Multimodal Medical Image Fusion Based on Parameter Adaptive PCNN and Latent Low-rank Representation

WANG Wenyan, ZHOU Xianchun*, YANG Liangjian

(*Nanjing University of Information Science and Technology*, *Nanjing 210044*)

Abstract: Medical image fusion has been developed as an efficient assistive technology in various clinical applications such as medical diagnosis and treatment planning. Aiming at the problem of insufficient protection of image contour and detail information by traditional image fusion methods, a new multimodal medical image fusion method is proposed. This method first uses non-subsampled shearlet transform to decompose the source image to obtain high and low frequency subband coefficients, then uses the latent low rank representation algorithm to fuse the low frequency subband coefficients, and applies the improved PAPCNN algorithm to fuse the high frequency subband coefficients. Finally, based on the automatic setting of parameters, the optimization method configuration of the time decay factor αe is carried out. The experimental results show that the proposed method solves the problems of difficult parameter setting and insufficient detail protection ability in traditional PCNN algorithm fusion images, and at the same time, it has achieved great improvement in visual quality and objective evaluation indicators.

Keywords: Image Fusion, Non-subsampled Shearlet Transform, Parameter Adaptive PCNN, Latent Low-rank Representation

1 Introduction

Due to the diversity of imaging mechanisms, different modes of medical images focus on different categories of organ or tissue information. Computed tomography (CT) imaging can precisely detect dense structures such as human bones and implants. Magnetic resonance (MR) imaging provides high-resolution anatomical information on soft tissues. In order to obtain enough information for accurate diagnosis, doctors usually need to sequentially analyze medical images captured by different modes, but this separation method may still cause inconvenience in many cases. Image fusion technology[1-2] has become an effective way to solve this problem. Its purpose is to generate a composite image that integrates complementary information contained in multiple images of different modalities.

In the past decades, various medical image fusion methods[3-6] have been proposed. Most medical image fusion methods are introduced under the framework based on multi-scale transform (MST) to pursue good perceptual results. A large number of research results show that choosing a reasonable multi-scale transformation method and an effective fusion scheme for improving the high and low frequency sub-band coefficients can significantly improve the performance of fusion methods based on the multi-scale transformation framework.

The non-subsampled shearlet (NSST) algorithm proposed in literature[7] is an improved result on the basis of retaining the advantages of shearlet transform, which avoids the pseudo-Gibbs phenomenon, has high operation efficiency and low complexity, and can Efficiently extract details such as edge texture in the source image; Johnson et al. developed a pulse-coupled neural network (PCNN) model[8-10], the standard PCNN model is usually simplified to reduce computational complexity while retaining basic visual Cortical characteristics, but the appropriate value of the parameters can only be estimated through manual adjustment or a large number of training, which has become a serious problem and restricts the further development of PCNN; literature[11] proposes within the framework of NSST, pulse-coupled neural The network (PCNN) adds fusion rules to effectively extract the gradient features and information retention of the image, but the setting of many parameters in the PCNN is also a major difficulty; literature[12] proposes an image segmentation method based on a simplified PCNN model (SPCNN). The size of PCNN free parameters can be automatically set to achieve higher segmentation accuracy; literature[13] improved the SPCNN model to obtain a parameter adaptive PCNN (PAPCNN) model and applied it to image fusion. Experiments show that the PAPCNN model has Faster convergence speed, and better results when applied in image fusion experiments.

Literature[14] proposes a low-rank representation (LRR) image fusion method, which can capture the global structure of the data, but due to the lack of dictionary learning of LRR, the local structure preservation ability is limited; literature[15] proposes a new method based on The image fusion method of dictionary learning and LRR, the algorithm uses the K-SVD algorithm to learn various sub-dictionaries, and then constructs a global dictionary, which has achieved good fusion results in both global and local structures; literature[16] proposes a method based on latent A low-rank representation (LatLRR) image fusion method, which uses a weighted average strategy and a summation strategy to fuse the global structure and local structure information, and better preserve the contour information in the source image.

