

Image Fusion Using Non-Subsampled Contourlet Transform Based on Activity and Local Gradient Energy Methods

Chunhua Guo, Dianyuan Han*

Communication and Media College, Weifang University, Weifang 261061, China

E-mail: wfhdyy@163.com

*Corresponding author:

Keywords: fusion, activity function, gradient energy, objective evaluation

Received: December 20, 2024

It is often difficult for a single image to obtain all the details of the same scene. To handle this problem, multiple images can be acquired through a variety of ways, and then the obtained images can be typically combined into one image by image fusion technology. For improving image fusion quality, a new image fusion method based on Non-Subsampled Contourlet Transform (NSCT) is proposed. Source images are initially decomposed via NSCT, the low frequency sub-band image and a series of high frequency sub-band images with different directions and different scales are obtained, low frequency sub-band image fusion is carried out based on the activity with local sharpness changes. In the high frequency region, fusion is carried out based on local gradient energy with edge strength, and finally the fusion image is reconstructed by NSCT inverse transform. Through the fusion experiments of multi-focus image and multimodal images, the proposed method is compared with the other methods such as NSCT_PC, NSCT_EN_PCNN, NSST_PCNN, CWT_SR and JBF. Visually, the fusion image obtained by the proposed method is clearer in detail and has a stronger sense of image hierarchy. Using objective evaluation such as peak signal to noise ratio, structural similarity, edge information retention, and information entropy, they are improved by at least 0.4%, 0.02%, 9.7%, and 1.4% respectively. Generally speaking, the method in this paper retains more important details and shows better fusion performance.

Povzetek: Članek predstavi metodo zlivanja slik na osnovi NSCT. Nizkofrekvenčne koeficiente združi z aktivnostjo in lokalnimi spremembami ostrine, visokofrekvenčne pa z lokalno gradientno energijo in močjo robov. Poskusi z multifokusnimi ter multimodalnimi slikami pokažejo boljšo jasnost, ohranitev robov, kontrast in entropijo.

1 Introduction

Different image sensors have redundancy and complementarity in space and time when imaging the same scene or imaging the same sensor in different ways [1]. In image processing, an important problem is how to optimally combine information into an image to maximize accurate and comprehensive scene description, image fusion is to solve this problem. Image fusion makes use of the redundancy and complementarity of single image information, and adopts certain fusion criteria to describe the new fusion image more accurately and more comprehensively [2]. At present, image fusion is widely used in medical imaging, remote sensing, computer vision and many other fields [3-6]. For example, in target recognition, the fusion of panchromatic image and multispectral image makes new image have higher spatial resolution and higher frequency resolution simultaneously. In medical diagnosis, the fusion of Computed Tomography (CT) and Magnetic Resonance Image (MRI) allows new image to show bone and soft tissue clearly. In remote sensing, the fusion of synthetic aperture radar image and millimeter wave radar image can

have the advantages of strong anti-jamming ability and high resolution at the same time.

According to different fusion ideas, the current dominant image fusion methods are mainly divided into two categories: fusion method based on spatial domain and fusion method based on transformation domain. The fusion method based on spatial domain is to fuse the pixel or color information directly, this method is unable to extract the clear part and edge part of image accurately, which is likely to lose the characteristic such as contour and detail of the source image and reduce the contrast of the fused image [7]. Multiscale and multiresolution analysis methods are commonly used in fusion methods based on transformation domain. Typical multiscale and multiresolution analysis methods include pyramid method [8], wavelet transform method [9] and contourlet transform method [10] et al. Because of its good time-frequency analysis characteristics, anisotropy and relative independence on different scales, wavelet transform can obtain better fusion effect than pyramid method, and thus becomes a common multiscale analysis method in the field of image fusion [11]. However, because wavelet transform has no translation invariance, which is likely to cause significant fluctuations or drift in the feature ex-

traction coefficients, and the directionality is also limited., it cannot effectively describe the contour and directional texture features of the image [12]. Do and Vetterli introduced the Contourlet transform in 2005 to address the limitations of wavelet transform, such as frequency aliasing and limited directionality [13], which has good time-frequency local characteristics, anisotropy, flexible multiscale and multi-directional decomposition of images. But it has the disadvantage of frequency aliasing and no translation invariance. Cunha and Zhou et al. put forward a Non-Subsampled Contourlet Transform (NSCT) in 2006 [10], which takes advantage of non-subsampled pyramid filter decomposition and non-subsampled directional filter Banks. This algorithm has been widely used in many fields.

In the image fusion process based on multiscale and multiresolution analysis, fusion rule selection is another significant factor affecting the fusion effect [14]. For the low frequency coefficients, weighted average is most commonly used, but this method is easily lose some useful information in the low frequency of image, which brings down the contrast of fusion image to a certain extent. For the high frequency coefficients, adopting the coefficients of maximum absolute value or maximum energy is the simplest, but it ignores the correlation between the coefficients, which is likely to cause the false selection of fusion coefficients [15].

To address these challenges, this paper proposes an approach based on NSCT for more effective feature extraction. This method adopts an activity function for the low frequency region and combines it with the local sharpness variation to form a low frequency measurement operator, thereby enhancing the expression of contour information in the fused image. For the high frequency region, a detail enhancement operator based on local gradient energy and combined with edge strength is proposed to further improve the extraction and utilization of edge texture features.

2 Non-Subsampled contourlet transform

NSCT has the characteristics of multi-scale, multi-direction, anisotropy, translation invariance and so on, which is a super perfect multi-scale transform method. It is composed of the Non-Subsampled Pyramid Filter Banks (NSPFB) and Non-Subsampled Directional Filter Banks (NSDFB), and using à trous algorithm [16] to achieve a very flexible multi-scale transform.

The structure diagram and frequency decomposition diagram of NSCT are shown in Fig. 1. First, NSPFB performs multi-scale and multi-resolution transform, and then NSDFB performs multi-direction transform. In the decomposition of NSCT, the decomposition filter used in each level is composed of the up-sampling decomposition filter used in the previous level of decomposition. In the NSPFB transform of image, the lowpass subband image generated by each level of decomposition is obtained by lowpass filtering the low-pass subband image of the previous level after up-sampling, while the band-pass subband image generated by each level of decomposition is obtained by high-pass filtering the low-pass subband image of the previous level after up-sampling. In the process of NSDFB directional decomposition of images, the two-dimensional frequency domain plane can be divided into several directional wedge-shaped blocks after each level filter undergoes up-sampling, then NSDFB can achieve accurate multi-direction decomposition in the frequency domain. Similarly, in the reconstruction of NSCT, each level of synthetic filter is obtained by up-sampling the synthetic filter used in the previous level of reconstruction. It is because in the decomposition and reconstruction process, no up-sampling and down-sampling operations are performed on the signal, but up-sampling operations are performed on the corresponding decomposition filter and synthesis filter, which makes NSCT have translation invariance.

