel.: +0-000-000-0000; fax: +0-000-000-0000; e-mail: author@university.edu

## 2. Related knowledge

### 2.1 Image Fusion based on MDLatLRR Fusion Framework [6,7]

Fusion image multi-level decomposition (MDLatLRR) is used to decompose the source image to extract the basic information and outstanding details of the image. It is mainly realized through two steps: first, the details of the image are reconstructed through the adaptive fusion strategy and reshaping operator, and the basic information of the image is fused through the average strategy; finally, the image is obtained by fusing the details and basic  $W$  information of the image.

A fusion framework based on image multi-level decomposition (MDLatLRR) is used to fuse infrared and VIs. Firstly, the projection matrix  $L$  is learned by LatLRR,[7] and then the matrix is used to extract the detail and basic information of the input image at several representation levels. MDLatLRR is used to extract multi-level prominent features of the image, and finally the fused image is specially used to deal with the details and basic information of the adaptive fusion strategy reconstruction.

The MDLALRR framework is generic and can be used to provide an effective decomposition method for extracting multi-level features for any number of input images. It can also be applied to other image processing fields with different projection matrices.

### 2.2 Image Fusion based on DWT [9]

Compared with Fourier, wavelet transform (WT) not only has the idea of local short-time Fourier, but also overcomes the disadvantage that the size of local window does not change with frequency. The traditional WT algorithm mainly makes use of the theoretical basis that the human eye is sensitive to the change of local contrast of image or signal, and selects the region of interest in the original image according to certain image fusion rules. and select the most prominent features in the region of interest, and retain these features in the final composite image. For the fusion of infrared grayscale image and VI, the visible details of IR and VI are effectively fused by WT.

However, the subsampling operation of the traditional WT will divide the picture into different levels, resulting in visual distortion of the image. The traditional WT also does not have the characteristics of transfer invariance, so it can not be expressed accurately. Stationary Wavelet Transform (SWT) is used to solve the transfer invariance of WT, and Discrete Cosine Transform (DCT) and local spatial frequency (LSF) are used to express part of the image features to solve the problem of image distortion in order to achieve better fusion results.

## 3 Fusion of visible light image and grayscale IR

The common WT will process the color image in grayscale, in order to reduce the dimension of image information, reduce the matrix dimension of recognition and improve the operation speed. In this paper, based on the principle of image compression coding and decoding, intra-frame data compression technology is used to reduce the amount of VI data, compress a large amount of data, and improve the operation rate to reduce the image fusion time. In order to make the details and feature information of the image more detailed, this paper fuses the image by using Discrete Wavelet Transform (DWT) and High-Pass Filtering (HPF) at the same time.

### 3.1 2-D DWT

Two-dimensional discrete WT (2-D DWT) is proposed based on continuous WT. In order to use the computer to calculate, the continuous function is sampled. The 1-D discrete WT of the new discrete function is represented by formula (1):

$$W_{\psi}(j,k)=\frac{1}{\sqrt{M}}\sum_x f(x)\psi_{j,k}(x) \quad (1)$$

For  $j > j_0$ :

$$f(x)=\frac{1}{\sqrt{M}}\sum_k W_f(j_0,k)f_{j_0,k}(x)+\frac{1}{\sqrt{M}}\sum_{j=j_0}^M\sum_k W_{\psi}(j,k)\psi_{j,k}(x) \quad (2)$$

$x=0,1,2\cdots M-1$

$f(x)$ ,  $f_{j_0,k}(x)$ ,  $f_{j,k}(x)$ ,  $\psi_{j,k}(x)$  is a function of discrete variables.  $M$  is the sequence length.

In order to extend the 1-D discrete cosine transform to 2-D scale, it is necessary to define the scale function and the shift function:

$$\varphi_{j,m,n}(x,y)=2^{j/2}\varphi(2^jx-m,2^jy-n) \quad (3)$$

$$\psi_{j,m,n}^i(x,y)=2^{j/2}\psi^i(2^jx-m,2^jy-n) \quad (4)$$

$i=\{H,V,D\}$

$\varphi(x,y)$  is a 2-D scale function,  $\psi^H(x,y)$ ,  $\psi^V(x,y)$ ,  $\psi^D(x,y)$  are three 2-D wavelet metric functions:

$$\varphi(x,y)=\varphi(x)\varphi(y) \quad (5)$$

$$\psi^H(x,y)=\psi(x)\varphi(y) \quad (6)$$

$$\psi^V(x,y)=\varphi(x)\psi(y) \quad (7)$$

$$\psi^D(x,y)=\psi(x)\psi(y) \quad (8)$$

The three 2-D measurement functions vary with the intensity and grayscale of the image in three directions,  $\psi^H$  changes horizontally,  $\psi^V$  changes vertically, and  $\psi^D$  changes diagonally.