This paper first combines the advantages of NSST transformation, latent low-rank representation and PAPCNN model, and then proposes an improved PAPCNN time decay factor αe method, which adjusts the decay speed of the dynamic threshold Eij to achieve better fusion Effect. This algorithm not only solves the problem of retaining detail information in the low-frequency subband of NSST domain, but also overcomes the difficulty of setting free parameters in the traditional PCNN model, which can not only capture target information, but also retain detail contour information.

2 Related Work

2.1 Non-subsampled Shearlet Transform (NSST)

The NSST transform completes the multi-scale and multi-directional decomposition of the source image through the non-subsampled pyramid filter combination filter (NSPF) and shearlet filter (SFB). NSST performs non-subsampling operations in the process of image decomposition and reconstruction, so it has translation invariance, can effectively overcome the pseudo-Gibbs effect when reconstructing images, and SFB can achieve direction localization.

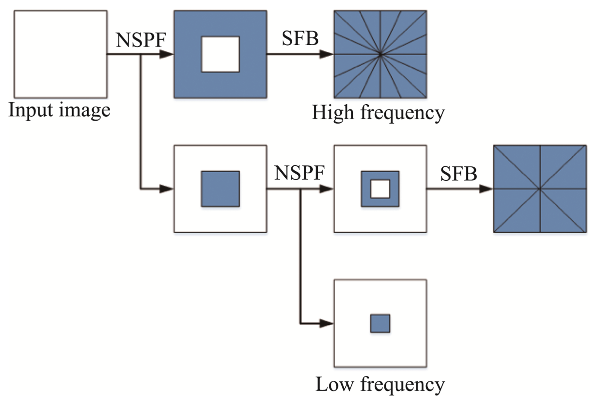


Fig.1 NSST Exploded Diagram

Fig.1 is a schematic diagram of NSST transformation decomposition. NSST has translation invariance, which is sensitive to edge contour feature information[17], and can effectively extract the detailed information of the source image[18], which is suitable for image processing. Based on the above characteristics, NSST is selected as the multi-scale transformation method for image fusion in this paper.

2.2 Improvement of Adaptive PCNN (PAPCNN) and Parameters *αe*

The basic PCNN model is applied to image fusion, the key problem is to determine the parameters. Such as time decay factor, connection strength and amplitude, etc., to avoid large experimental errors caused by manual selection of parameters, Chen et al.[19] proposed a simplified parameter adaptive PCNN (SPCNN) model, which is used in the field of image segmentation Has been widely used. Many experimental results show that the SPCNN model is also effective in the field of image fusion, especially when the PAPCNN model is applied to the fusion of high-frequency coefficients of multi-scale transformation, the overall image fusion quality is greatly improved.

The SPCNN model is described as follows:

$$F_{ij}[n]=S_{ij} \tag{1}$$

$$L_{ij}[n]=V_L\sum_{kl}W_{ijkl}Y_{kl}[n-1] \tag{2}$$

$$U_{ij}[n]=e^{-\alpha_f}\,U_{ij}[n-1]+F_{ij}[n](1+\beta L_{ij}[n]) \tag{3}$$

$$Y_{ij}[n]=\begin{cases}1, ij\ \mathrm{U}_{ij}[n]>E_{ij}[n-1]\\ 0, \text{otherwise}\end{cases} \tag{4}$$

$$E_{ij}[n]=e^{-\alpha_e}\,E_{ij}[n-1]+V_E Y_{ij}[n]) \tag{5}$$

In the above SPCNN model, $F_{ij}[n]$ and $L_{ij}[n]$ represent the input and connection input of n iterations at position (i, j) respectively. In the iterative process, S_{ij} is the external input, and is fixedly connected to $F_{ij}[n]$. The previous firing state of $L_{ij}[n]$ adjacent neurons determine the current strength through synaptic weights. Fig.2 shows the basic structure of the SPCNN model.