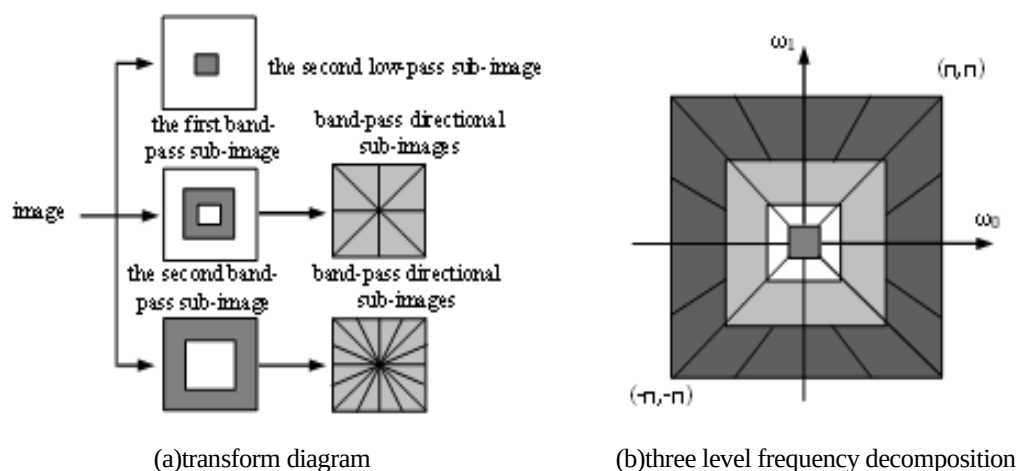


Figure 1: a and b Transform diagram and frequency composition diagram of NSCT

Image after NSCT decomposition, the size of all decomposed images is the same as that of the source image. The multi-scale and multi-direction decomposition at each level ensures the characteristics of anisotropy and

enables the NSCT to represent the image sparsely. Obvious features such as curves, edges and contours are represented as the magnitudes of sub-image coefficients. Especially the detailed features such as edges and textures

are manifested as a few frequency coefficients with larger values in the NSCT domain. In other words, frequency coefficients with larger values contain more information such as edges and textures. Compared with wavelet transform, NSCT has better sparse representation ability and more concentrated energy. Therefore, using NSCT in image fusion can better extract features of each band and obtain more information.

3 The fusion method

Fusion rule is another important issue to determine the fusion effect. A new image fusion method based on NSCT is put forward in this paper.

3.1 Fusion rule of low frequency sub-image

The change of local information not only reflects the difference of each pixel in the image, but also describes the significance degree of the image relative to the background information from the side. Therefore, aiming at the low frequency subband decomposed by NSCT, this paper proposes a new definition of local change information -- activity function (AF), taking the local energy weighting of coefficients as the activity degree not only reflects the spatial correlation between pixels, but also considers the significance degree of coefficients, its expression is as follows:

$$AF(x, y) = \sum_{(x', y') \in \Omega_1} \omega(x', y') |C(x + x', y + y')|^2 \quad (1)$$

Where, $AF(x, y)$ represents the activity of low frequency coefficient at the position (x, y) . $C(x, y)$ is the low frequency subband coefficient of source image after decomposition by NSCT. Ω_1 defines the window range, which is a sliding window with a value of 3×3 [17], $\omega(x', y')$ represents the weight of the window and is the contrast sensitivity function, here it is taken as the Michelson contrast and defined as $\omega(x', y') = (C_{\max} - C_{\min}) / (C_{\max} + C_{\min})$, $C_{\max}$ and $C_{\min}$

are respectively the maximum value and minimum value of subband coefficients in the corresponding window.

The low frequency subband contains the variations of pixel brightness and gray level. To further enhance the information of the low frequency subband and reflect the local contrast changes, the Local Sharpness Change (LSC) is introduced. The local contrast variation of image is reflected by calculating the neighborhood sharpness change (SC) of (x, y) . LSC is defined as equation 2.

$$LSC(x, y) = \sum_{m=-M}^M \sum_{n=-N}^N (C(x, y) - C(x_1, y_1))^2 \quad (2)$$

In which, the values of M and N are 3 and 3 respectively [18].

Thus, the low frequency subbands are fused through the low frequency measurement operator (LMO), that is equation 3.

$$LMO(x, y) = (AF(x, y))^{\alpha_1} \cdot (LSC(x, y))^{\beta_1} \quad (3)$$

In which, the parameters α_1 and β_1 are respectively used to adjust the weights of AF and LSC sizes in LMO .

In conclusion, the fusion rule of the low frequency subband is as equation 4.

$$LF(x, y) = \begin{cases} C_A & \text{if } [R_A(x, y)] > \frac{\tilde{M} \cdot \tilde{N}}{2} \\ C_B & \text{otherwise} \end{cases} \quad (4)$$

In which,

$$R_A(x, y) = \{(x_0, y_0) \in \Omega_2 \mid LMO_A(x_0, y_0) \geq LMO_B(x_0, y_0)\}$$

From the above, the fused low frequency subband $LF(x, y)$ is obtained. C_A and C_B represent the low frequency subband of image to be fused respectively, and Ω_2 represents the sliding window centered on (x, y) with a size of $\tilde{M} \times \tilde{N}$. The values of $\tilde{M}$ and $\tilde{N}$ are 7 and 7 respectively [19].

3.2 Fusion rule of high frequency sub-image

The high frequency sub-image decomposed by NSCT mainly includes the detail information such as the contour structure and edge of image. This information is usually presented as the coefficients with large absolute value or large modulus, corresponding to the significant features in a certain direction interval, and they can well describe the structure information of image. The quality of extracting high frequency subband information directly affects the image fusion effect. To highlight the texture information of the high frequency subbands, local gradient energy (LGE) is introduced. The greater the local gradient energy, the more detailed information the image contains. By calculating the gray level changes at the position (x, y) of high frequency subband, the amount of its detailed information is reflected, that is as equation 5.

$$LGE(x, y) = \sum_{m, n \in S} |G(x + m, y + n)|^2 \quad (5)$$

Where, S is the window size and its value is 3 [19].