The DWT formula of  $M \times N$  image  $f(x,y)$  is:

$$W_{\varphi}(j_0,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\varphi_{j_0,m,n}(x,y) \quad (9)$$

$$W_{\psi}^i(j,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\psi_{j,m,n}^i(x,y) \quad (10)$$

$i=\{H,V,D\}$

In the same way as  $j \geq j_0$ ,  $f(x,y)$  is obtained by inverse transform of DWT:

$$f(x,y) = \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(j_0, m, n) \varphi_{j_0, m, n}(x, y) + \frac{1}{\sqrt{MN}} \sum_{i=H, V, D} \sum_{j=j_0}^{\infty} \sum_m \sum_n W_\psi^i(j_0, m, n) \varphi_{j, m, n}^i(x, y) \quad (11)$$

Like a DWT, using separable 2-D scales and wavelet functions, the rows of 2-D functions are decomposed by 1-D DWT, and then the columns are decomposed by 1-D DWT [12]. After decomposition, the low-frequency components LL, HL and high-frequency components LH and HH of the original image in horizontal and vertical directions can be obtained. In this paper, the image is decomposed by two-level WT, that is, the low-frequency component LL is further transformed by 2-D DWT, which can effectively quantify the image to meet the requirements of feature extraction. The specific decomposition process is shown in the Fig.1.

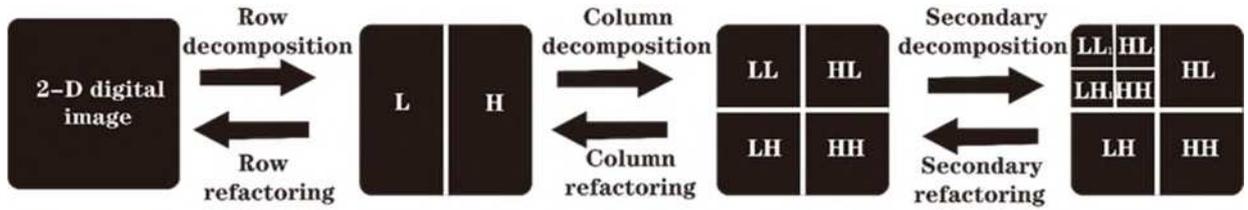


Fig.1. 2-D discrete wavelet transform

### 3.2 Image coding and compression

Redundant data is often produced in the process of image processing. generally, in order to improve the transmission efficiency of the image and the operation speed of the image, the redundant data is often deleted in the process of image processing. deleting these redundant data can effectively reduce the image data without destroying the effective information of the image [13]. This paper mainly deletes the visual redundancy and spatial redundancy of the color image without destroying the details and feature information of the image. Image coding can gradually reduce the amount of data, improve the transmission rate and accelerate the calculation rate. In practical application, it is also necessary to decode it and restore it to 2-D data image form. In order to ensure the fidelity of the decoded image relative to the original image, that is, the image acquired after image compression and decoding has no information loss, it is necessary to propose an objective fidelity criterion.

#### 1 Root Mean Square Error

Root Mean Square Error (RMSE) is a commonly used fidelity criterion. The RMSE between the original image  $f(x,y)$  and the compressed and decoded image  $\hat{f}(x,y)$  is:

$$e_{rms} = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]^2} \quad (12)$$

The size of the image is  $M \times N$ .

## 2 Mean square signal-to-noise

Because there may be noise in the process of image compression and decoding, it is proposed  $\hat{f}(x,y)=f(x,y)+n(x,y)$  that  $n(x,y)$  is the noise signal, and the root mean square signal-to-noise ratio of the output image is:

$$\text{SNR}_{\text{rms}} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x,y)^2}{\sqrt{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y)-f(x,y)]^2}} \quad (13)$$

The size of the image is  $M \times N$ .

### 3.3 High-Pass Filtering method

The HPF method uses the High-Pass filter to suppress the low-frequency part of the image, and the ideal high-pass filter filters the low-frequency image signal by setting the threshold. In this paper, the high-frequency information in the quantized image is extracted by the high-pass filter, and the high-frequency information of the IR and the VI is superimposed, and the mean value of the low-frequency information in the quantized image is processed to reduce the influence of the fused image obtained by the HPF method on the details of the final fused image.

### 3.4 Fusion steps and parameter settings

Step 1, the image editor is used to scale the IR and VI to the same scale.

Step 2, the two images are decomposed respectively by 2-D DWT, and the LL subband of the decomposed image is decomposed by second-order WT.

Step 3, the decomposed images are quantized respectively, and the image coding technology of discrete cosine transform is used to compress them, so as to remove the information ignored by human visual system and reduce the amount of image data.

Step 4, HPF is used to filter the compressed data, low-frequency data is averaged, and high-frequency data is retained to obtain feature image A;

Step 5, the compressed image data are respectively de-quantized, and the look-up table method is used to decode them. The de-quantized data are obtained by discrete cosine transform to form the decoded image.

Step 6, the inverse WT is performed on the decoded image to obtain the fusion image, which is used to express the fusion detail information.

Step 7, feature image A and detail image B are fused.