The internal activity $U_{ij}[n]$ is composed of its previous iteration value $e^{-\alpha_f}\,U_{ij}[n-1]$ and the nonlinear modulation of the input $F_{ij}[n](1+\beta L_{ij}[n])$ and connection input, where the parameter α_f is the exponential decay coefficient of the internal activity and the parameter β is the connection strength. The output $Y_{ij}[n]$ of the PCNN model determines the trigger event of the model, and its output has two states: fired ($Y_{ij}[n]$=1) and unfired ($Y_{ij}[n]$=0), this state is determined by the relationship between the internal activity $U_{ij}[n]$ and the dynamic threshold $E_{ij}[n-1]$ of the previous iteration. As shown in formula(4), when $U_{ij}[n]>E_{ij}[n-1]$, fired ($Y_{ij}[n]$=1). Substitute the value of the output $Y_{ij}[n]$ into the formula(5) to update the dynamic threshold $E_{ij}[n]$, where the parameter αe is the time decay factor, and V_E is the amplitude of $E_{ij}[n]$. In SPCNN, the initial values of $Y_{ij}[0]$, $U_{ij}[0]$, and $E_{ij}[0]$ are all 0. Therefore, all neurons with non-zero intensity fire in the first iteration. The architecture of the SPCNN model is shown in Fig.2.

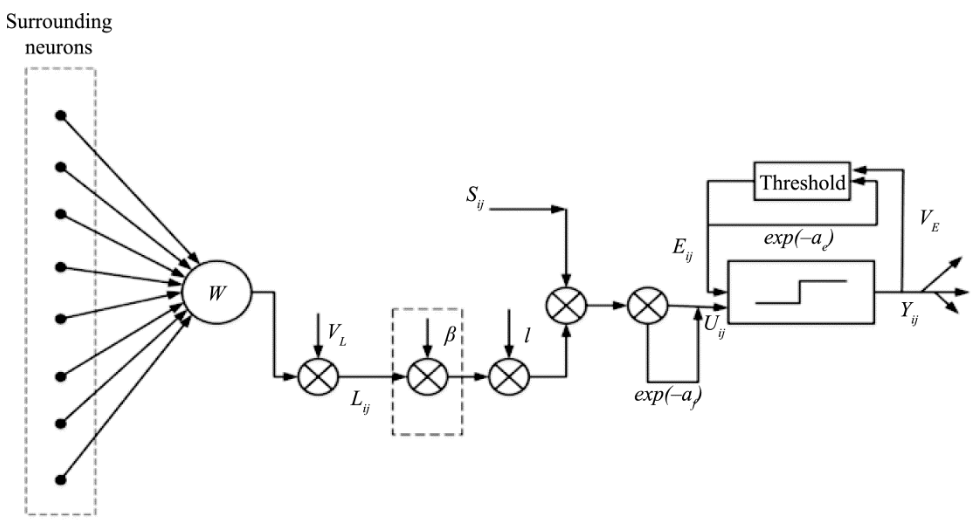


Fig.2 PAPCNN Structural Model

There are five parameters in the SPCNN model: connection strength β, amplitude of connection input L and dynamic threshold E, internal activity U and time decay factors V_L and V_E of dynamic threshold E. In addition, it can be obtained from the simplified formula of the SPCNN model that both α_f and α_e act as weight factors of $\sum_{kl}W_{ijkl}Y_{kl}[n-1]$, therefore, these two parameters can be processed in the model as a whole λ. So there are actually only four parameters: α_e , α_f , β and V_E.

In the paper we improved the parameters αe. The parameter αe is a time decay constant of $E_{ij}[n]$, and images with different gray levels require completely different parameters αe. In high-intensity areas, smaller αe is more suitable for fusion of different parts of the image; for darker images, larger parameters αe can achieve better fusion results. In this section, we introduce an adaptive αe computation method based on the Stevens power law[20]:

$$S=K\times I^n \tag{6}$$

Among them, S is the perceived strength, K is a constant, I is the physical strength, and the index n changes with the change of objective conditions.

When using formula(6) to explain the relationship between perceived brightness and actual brightness (the gray value of the grayscale image is the actual brightness), select n=0.5. In fact, it can be understood as a process of nonlinear mapping from the spatial domain to the visual domain. The gray value of the high gray area in the spatial domain is compressed, and the gray value of the low gray area is increased to complete the process of mapping to the visual domain. Generally, the target pixels in the fused image are mostly high gray values. Since the gray values are compressed, to prevent the dynamic threshold value E_{ij} from decaying too fast during the mapping process and causing the neurons corresponding to non-target pixels to misfire, a smaller parameter αe is used. to reduce the E_{ij} attenuation speed; on the contrary, the low gray area contains a large number of non-target pixels, so use a larger αe to increase the gray level in the area to avoid information redundancy.