$G(x, y)$ is the gradient of pixel (x, y) , $LGE(x, y)$ represents the local gradient energy of the high frequency subband at the position (x, y) .

$G(x, y)$ is the first-order difference in the x and y directions, representing the transverse difference feature and the longitudinal difference feature respectively, and is defined as equation 6.

$$G(x, y) = |d(x, y) - d(x + 1, y)| + |d(x, y) - d(x, y + 1)| \quad (6)$$

Where, $d(x, y)$ is the coefficient of the high frequency subband at the position (x, y) .

Since LGE, as an estimator describing the detailed information of images, lacks the extraction and description of large-scale structural information such as contours, Edge Strength (ES) is introduced. By calculating the amplitude of the gradient of edge pixels in the high frequency subband, the layering of its structure and edge contour is highlighted, that is as equation (7).

$$ES(x, y) = \sqrt{d_x(x, y)^2 + d_y(x, y)^2} \quad (7)$$

Where, $d_x = d * h_x$, $d_y = d * h_y$ respectively represent the convolution results of the high frequency subband pixel $d(x, y)$ with the Scharr operator [20] in the x and y directions.

$$h_x = \begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix}, h_y = \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ 3 & 10 & 3 \end{bmatrix}$$

represent the Scharr operators in the x and y directions respectively.

Thus, through the high-frequency measurement operator (HMO)

$$HMO(x, y) = LGE(x, y)^{\alpha_2} \cdot ES(x, y)^{\beta_2} \quad (8)$$

the high frequency subbands are fused. In (8), the parameters α_2 and β_2 are respectively used to adjust the weights LGE and ES in HMO .

By comparing the value of HMO in the high frequency subbands, the fusion rule of high frequency subbands is obtained as

$$HF^{j,k}(x, y) = \begin{cases} d_A^{j,k}(x, y) & \text{if } HMO_{d_A^{j,k}}(x, y) > HMO_{d_B^{j,k}}(x, y) \\ d_B^{j,k}(x, y) & \text{otherwise} \end{cases} \quad (9)$$

Where, $HF^{j,k}(x, y)$ is the fused image of the high frequency subband at scale j and direction k (horizontal, vertical or diagonal direction), $d_A^{j,k}(x, y)$ and $d_B^{j,k}(x, y)$ are the high frequency subbands corresponding to the scale j and direction k of the source images A and B respectively.

3.3 Overall framework of image fusion

According to the above method, the fusion framework based on NSCT is shown in Fig. 2.

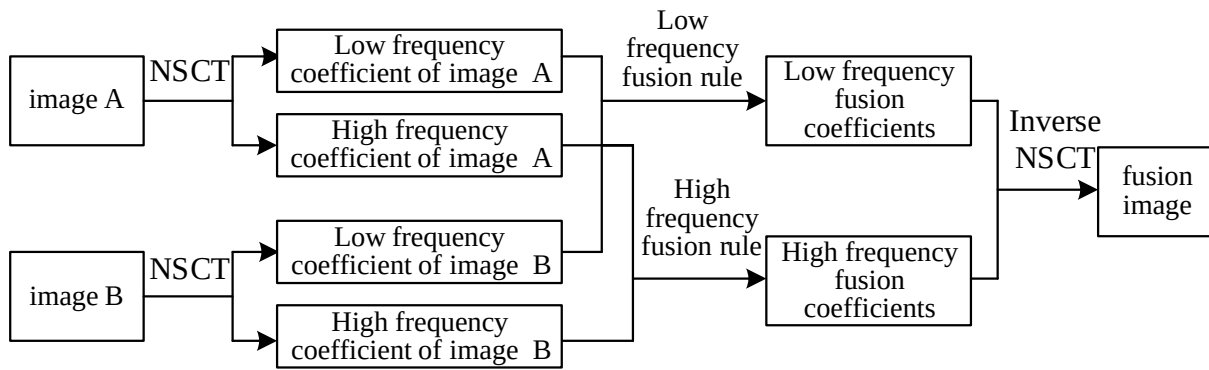


Fig. 2 Overall framework of image fusion

Firstly, the registered source image A and B are transformed by NSCT at level L respectively, here the value of L is set to 3 [16], then the low frequency subband coefficients $\{C_{j_0}^A(x, y), C_{j_0}^B(x, y)\}$ and the high frequency subband coefficients $\{d_{j,k}^A(x, y), d_{j,k}^B(x, y)\}$ ($j \geq j_0$) are obtained. Furthermore, as activity reflects the energy changes in a local area, and the local sharpness change enhances the contrast variation of activity. Therefore, for the low frequency subband, the LMO rule based on the activity function with local sharpness change is adopted for fusion to obtain $LF(x, y)$. While local gradient energy represents the difference changes, and edge strength represents the weighted difference of large-scale structures, both of them reflect the amount of image details. In high frequency subband, the HMO rule based on local gradient energy with edge strength is adopted for fusion to obtain $HF^{j,k}(x, y)$. Finally, the high frequency coefficients and low frequency coefficients are reconstructed by NSCT inverse transform to get the fusion image.

4 Experimental results and analysis

The superiority of the method proposed in this paper is verified through experiments on self-built datasets. The test images included multi-focus Lena images, CT and

MRI images, FA (Fluorescent Angiography) images and RF (Red-Free) images, etc., and were analyzed through visual and objective indicators. The experiment was conducted on CPU, with the code written in MATLAB, and the running time was within 2 seconds.

4.1 Objective evaluation indicators

This paper uses the commonly used measurement indicators to comprehensively and quantitatively evaluate the performance of different fusion methods, they are respectively [21]–[23]: peak signal to noise ratio (PSNR), structural similarity (SSIM), root mean square error (RMSE), Edge Information Retention (EIR), Information Entropy (IE). Among them, PSNR measures the similarity between the source image and the fused image by calculating the mean square error, the higher its value is, the smaller the distortion generated during the fusion process is, and the more similar the source image is to the fused image. SSIM measures the structural similarity between the source image and the fused image, the larger its value is, the more similar the source image is to the fused image. RMSE calculates the standard deviation between the fused image and the standard image, the smaller the value, the closer the fused image is to the standard image. EIR measures the retention of edge information and detailed textures of the source image, the higher the value, the more detailed and texture information the fused image contains, and the better the quality of

visual information obtained from the source image. IE reflects the information and clarity contained in the fused image, the higher its value, the more information and richer details the fused image contains, and the better the fusion performance.