## 4 Evaluation and experimental results

### 4.1 Evaluation system

To judge whether the method proposed in this paper is effective or not, the images fused by the above three methods are comprehensively evaluated and analyzed. The traditional comprehensive evaluation analysis can be divided into two categories: objective evaluation method and subjective evaluation method. The subjective evaluation method evaluates the image quality through the subjective evaluation of human eyes, but it is easily affected by environment, emotion and so on, which leads to the shortcomings of unintuitive evaluation results and poor stability. The objective evaluation method can quantitatively evaluate the image quality based on a certain mathematical model with strong stability and simple operation. In this paper, the objective evaluation method is used to evaluate the quality of the fused image, and whether the proposed method is effective or not is discussed. The evaluation systems used in this paper are: Image Gradient (IG); Peak Signal-to-noise Ratio (PSNR); Structural Similarity (SSIM); Mutual Information (MI); Information Entropy (EN). [2, 16-21]

### 1 Image Gradient

IG makes use of the content with high sensitivity of the human eye, which well represents the structure information of the fused image. For the 2-D grayscale image [22], the horizontal gradient and vertical gradient can be decomposed and calculated first:

$$\nabla f_H = f(m+1, n) - f(m, n) \quad (14)$$

$$\nabla f_V = f(m, n+1) - f(m, n) \quad (15)$$

Then take the sum of the two as the point gradient:

$$\nabla f = |\nabla f_H| + |\nabla f_V| \quad (16)$$

The point gradient value can obtain the edge information of the fused image features, and the larger the gradient value is, the better the algorithm is for image feature edge extraction.

### 2 Peak Signal-to-noise Ratio

PSNR is measured in decibels, which is the ratio of the energy of the peak signal to the average energy of the noise:

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \hat{f}(x, y)]^2 \quad (17)$$

$$PSNR = 10 \log_{10} \left( \frac{2^{\text{bits}} - 1}{MSE} \right) \quad (18)$$

MSE is the mean square error of the image, and bits is the number of bits per sampling value, bits=8 in this paper. The larger the PSNR, the higher the fidelity of the fused image and the better the image quality [23].

### 3 Structural Similarity

SSIM is an index to measure the similarity of two images. At the same time, if the two images are before and after compression, SSIM can evaluate the image quality of the compressed image. SSIM is mainly for brightness, contrast and image structure comparison [24, 25]. The average gray level of the image is used as the estimation of luminance measurement:

$$\mu_x = \frac{1}{H \times M} \sum_{i=1}^H \sum_{j=1}^M X(i, j) \quad (19)$$

Luminance contrast function  $l(x, y)$ :

$$l(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1} \quad (20)$$

Contrast function  $c(x,y)$ :

$$c(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2} \quad (21)$$

Structural contrast function  $s(x,y)$ :

$$s(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \quad (22)$$

$$SSIM(x,y)=f[l(x,y),c(x,y),s(x,y)]=[l(x,y)]^\alpha [c(x,y)]^\beta [s(x,y)]^\gamma \quad (23)$$

Where  $\mu_x$   $\mu_y$  is the average grayscale of the two images,  $\sigma_x$   $\sigma_y$  is the square check of the two images, and  $\sigma_{xy}$  variance represents the covariance of the image.  $C_1$   $C_2$   $C_3$  is constant, in order to avoid the case that the denominator is 0. Weights  $> 0$ , which are used to adjust the weight of each module. When  $\alpha$  ,  $\beta$  ,  $\gamma=1$ , and  $C_3=C_2/2$  to get:

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \quad (24)$$

The value range of the function is [0, 1]. The closer the value of the function is to 1, the higher the similarity between the two images.

#### 4 Mutual Information

For how to calculate the amount of information transmitted before and after image fusion, the key lies in the problem of how much information is obtained from the fused image, and the main problem is to solve the problem of information transmission. The amount of received information depends on how much uncertainty is eliminated, which is the amount of MI [21], defined as:

$$I(x,y)=I(x)-I(x|y) \quad (25)$$

$I(x,y)$  indicates the amount of information obtained after transmission,  $I(x)$  represents the amount of information before transmission, and  $I(x|y)$  indicates the amount of information lost during transmission.

#### 5 Information Entropy

EN is a common index to evaluate image quality. generally speaking, the greater the image information entropy, the more information the image contains, the better the image quality. However, if there is a lot of noise in the image, the information entropy will increase significantly, so it is impossible to determine the quality of the image. Formula of image information entropy [21]:

$$H=-\sum_{i=0}^{n-1} p_i \log_2 p_i \quad (26)$$

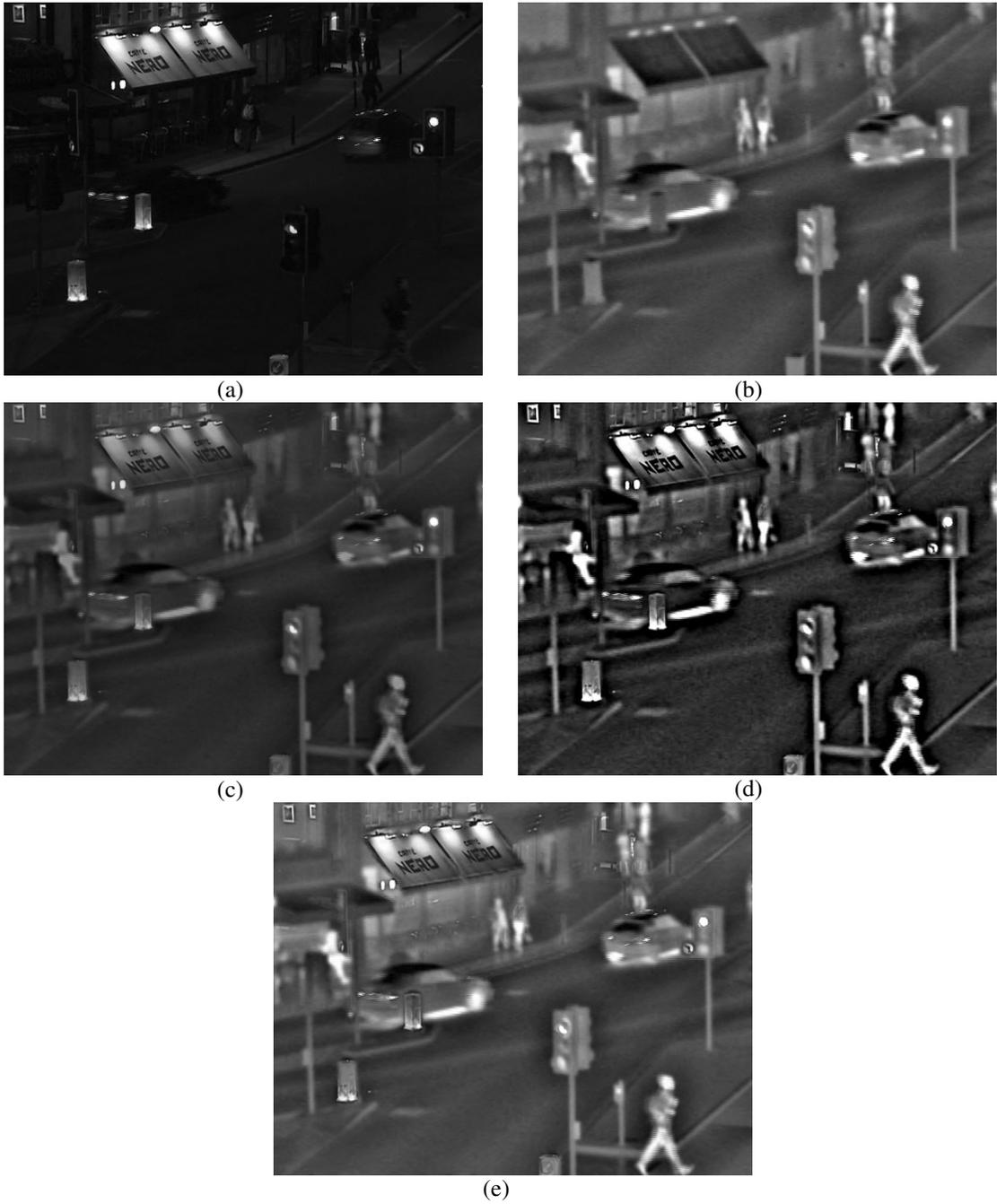
$p_I$  refers to the probability that the grayscale value  $I$  appears in the image, and  $I$  refers to the grayscale of the image, which generally takes a value of 255.

#### *4.2 Evaluation index system and experimental results*

The above three different methods are used to evaluate the quality of the fused image to verify the effectiveness of the proposed method.

The first group of infrared and VIs are shown in fig.2 (a) and (b) with a resolution of  $632 \times 496$ . This image is a commonly used test image. The images obtained by different fusion methods can be seen that the effect of the proposed method is more obvious than the other two, and there is no ghosting in feature extraction. Obviously, the proposed method is effective and superior to the other two methods, and has a better feature extraction for the fused image.

The image quality of fused image fig.2 (c), (d) and (e) is evaluated. Table 1 shows the specific values of the proposed image quality evaluation index for different images. It can be seen from the table that the gradient information of the image obtained by the proposed fusion method is richer, the structure similarity with the original image is higher, and the information entropy is also larger. There is no obvious noise on the image, which indicates that the image quality is better.



**Fig.2.** Different methods fuse VI and IR for the first group. (a) VI. (b)IR. (c) MDLatLRR. (d)DWT. (e) Proposed method.