The highest grayscale T_{Otsu} of the image background and the grayscale distribution of the background σ_b play a crucial role in the quality of the fusion.

The specific analysis process is as follows:

For the highest grayscale T_{Otsu} of the background. If T_{Otsu} is close to or even completely overlaps with the lowest gray level of the target pixel, a smaller αe is required to achieve the best fusion effect. Otherwise, a large αe parameter is difficult to achieve the ideal fusion effect on the edge of the image. Therefore, the relationship between αe and T_{Otsu} is set as an inverse relationship in this paper.

The grayscale distribution of the background σ_b. When T_{Otsu} is close to or completely overlaps with the lowest grayscale of the target pixel, if σ_b is small, it means that most of the background pixels are close to the grayscale of the target pixel, so a smaller αe is needed for image fusion; if σ_b Larger, larger αe will give better fusion results. On the other hand, when the lowest gray level of the target pixel is much larger than T_{Otsu}, σb has little effect on the fusion quality.

To sum up, in this paper, αe and σb are set as a proportional relationship. αe is negatively correlated with T_{Otsu}, and positively correlated with σb. Through a large number of experiments, it is found that when the relationship between the three satisfies formula(7), the algorithm can achieve the best state between fusion quality and computational efficiency:

$$\alpha_e=\begin{cases}\left(\dfrac{\sigma_b}{T_{Otsu}}\right)2, if\ \dfrac{\sigma_b}{T_{Otsu}}>0.1\\ \dfrac{\sigma_b}{T_{Otsu}}, else\end{cases} \tag{7}$$

Among them, T_{Otsu} is the highest gray level of the image background, and σb is the background standard deviation.

The above PAPCNN model is mainly used in the field of image segmentation. In the paper, experiments show that it is also effective for image fusion. Especially when fusing high-frequency coefficients obtained by multi-scale transformation, it has a greater improvement in visual and objective indicators than traditional algorithms. The essence of PCNN in image segmentation is to transfer the intensity information of a pixel to adjacent pixels, so that the information of adjacent pixels can be coupled and comprehensively utilized. For the high-frequency coefficient fusion of multi-scale transformation, the absolute value of the coefficient represents the activity level information to a certain extent, and there is a positive correlation between the activity level and the absolute value of the coefficient. The PCNN model performs image segmentation based on pixel intensity and can also distinguish the absolute value of high-frequency coefficients of multi-source images. The fusion of high-frequency coefficients by PCNN inherits its advantages in image segmentation, that is, the activity level information of adjacent pixels can be transferred to each other through the connection strength β, so as to obtain a more stable activity measure. Therefore, the image segmentation based on PCNN can be extended to the field of image fusion based on PCNN. In this paper, the above-mentioned PAPCNN model is selected as the fusion scheme of multi-scale transformation high-frequency coefficients.

2.3 Latent Low-rank Representation

The traditional LRR algorithm only considers the global structure of the image and the problem of limited application range. Liu[21] and others proposed the LatLRR algorithm based on the LRR theory, which can extract the global and local structures from the source image, and improve the protection ability of the main information and local salient information. The algorithm can be simplified to:

$$\begin{aligned}&\min_{Z,L,E} \|Z\|_*+\|L\|_*+\lambda\|E\|_1\\&s.t., X=XZ+LX+E\end{aligned} \tag{8}$$

Among them, λ is the balance coefficient and $\|\bullet\|_*$ is always greater than zero, which is the nuclear norm, $\|\bullet\|_1$ is the sum of the singular values of the matrix, and is the L1 norm. X is the observed data matrix, Z is the low-rank coefficient matrix, L is the significant coefficient matrix, and E is the sparse noise part. Eq. (8) can be solved by inexact augmented Lagrangian multipliers (ALM). In Eq. (8), XZ represents the low-rank part of the image, and LZ represents the salient part of the image. The source image is subjected to the LatLRR algorithm to obtain low-rank and salient parts, and represent low-rank parts and salient parts, respectively. The low-rank part contains more global structural information in the source image. In order to better protect the contour information in the image, the weighted average strategy is used to process the low-rank part; the salient part contains the local structure information and salient features of the source image, and the source image The salient features of must be preserved in the fused image as much as possible, so the summation strategy is used to process the salient parts.