4.2 Parameter setting

Adjusting one parameter is achieved by fixing other parameters and evaluating it from RMSE, SSIM, EIR, and IE to determine the optimal value of the parameter. Below, we take the image fusion of the left-focused Lena and right-focused Lena as an example to analyze the optimal parameters.

The measurement operator of high frequency subbands and low frequency subbands obtained by NSCT is used as the measurement method for the contour structure and edge texture information of the image, its parameter values directly affect the fusion quality and determine the effect of the final fused image. In order to determine the optimal values of the weight parameters in *HMO* and *LMO*, the weight parameters α_1 and β_1 in the *LMO* were analyzed first. The parameter β_1 was fixed as 1, and the value of parameter α_1 was set within the range of 0 to 0.01. The experimental results are shown in Table 1. It can be seen that: (1) with the increase of the parameter α_1 , each evaluation index changes accordingly. When $\alpha_1=0.006$, RMSE reaches the minimum value, SSIM, EIR and IE reach the maximum values, and the quality of the fused image improves well. Similarly, when $\alpha_1=0.006$, adjust the value of the parameter β_1 to achieve the best fusion effect performance. Eventually, the values of α_1 and β_1 are 0.006 and 1.05 respectively. (2) The parameters α_2 and β_2 in *HMO* were analyzed. Through experiments, it was found that whether it was RMSE, SSIM, EIR or IE, the fused image was not greatly affected by the values of α_2 and β_2 . Therefore, both α_2 and β_2 took the default value of 1.

Table 1: Comparison of multi-focus Lena fusion results at different values of α_1

α_1	RMSE	SSIM	EIR	IE
0.001	3.299	0.5124	0.685	7.795
0.002	3.298	0.5124	0.685	7.798
0.003	3.298	0.5125	0.687	7.803
0.004	3.297	0.5125	0.688	7.808

0.005	3.296	0.5125	0.688	7.811
0.006	3.296	0.5126	0.689	7.812
0.007	3.297	0.5126	0.688	7.810
0.008	3.297	0.5126	0.687	7.805
0.009	3.298	0.5126	0.685	7.802

Note: Bold indicates the optimal value

4.3 Frame structure experiment

Three representative multi-scale image fusion frameworks were applied to the multi-focus Lena image fusion, and the same fusion rules of high frequency and low frequency were adopted to conduct experiments on these three fusion frameworks and NSCT framework to prove the effectiveness of NSCT framework in this paper. These three multi-scale image fusion frameworks include the Pyramid Method (PM), Discrete Wavelet Transform (DWT), and Contourlet Transform (CT). The experimental results are shown in Figure 3, it can be seen from the local magnification of the cap edge at the lower right corner in Figure 3(c)- 3(f) that: Fig. 3(c) using the PM framework, the energy loss of the image is relatively large, which leads to a decrease in the contrast of the fused image and the appearance of slight artifacts at some contour edges. In Fig. 3(d) and Fig. 3(e), the DWT and CT frameworks are respectively applied, there are slight distortion and aberration at some edges. the fused images have some random noise, and the graininess in the detail areas is relatively strong. In Fig. 3(f), the NSCT framework is used. compared with the previous three multi-scale frameworks, both contrast and fidelity have been significantly improved, and the texture features of the source image have also been restored.

Table 2 shows the comparison of objective evaluation for multi-focus Lena image fusion under several multi-scale frameworks. It can be seen that compared with other multi-scale frameworks, the NSCT framework performs better in preserving edge details, as confirmed by PSNR, SSIM, RMSE, EIR and IE. Based on the authenticity of the texture features of the source image, the NSCT framework enhances the expression of texture features and details during the fusion process. This makes it particularly suitable for high-detail preservation applications, such as medical imaging etc.





Figure3: a-f Multi-focus Lena image fusion under several multi-scale frameworks

Table 2: Comparison of multi-focus Lena image fusion under several multi-scale frameworks

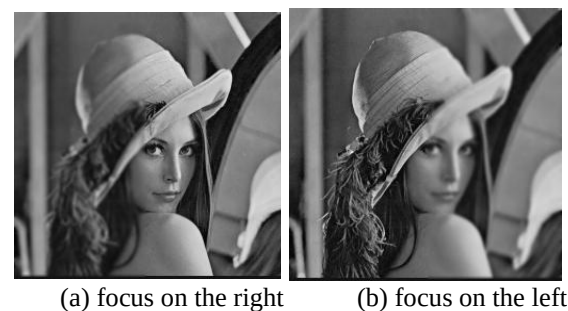
Framework	PSNR	SSIM	RMS	EIR	IE
k			E		
PM	10.27	0.512	3.325	0.65	7.64
	7	5		2	5
DWT	10.27	0.512	3.310	0.66	7.76
	8	5		1	5
CT	10.27	0.512	3.309	0.66	7.77
	9	4		8	5
NSCT	10.28	0.512	3.305	0.67	7.78
	1	6		4	9

4.4 Analysis of fusion results

To test the effectiveness of the algorithm proposed in this paper, it is compared with several frequency-domain image fusion methods proposed in recent years, including NSCT_PC [18], NSCT_EN_PCNN [17], NSST_PCNN [24], CWT_SR [11] and JBF [25]. NSCT_PC represents the fusion rule based on Phase Congruency (PC) and Laplace energy in the NSCT domain; NSCT_EN_PCNN represents the fusion rule based on the PCNN rule and image entropy in the NSCT domain; NSST_PCNN represents the fusion method based on the PCNN and energy strategy in the NSST domain; CWT_SR represents the fusion method based on sparse representation in the domain of complex wavelet transform; JBF represents a joint bilateral filtering fusion method based on enhancing edges and contrast.

Figure4 shows the artificial multi-focus Lena image fusion results of the method proposed in this paper and five representative frequency-domain fusion methods. It can be seen that in Fig. 4(c), Fig. 4(d), Fig. 4(e) and Fig. 4(f), a large amount of energy is lost during the fusion process,

resulting in a decrease in the contrast of the fused image. It can be seen from the local magnification images that the fused image shows insufficient performance in the detail texture of the hair, and some even have distortion. In Fig. 4(g), the JBF method significantly improves the fusion effect. However, there are greater differences and not smooth in local changes, and some details have discontinuous edge points. Compared with the previous several methods, the image texture fused by the method in this paper is clearer and the image layering is stronger. Table 3 shows the objective evaluation comparison of multi-focus Lena image fusion by different methods. It can be seen that the value of RMSE by the proposed method is the smallest, that is to say, the fusion result obtained by the proposed method is the closest to the standard image. In addition, the proposed method also outperforms other methods in terms of PSNR, SSIM, EIR, and IE evaluation indicators, which indicates that the fusion effect of the proposed method outperforms that of other methods in terms of detail contour and pixel intensity. The experimental results show that the performance of the proposed method is better than that of other methods both in subjective analysis and objective indicators.