**Table 1.** Fusion quality indexes with different methods for the first pair of IR and VI.

	IG	PSNR	SSIM	MI	EN
MDLatLRR	0.0160	36.7662	0.5629	6.6581	6.0184
DWT	0.0204	28.6539	0.3777	5.9944	6.7327
Proposed	0.0255	34.0339	0.6368	12.5604	6.7713



**Fig.3.** Different methods fuse VI and IR for the second group. (a) VI. (b) IR. (c) MDLatLRR. (d) DWT. (e) Proposed method.

**Table 2.** Fusion quality indexes with different methods for the second pair of IR and VI.

	IG	PSNR	SSIM	MI	EN
MDLatLRR	0.0173	36.7262	0.6713	4.2850	6.9649
DWT	0.0154	39.7223	0.7465	8.1773	6.9416
Proposed	0.0238	32.1922	0.7056	6.4231	6.9785

The second group of IRs and light images are common infrared test images as shown in fig. 3. The picture environment is a city road with light source, and the image resolution is  $632 \times 496$ . MDLatLRR (c), DWT (d)proposed (e) in are used respectively. Compared with the other two fusion methods of DWT, more information is obtained. There is no obvious difference between the MDLatLRR and the proposed under the condition of subjective analysis, so the objective evaluation method is used.

Table 2 shows the five indicators of the fused images of each group, which directly reflects the image quality of the images obtained by different methods. The data in the punching table can be seen intuitively that the image quality of the proposed method is better than that of the cyclic neural network. at the same time, the fusion image obtained by the proposed method has greater structural similarity and interactive information with the IR (a). Although the peak signal-to-noise ratio and structural similarity of the DWT are higher, the information entropy of the image obtained by the DWT is less. At the same time, the detail information of the obtained image is less, and the artifact problem of the image is more serious. To sum up, the detailed information of the image obtained by the proposed method is rich.

The third group of IRs and VIs, the main body of the image is street lights and pedestrian rural roads, the image environment is mostly trees, there is more obvious infrared radiation, is also one of the common test images, the image resolution is  $632 \times 464$ . In fig.3, it can be seen that the fusion image using DWT has poor processing of detail information, and the method proposed in this paper is rich in extracting the details of the image.

It can be seen from Table 3 that the SSIM and MI of the fusion image (e) obtained by the proposed method are significantly higher than those of the other two methods. The IG is also higher than that of the other two methods, indicating that the overall structure of the fused image (e) is better, but due to the existence of some noise in the original IR and VR, the two parameters of PSNR and EN have some limitations in the evaluation of image quality.



**Fig.4.** Different methods fuse VI and IR for the third group. (a) VI. (b) IR. (c) MDLatLRR. (d) DWT. (e) Proposed method.

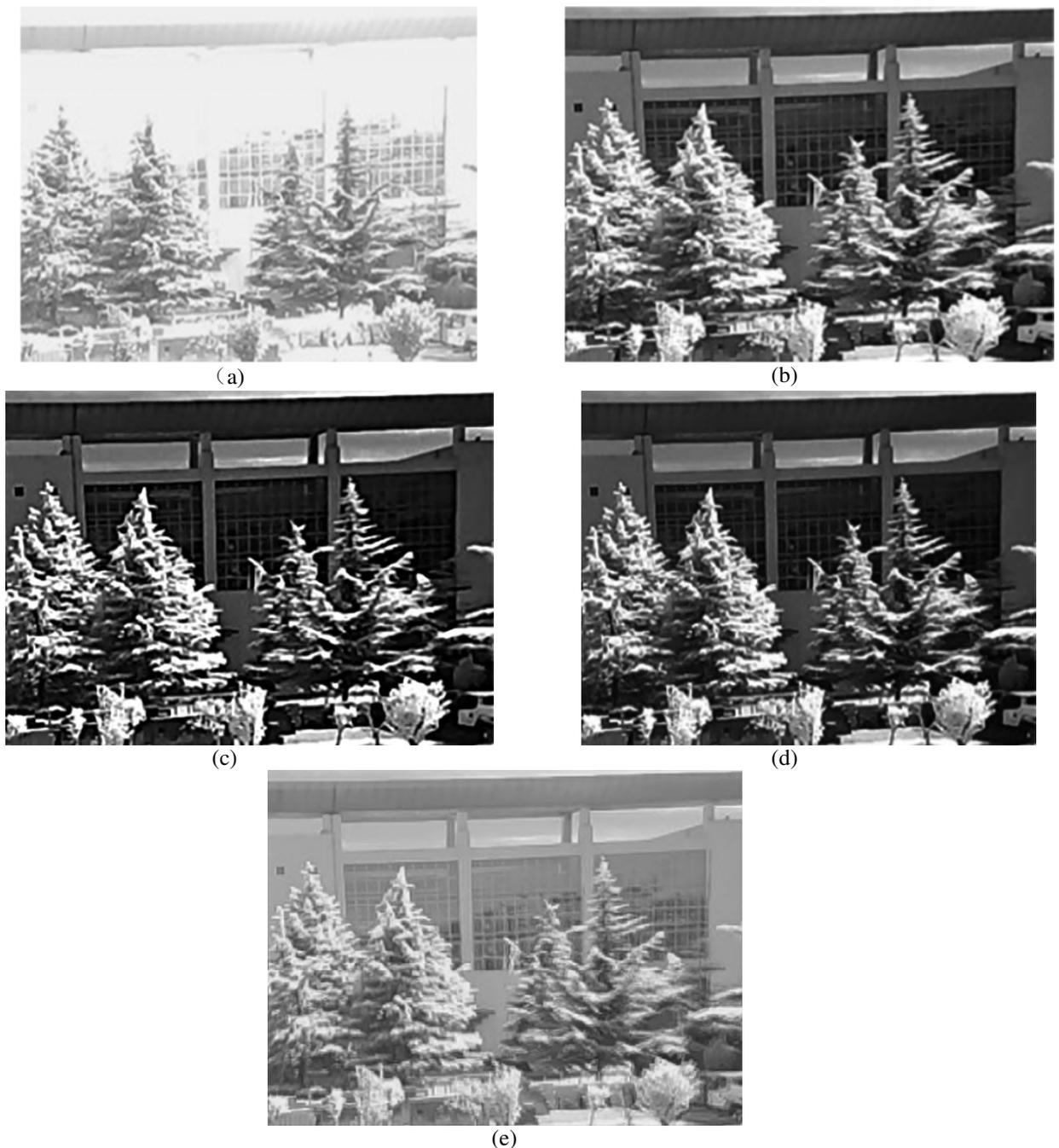
**Table 3.** Fusion quality indexes with different methods for the third pair of IR and VI