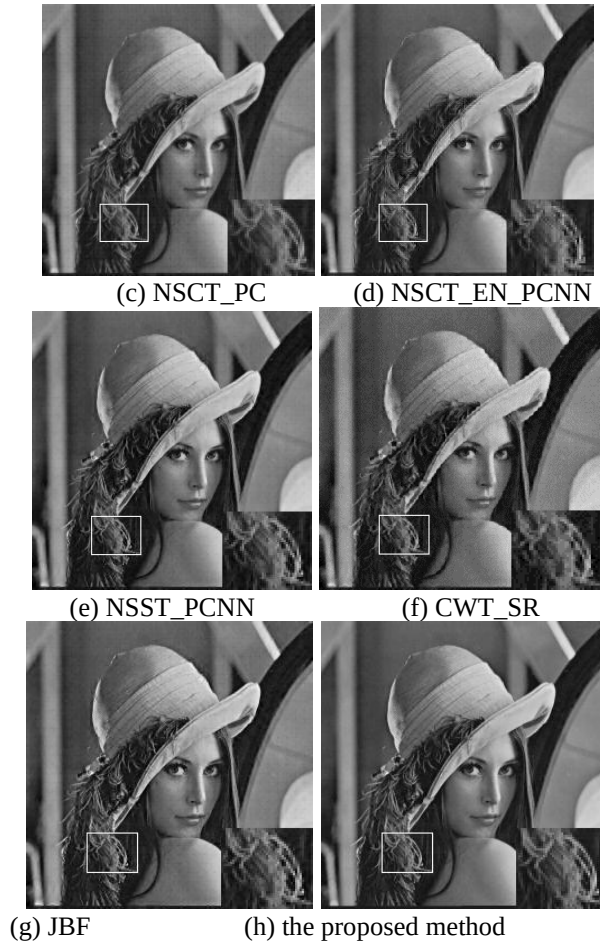


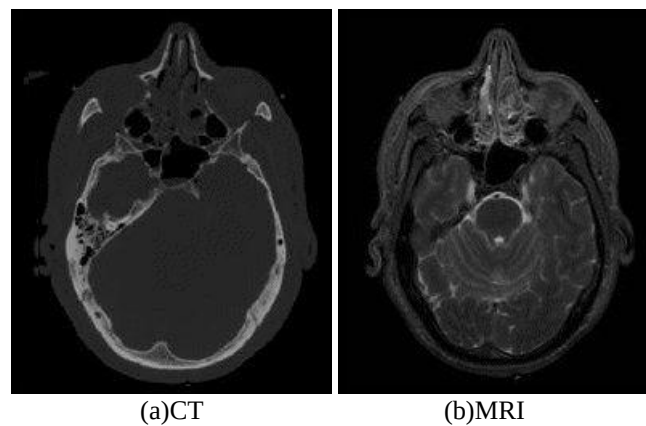
Figure 4: a-h Fusion results of different methods for multi-focus Lena image

Table 3: Objective efficacy comparison of different methods for multi-focus Lena fusion results

Methods	PSNR	SSIM	RMS	EIR	IE
			E		

NSCT_PC	10.27	0.512	3.31	0.66	7.74
	6	4	1	8	8
NSCT_EN_PCNN	10.27	0.512	3.31	0.66	7.75
N	7	5	2	4	6
NSST_PCNN	10.27	0.512	3.30	0.67	7.78
	8	5	8	1	2
CWT_SR	10.27	0.512	3.30	0.67	7.77
	8	3	9		5
JBF	10.27	0.512	3.30	0.67	7.78
	9	5	8	2	5
Proposed method	10.28	0.512	3.30	0.67	7.78
	1	6	5	4	9

CT images focus on describing the contour features of bones, while MRI images focus on describing the detailed texture features of soft tissues. Figure 5 shows the fused images of brain CT images and MRI images obtained by different methods. It can be seen that in Fig. 5(c), 5(d), 5(e), and 5(f), the fused images show edge distortion or blurring in terms of brain fiber texture information. It can also be observed from the local magnified images that the fused images have edge artifacts and information loss, which seriously affects the accuracy of medical diagnosis. This situation is improved in Fig. 5(g), but it is still insufficient in the presentation of texture details. In Fig. 5(h), the detail texture of soft tissue structure in the fused image is clearer compared with the previous several methods, in the local magnified image, the gully texture has better contrast. Table 4 shows the objective evaluation comparison of CT/MRI image fusion by different methods. It can be seen that the method proposed in this paper has achieved considerable results in various indicators such as PSNR, SSIM, EIR, and IE, which indicates that the proposed method has better performance in aspects such as image feature transfer and texture detail expression. The experimental results show that in CT/MRI image fusion, the method proposed in this paper outperforms other methods in both subjective analysis and objective indicators.