	IG	PSNR	SSIM	MI	EN
MDLatLRR	0.0258	33.6164	0.7652	13.0446	6.2593
DWT	0.0226	35.4293	0.6953	15.9687	6.1788
Proposed	0.0302	30.8597	0.8223	21.1437	6.1588

The fourth group of IR and VI are shown in fig. 5 (a) (b). The resolution is  $770 \times 562$ . The picture is an office building in the sun and has more detail information. Due to the exposure problem, the VI lost a lot of information. Three groups of images (c) (d) (e) were

obtained after the IR and VI was fused. By comparing the other two groups of images, it can be seen that the fusion image obtained by the proposed method has more detailed information. It has good ability of feature extraction.

Table 4 lists the fusion quality coefficient of the fourth group of fused images, and there is some noise in the fused image. PSNR and EN have some limitations in the evaluation of image quality, and SSIM and MI are obviously higher than the other two methods, which shows that the proposed method eliminates more uncertain information, improves the amount of image information, and has less image distortion and higher similarity with the original image. The results show that the quality of the fused image obtained by the proposed method is high.



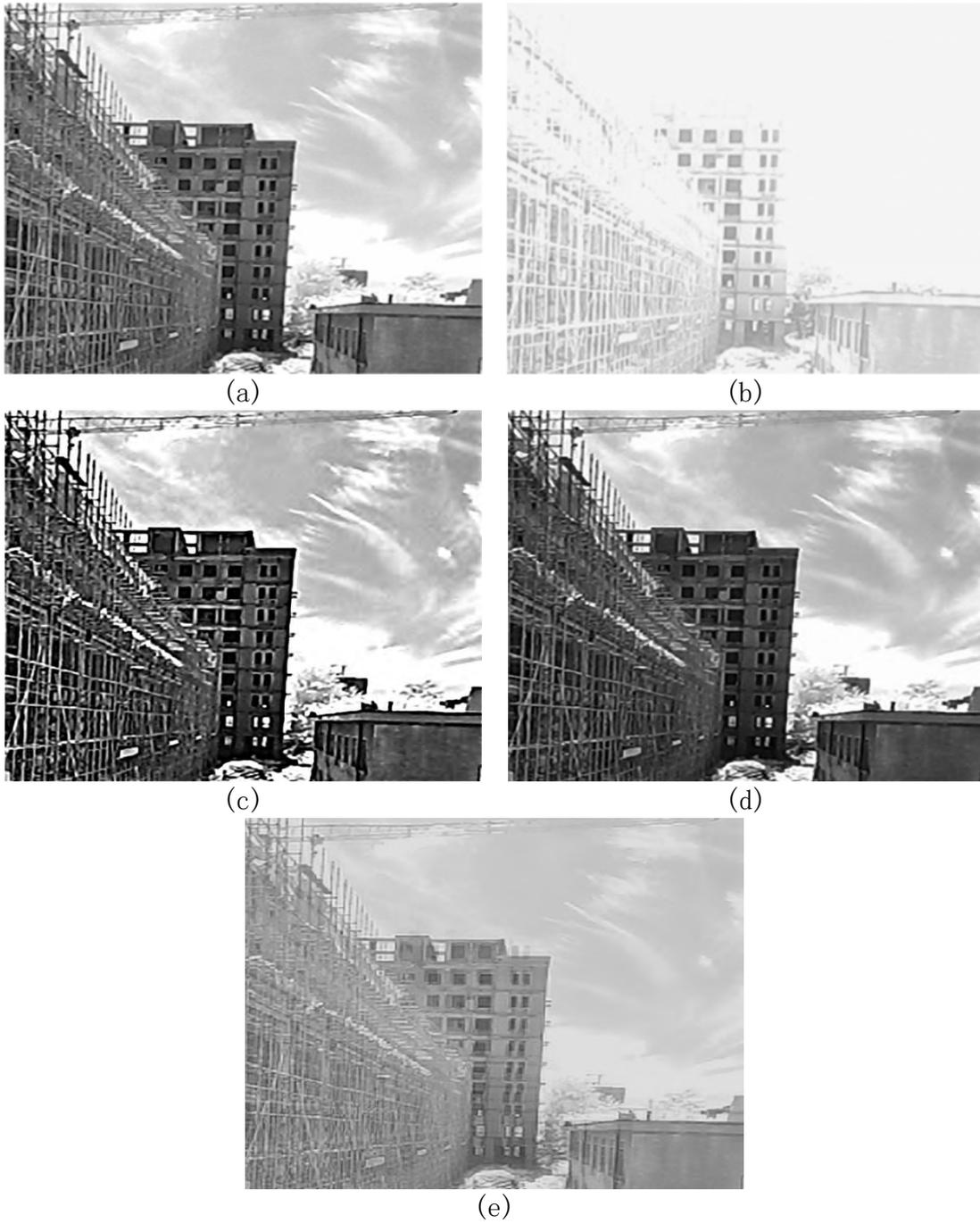
**Fig.5.** Different methods fuse VI and IR for the fourth group. (a) VI. (b)IR. (c) MDLatLRR. (d)DWT. (e) Proposed method