(a)CT

(b)MRI

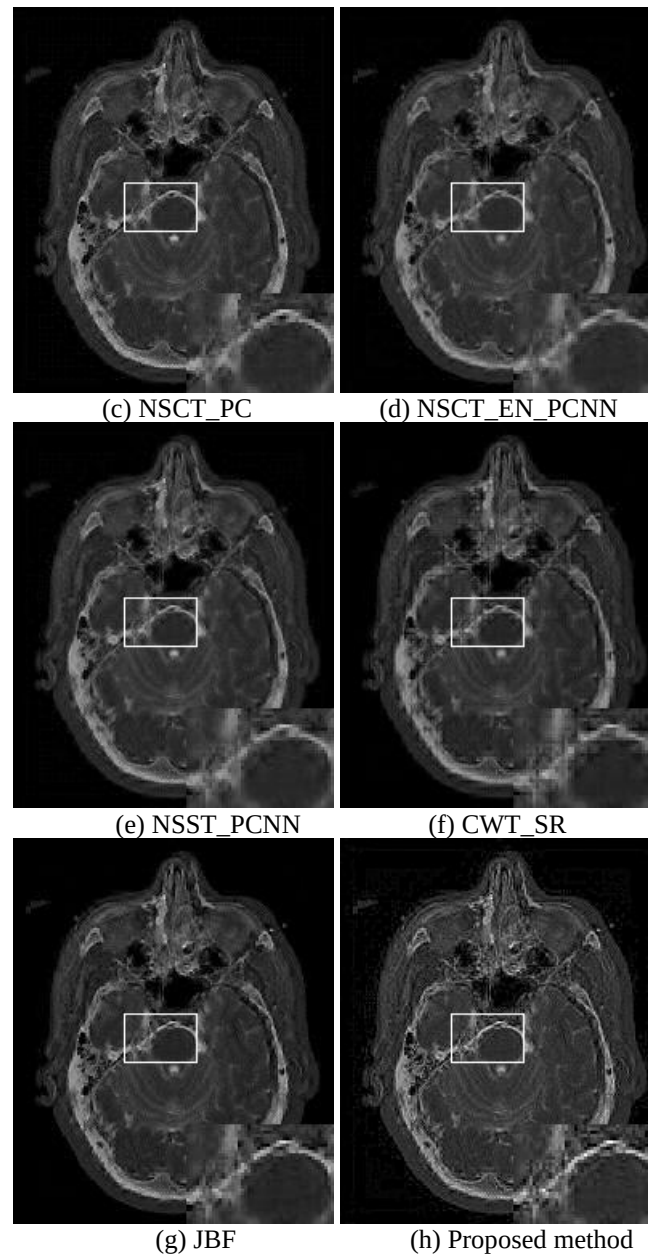


Figure 5: a-h Fusion results of different methods for CT image and MRI image

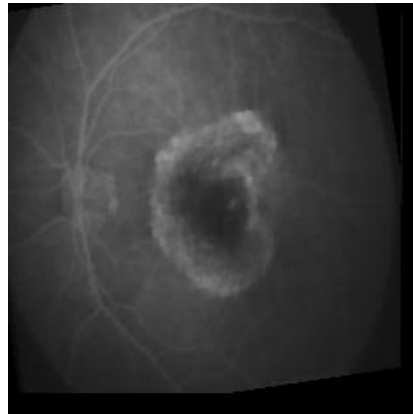
Table 4: Objective efficacy comparison of different methods for CT image and MRI image

Methods	PSNR	SSIM	EIR	IE
NSCT_PC	9.847	0.5235	0.668	6.781
NSCT_EN_PCNN	9.826	0.5233	0.656	6.665
NSST_PCNN	9.845	0.5235	0.665	6.783
CWT_SR	9.839	0.5232	0.664	6.776
JBF	9.851	0.5237	0.673	6.826
Proposed method	9.856	0.5238	0.684	6.857

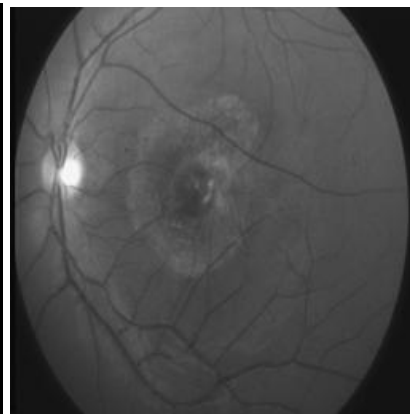
Retinal FA (Fluorescent Angiography) images are the images of retinal blood vessels obtained by injecting sodium fluorescein contrast agent, and RF (Red Free) images are the images obtained by irradiating the retina with short-wavelength light. They play an important role in the diagnosis and monitoring of fundus diseases. Figure 6 and Figure 7 show the fusion images of FA images and RF images obtained by different methods. It can be seen that the other methods have problems such as fuzzy pseudo-shadow and information loss in the soft tissues of the fundus vessels at different level, and even vascular edge distortion in the local magnified images, which is not beneficial to the accuracy of medical diagnosis. The method proposed in this paper, relatively speaking, provides a more accurate description of the FA/RF image and better preserves the small structure information of the vascular tissues. Tables 5 and Tables 6 show the objective

evaluation comparison of FA/RF image fusion by different methods. It can be seen that the method proposed in this paper has obvious advantages in PSNR, SSIM, EIR, and IE indicators, which indicates that the proposed method has greatly improved the expression of the detailed texture of fundus vascular tissue in FA and RF

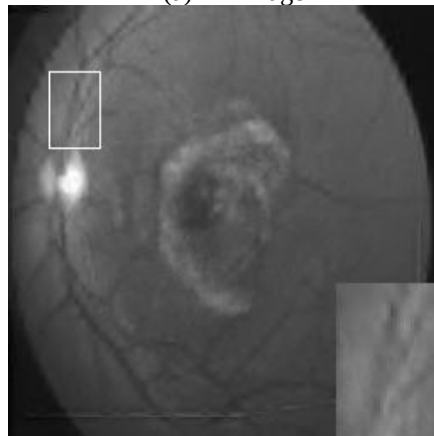
images and enhanced the contrast of the fused images. It is convenient to observe the diseased region of the patient. The experimental results show that the proposed method outperforms the other methods in both subjective analysis and objective evaluation in FA/RF image fusion.



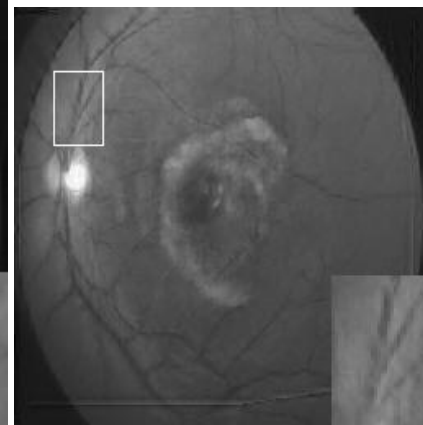
(a) FA image



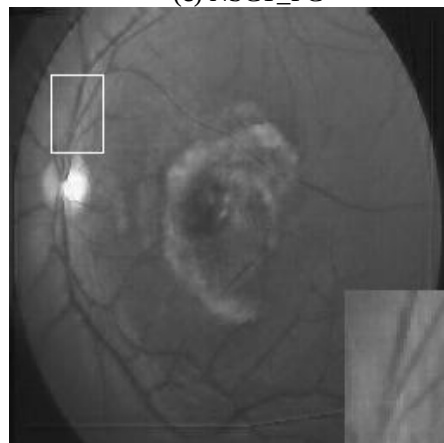
(b) RF image



(c) NSCT_PC



(d) NSCT_EN_PCNN



(e) NSST_PCNN



(f) CWT_SR

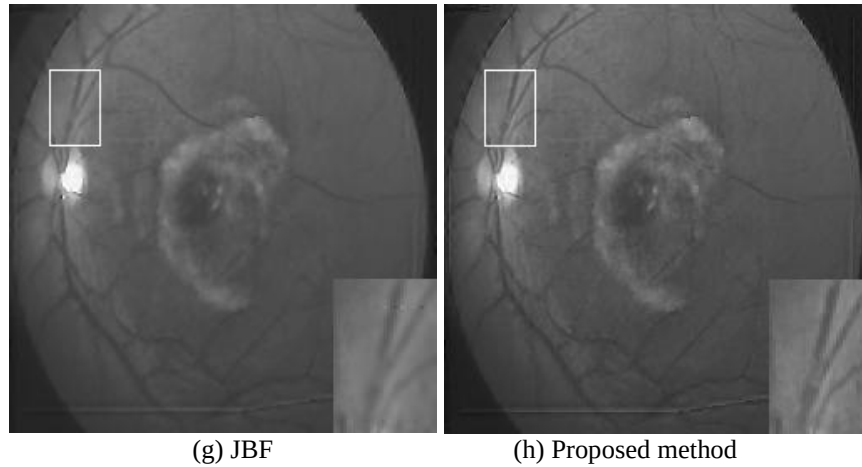
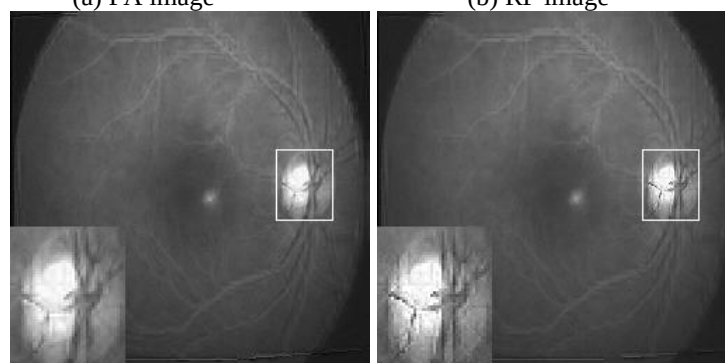
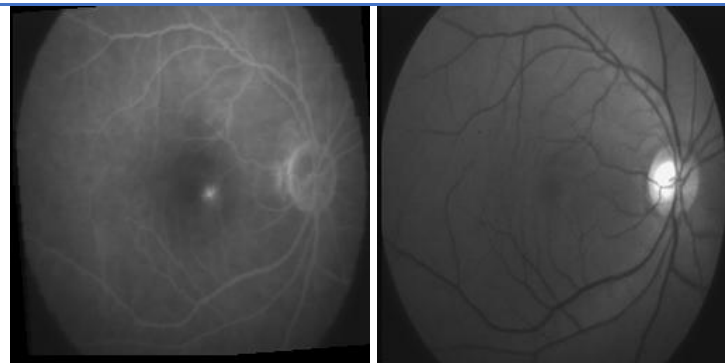


Figure 6: a-h Fusion results of different methods for FA image and RF image

Table 5: Objective efficacy comparison of different methods for FA image and RF image

Methods	PSNR	SSIM	EIR	IE
NSCT_PC	8.547	0.5051	0.486	6.588
NSCT_EN_PCNN	8.604	0.5053	0.531	6.604
NSST_PCNN	8.635	0.5054	0.549	6.628
CWT_SR	8.578	0.5052	0.536	6.606
JBF	8.642	0.5054	0.558	6.725
Proposed method	8.673	0.5055	0.612	6.817



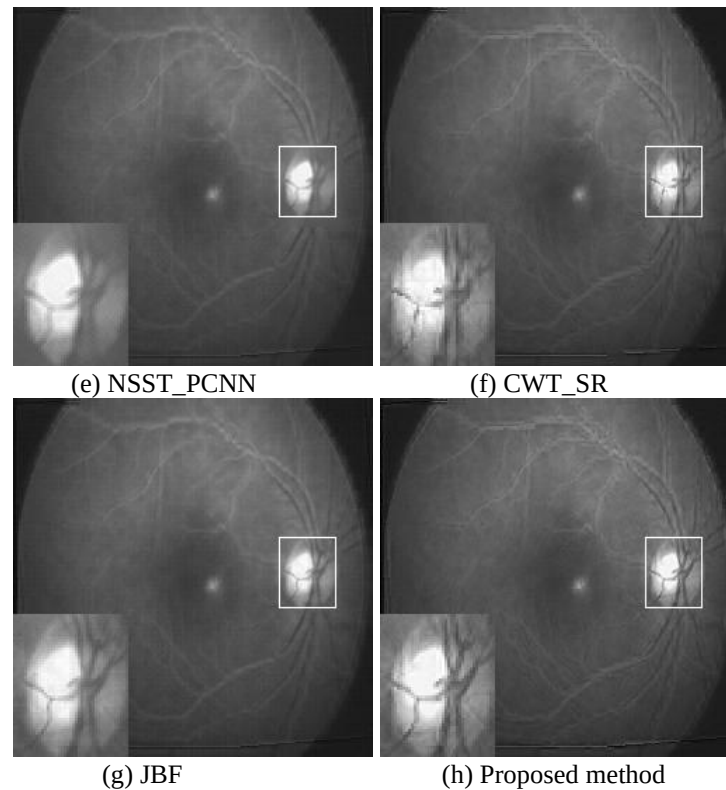


Figure 7: a-h Fusion results of different methods for FA image and RF image

Table 6: Objective efficacy comparison of different methods for FA image and RF image

Methods	PSNR	SSIM	EIR	IE
NSCT_PC	8.512	0.5053	0.529	6.758
NSCT_EN_PCNN	8.535	0.5044	0.539	6.789
NSST_PCNN	8.557	0.5045	0.541	6.791
CWT_SR	8.525	0.5054	0.534	6.783
JBF	8.564	0.5045	0.546	6.788
Proposed method	8.582	0.5046	0.572	6.804

As mentioned above, the fusion effect of the method proposed in this paper is better than the five popular image fusion methods and three image fusion frameworks in the field of frequency domain in recent years, and it achieves better visual effects, thereby ensuring the reliability of subsequent processing.