**Table 4. Fusion quality indexes with different methods for the fourth pair of IR and VI.**

	IG	PSNR	SSIM	MI	EN
MDLatLRR	0.0577	35.9456	0.1624	6.2540	7.2713
DWT	0.0428	40.9239	0.2013	6.5969	7.5157
Proposed	0.0236	36.3095	0.4021	7.0331	5.5705

The fifth group of IR and VI is shown in fig. 6 (a) (b) with a resolution of  $740 \times 580$ . This is a construction site scene under the scorching sun, including part of the sky. The images obtained by different fusion methods are shown in figure 6 (c) (d) (e). See the distant details of the features, such as trees, houses, etc., have a better expression, other ways of performance is poor. The results show that this method can effectively extract the detail features of the image.

Table 5 lists the image quality indicators of different fusion methods for the fifth group of images. The PSNR and EN of the images obtained by DWT method are higher than those of the proposed method. However, for MI, the proposed method is higher than the other two fusion methods, indicating that more information is obtained from the original image, but the image quality is not as good as DWT fusion. For the fifth group of images, the quality of the fused image obtained by DWT method is better, and the image obtained by the proposed method can extract more information from the original image.



**Fig.6.** Different methods fuse VI and IR for the fifth group. (a) VI. (b)IR. (c) MDLatLRR. (d)DWT. (e) Proposed method

**Table 5.** Fusion quality indexes with different methods for the fifth pair of IR and VI.

	IG	PSNR	SSIM	MI	EN
MDLatLRR	0.0688	29.6521	0.7622	1.7102	7.2422
DWT	0.0386	37.5330	0.9422	4.2602	7.6753
Proposed	0.0293	35.9981	0.8182	16.8755	6.7747

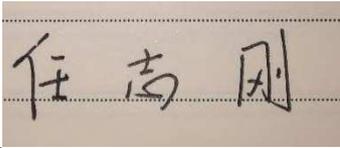
## 5. Conclusion

In this paper, the image is decomposed by 2-D DWT, the basic information of the image is extracted, and then the detail features of the image are extracted by HPF, and the basic information and detail features are fused. The fused image has more information of the original image, and the basic information of the original image can be extracted effectively at the same time. At present, the algorithm can extract the detail features of the fused image, but it cannot solve the problem of IR and VI registration, so it can further register the IR and VI, which can make image feature extraction more accurate and eliminate artifacts and other problems. The experimental results show that the proposed method can effectively extract the basic information and detail information of the original image, while maintaining high structural similarity and information interaction, which shows that the fusion image has less distortion and has a better ability of detail feature extraction compared with the MDLatLRR of region fusion and the DWT of pixel fusion. The amount of information of the image is improved.

### Conflict of interest statement

As corresponding author I, Zhigang Ren, hereby confirm on behalf of all authors that:

1. This manuscript, or a large part of it, has not been published, was not, and is not being submitted to any other journal.
2. If presented at or submitted to or published at a conference(s), the conference(s) is (are) identified and substantial justification for re-publication is presented below. A copy of conference paper(s) is(are) uploaded with the manuscript.
3. If the manuscript appears as a preprint anywhere on the web, e.g. arXiv, etc., it is identified below. The preprint should include a statement that the paper is under consideration at Signal Processing Letters.
4. All text and graphics, except for those marked with sources, are original works of the authors, and all necessary permissions for publication were secured prior to submission of the manuscript.