5 Conclusion

A new image fusion method based on NSCT is presented. Firstly, the fusion rule scheme of source image decomposed by NSCT is constructed, the low frequency fusion rule is based on activity with local sharpness changes, and the high frequency fusion rule is based on local gradient energy with edge strength. In this paper, artificial multi-focus images, brain CT and MRI images, FA image and RF image are tested. The results show that

the presented method has excellent fusion property for multi-focus and multimodal image fusion.

Acknowledgements

The authors would like to thank Weifang University teaching reform research project (No.2023YB003, No. 2024YB034), Weifang Science and Technology Development Project (2025GX016) for financially supporting this research.

References

- [1] Karim S, Tong J, Li J, et al. Current advances and future perspectives of image fusion: A comprehensive review. *Information Fusion*, 90: 185-217, 2023. <https://doi.org/10.1016/j.inffus.2022.09.019>.
- [2] Singh S, Singh H, Bueno G, et al. A review of image fusion: Methods, applications and performance metrics. *Digital Signal Processing*, 137: 104020, 2023. <https://doi.org/10.1016/j.dsp.2023.104020>.
- [3] Azam M A, Khan K B., Salahuddin S, et al. A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Computer in Biology and Medicine*, 144: 105253, 2022. <https://doi.org/10.1016/j.combiomed.2022.105253>.
- [4] Hermessi H, Mourali O, Zagrouba E. Multimodal medical image fusion review: Theoretical background and recent advances. *Signal Processing*, 2021, 183: 1-27. <https://doi.org/10.1016/j.sigpro.2021.108036>.
- [5] Zhang W, Jiao L, Liu. F. Adaptive contourlet fusion clustering for SAR image change detection. *IEEE*

- Transactions on Image Processing, 31: 2295-2308, 2022. <https://doi.org/10.1109/TIP.2022.3154922>.
- [6] Li X T, Guo H Y. Fusion of deep convolutional neural networks and brain visual cognition for enhanced image classification. Informatica, 49(16): 37-52, 2025. <http://dx.doi.org/10.31449/inf.v49i16.7787>.
- [7] Li H F, He X G, Tao D P, et al. Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning. Pattern Recognition, 79: 130-146, 2018. <https://doi.org/10.1016/j.patcog.2018.02.005>.
- [8] Toet A. Image fusion by a ratio of low-pass pyramid. Pattern Recognition Letters, 9(4): 245-253, 1989. [https://doi.org/10.1016/0167-8655\(89\)90003-2](https://doi.org/10.1016/0167-8655(89)90003-2).
- [9] Pu T, Ni G Q. Contrast-based image fusion using the discrete wavelet transform. Optical Engineering, 39(8): 2075- 2082, 2000. <http://dx.doi.org/10.1117/1.1303728>.
- [10] Cunha A. L, Zhou J P. The non-subsampled contourlet transform: theory, design, and applications. IEEE Transactions on Image Processing, 15(10): 3089- 3101, 2006. <https://doi.org/10.1109/TIP.2006.877507>.
- [11] Liu Y, Liu S P, Wang Z F. A general framework for image fusion based on multi-scale transform and sparse representation. Information Fusion, 24: 147-164, 2015. <https://doi.org/10.1016/j.inffus.2014.09.004>.
- [12] Qu X, Yan J, Yang G. Multifocus image fusion method of sharp frequency localized contourlet transform domain based on sum-modified-laplacia. Optics and Precision Engineering, 17(5): 1203-1211, 2009.
- [13] Do M N, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation. IEEE Transaction on Image Processing, 14(12): 2091-2106, 2005. <https://doi.org/10.1109/TIP.2005.859376>.
- [14] Tawfik N, Elnemr H A, Fakhr M, et al. Survey study of multimodality medical image fusion methods. Multimedia Tools and Applications, 80: 6369-6396, 2020. <https://doi.org/10.1007/s11042-020-08834-5>.
- [15] Ramlal S D, Sachdeva J, Ahuja C K, et al. Multimodal medical image fusion using non-subsampled shearlet transform and pulse coupled neural network incorporated with morphological gradient. Signal. Image and Video Processing, 12: 1479-1487, 2018. <https://doi.org/10.1109/TIM.2018.2838778>.
- [16] Ramlal S D, Sachdeva J, Ahuja C K, Khandelwal N. An improved multimodal medical image fusion scheme based on hybrid combination of nonsubsampling contourlet transform and stationary wavelet transform. International Journal of Imaging Systems and Technology, 29(2): 146-160, 2019. <https://doi.org/10.1002/ima.22310>.
- [17] Guo W. Research on medical image fusion algorithm based on nsct and neural network. Lanzhou Jiaotong University, 2024.
- [18] Zhu Z, Zheng M, Qi G, et al. A phase congruency and local laplacian energy-based multimodality medical image fusion method in NSCT domain. IEEE Access, 7: 20811-20824, 2019. <https://doi.org/10.1109/ACCESS.2019.2898111>.
- [19] Zhu J, Liu H, Li S, et al. Two-Channel medical image fusion combining local entropy and gradient energy. Journal of Computer-Aided Design & Computer Graphics, 36(6): 857-874, 2024. <https://dx.doi.org/10.3724/SP.J.1089.2024.19899>.
- [20] Liu X, Lin Y Z. Comparative study of multiple image edge detection operators applied to size measurement of passiflora edulia sims. Chinese Journal of Tropical Crops, 43(12): 2554-2563, 2022. <https://doi.org/10.3969/j.issn.1000-2561.2022.12.018>.
- [21] Zhu Y. No-reference image quality assessment based on statistical information. Huaqiao University, 2021.
- [22] Liu Y. No-reference evaluation of visible and infrared fusion image sharpness. China University of Petroleum, 2020.
- [23] Xydeas C S, Petrovic V. Objective image fusion performance measure. Electronics Letters, 36(4): 308-309, 2000. <http://dx.doi.org/10.1049/el:20000267>.
- [24] Tan W, Zhang J J, Xiang P, et al. Infrared and visible image fusion via NSST and PCNN in multiscale morphological gradient domain. Proceedings of the SPIE 11353, Optics, Photonics and Digital Technologies for Imaging Applications VI. 113531E, 2020. <https://doi.org/10.1117/12.2551830>.
- [25] Li X S, Zhou F Q, Tan H S, et al. Multimodal medical image fusion based on joint bilateral filter and local gradient energy. Information Sciences, 569: 302-325, 2021. <https://doi.org/10.1016/j.ins.2021.04.052>.