Signature \_\_\_\_\_ Date 2021.10.21

All authors each made a significant contribution to the research reported and have read and approved the submitted manuscript.

## References:

- [1] Yang Guang et al., infrared and visible image fusion based on multi-features. *Optical Precision Engineering*, 2014. 22 (02): pages 489-496.
- [2] Zhou Yu people, Research on Infrared and visible Image Fusion algorithms, 2014, Graduate School of Chinese Academy of Sciences (Changchun Institute of Optics and Precision Machinery and Physics). Page 111.
- [3] Li Quanzhong, Research on Multi-source Image Fusion and Evaluation based on artificial Neuron Perceptual Model, 2014, University of Electronic Science and Technology. Page 85.
- [4] Zhang Baohui, Infrared and visible Image Fusion system and Application Research, 2013, Nanjing University of Science and Technology. Page 117.
- [5] Yang Yang, Image Fusion algorithm Research based on Multi-scale Analysis, 2013, Graduate School of Chinese Academy of Sciences (Changchun Institute of Optical Precision Machinery and Physics). Page 124.
- [6] Lu Yongfang, Lu Ke and Hou Xiangwen, Research on Image Fusion method based on IHS transform. *Science and Technology Bulletin*, 2012. 28 (06): page 212-214.
- [7] Li Mei et al. Infrared and visible image fusion method based on NSCT and PCNN. *Optoelectronic Engineering*, 2010. 37 (06): page 9095.
- [8] Li, H., X. Wu and J. Kittler, MDLatLRR: A Novel Decomposition Method for Infrared and Visible Image Fusion. *IEEE Transactions on Image Processing*, 2020. 29: P. 4733-4746.
- [9] Jin, X., et al., Infrared and visual image fusion method based on discrete cosine transform and local spatial frequency in discrete stationary wavelet transform domain. *Infrared Physics & Technology*, 2018. 88: P. 1-12.
- [10] Infrared and visible image fusion method based on Yin Xiang and Ma Jun, wavelet packet and non-subsampled contour wave. *Laser Magazine*, 2018. 39 (01): pp. 123-127.
- [11] Zhao Sihui, IR and visible Image Fusion based on Multi-Wavelet transform, 2015, Shenyang University of Technology. Page 61.
- [12] Anfu, etc., infrared polarization and light intensity image fusion based on DWT. *Application of Optoelectronic Technology*, 2013. 28 (02): pages 18-22.
- [13] Song Yanfeng, Shao Xiaopeng and Xu Jun, infrared image enhancement algorithm based on dual-platform histogram. *Infrared and Laser Engineering*, 2008 (02): 308-311.
- [14] Zhang Qiang, Zhou Haiyin and Wang Jionqí, wavelet transform image fusion based on local variance and high-pass filtering. *Computer Simulation*, 2008 (08): 223-226pp.
- [15] Wei Jun and li Bicheng, remote sensing image fusion based on IHS transform, wavelet transform and high-pass filtering. *Journal of University of Information Engineering*, 2003 (02): 46-50.
- [16] Yang Yanchun, Li Jiao and Wang Yangping, review of image fusion quality evaluation methods. *Computer Science and Exploration*, 2018. 12 (07): pages 1021-1035.
- [17] Wu Jinjian, Image Information perception and Image quality Evaluation based on Human Vision system, 2014, Xi'an University of Electronic Science and Technology. Page 163.
- [18] Wang Yuqing and Wang Suojian, quality evaluation of infrared and visible light fusion images. *China Optics*, 2014. 7 (03): pages 396-401.
- [19] Zhang Xiaoli, Li Xiongfei and Li Jun, correlation analysis and performance evaluation of fusion image quality evaluation index. *Journal of Automation*, 2014. 40 (02): pp. 306-315.
- [20] Zhang Yong and Jin Weiqi, objective evaluation method of night vision fusion image quality. *Infrared and Laser Engineering*, 2013. 42 (05): 13601365.
- [21] Zhang Xiaofeng, quality Evaluation of visible and infrared image Fusion, 2007, Huazhong University of Science and Technology. Page 61.
- [22] Wang Yuqing, objective evaluation method of image quality based on gradient complex matrix. *Computer Technology and Development*, 2013. 23 (01): pp. 63-66.
- [23] Xiao Hui, Research on Infrared and visible Image Fusion algorithm based on Wavelet transform, 2009, Changchun University of Technology. Page 53.
- [24] Yuan Jinsha et al., research status and prospect of infrared and visible image registration. *Laser and infrared*, 2009. 39 (07): pp. 693-699.
- [25] Wang Kunpeng, Xu Yidan and Yu Qifeng, classification and status of infrared and visible image registration methods. *Infrared technology*, 2009. 31 (05): pp. 270-274.