The LatLRR algorithm is based on the processing process of the global and local structures, X_{lrr} and X_s overcomes the problems of unobvious salient features, insufficient protection of main information, and artifacts after the fusion of traditional low-frequency algorithms. In this paper, it is considered effective to use LatLRR algorithm for low-frequency sub-band information fusion. Especially for the low-frequency sub-band information after NSST decomposition, the LatLRR algorithm approximates the low-frequency sub-band coefficients from both global and local aspects, which can represent the low-frequency information of the image in detail, and the fused image is more natural. Based on the above considerations, we adopt the LatLRR algorithm model as a fusion scheme for multi-scale transform low-frequency coefficients.

3 NSST-PAPCNN-LatLRR Fusion Algorithm

The specific fusion steps of this algorithm are shown in Fig.3. In this paper, the source image CT and MRI are fused, and the detailed fusion process includes four steps: NSST decomposition, high-frequency sub-band fusion, low-frequency sub-band fusion, and NSST reconstruction.

Step 1: NSST decomposition

First, perform NSST decomposition on the two source images A and B to obtain the decomposed subband coefficients $\{H_A^{l,k}, L_A\}$ and $\{H_B^{l,k}, L_B\}$. $H_A^{l,k}$ indicates the high-frequency sub-band coefficients of the source image A in the decomposition level l and decomposition direction k, and L_A represents the low-frequency sub-band coefficients of the source image A after decomposition. Similar to A, $H_B^{l,k}$ and L_B represent the corresponding subband coefficients of image B after decomposition.

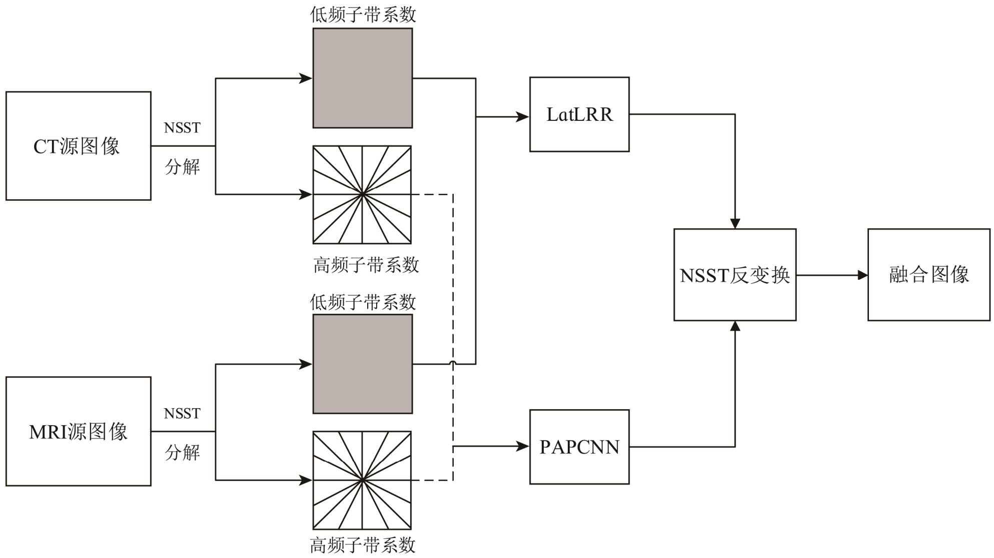


Fig.3 NSST-PAPCNN-LatLRR Algorithm

Step 2: Fusion of high frequency subband coefficients

In this paper, the improved parameter adaptive PCNN (PAPCNN) model in Section 2 is used to fuse high-frequency subbands. Based on the above discussion, the absolute value map of the high-frequency subband coefficients is selected as the network input of the fusion process, that is, the feed input is $F_{ij}[n]=| H_S^{l,k} | S\in\{A, B\}$. During the iteration of high-frequency information, the overall activity level of high-frequency coefficients was measured by the total firing time. According to the basic formula of PAPCNN, the cumulative trigger time can be calculated by the sum of the last trigger time and the final output of this time:

$$T_{ij}[n]= T_{ij}[n-1]+Y_{ij}[n] \tag{9}$$

Among them, N is the total number of iterations, and $T_{ij}[n]$ is the firing time of the neuron. $T_{A,ij}^{l,k}$ indicates the ignition time of the high-frequency sub $H_A^{l,k}$ in the PAPCNN model, and $T_{B,ij}^{l,k}$ has the same meaning for $H_B^{l,k}$. The fusion band is calculated by formula (10):

$$H_F^{l,k}(i,j)=\begin{cases} H_A^{l,k}(i,j), \text{ if } T_{A,ij}^{l,k}[N]\geqslant T_{A,ij}^{l,k}[N] \\ H_B^{l,k}(i,j), \text{ otherwise}\end{cases} \tag{10}$$

Formula (10) shows that the final high-frequency sub-band coefficient is represented by the high-frequency sub-band coefficient with a larger ignition number, and the sub-band coefficient with a longer ignition time (that is a larger ignition number) is the final high-frequency sub-band coefficient. In this paper, the optimal high-frequency fusion coefficient is obtained by adaptively adjusting the size of the time decay factor αe.

Step 3: Fusion of low frequency subband coefficients

The low-frequency sub-band coefficients still contain some important details, so the fusion process of low-frequency sub-band coefficients is also a key part of image fusion. In this paper, the latent low-rank representation is selected to fuse low-frequency sub-band coefficients[21]. First, the low-frequency part is decomposed to obtain the low-rank part L_{C_lrr} and the salient part L_{C_s}, among which $C\in\{A, B\}$; then the low-rank part and the salient part are fused using the weighted average fusion strategy and the sum fusion strategy respectively, and the fused low-rank part L_{lrr} and the salient part L_{C_s} are obtained; Finally, the low-rank part and the significant part are accumulated to obtain the fused low-frequency coefficients. The specific implementation steps of low-frequency sub-band fusion are shown in Fig.4:

Step 4: NSST reconstruction

Finally, inverse NSST reconstruction is performed on the fused high-frequency sub-band coefficients $H_F^{l,k}$ and low-frequency sub-band coefficients L_F to obtain the fused image F.

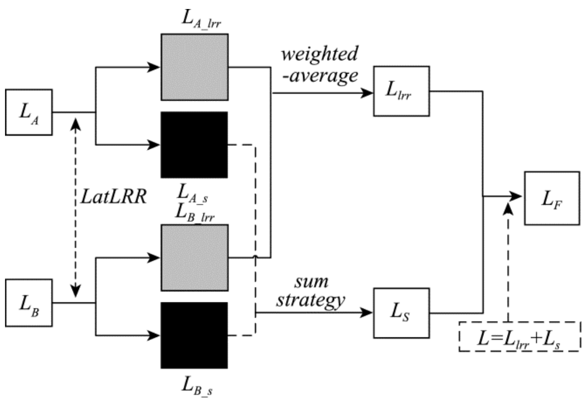


Fig.4 Low Frequency Subband Flow Chart Based on Potential Low Rank Representation

4 Experimental Results and Analysis

4.1 Image Fusion Quality Evaluation Index

In practical applications, image fusion quality evaluation methods are generally divided into two categories: subjective evaluation methods and objective evaluation methods. Subjective evaluation mainly relies on the subjective evaluation of the fusion image by the human visual system. It is feasible for some specific fusion scenes, but it has the disadvantages of strong subjectivity and one-sidedness. The objective evaluation method analyzes the performance of different fusion methods through objective indicators, and the evaluation results are more instructive. In this paper, we select 6 widely used objective fusion metrics.

(1) Spatial Frequency

spatial frequency（Spatial Frequency）[22] reflects the change rate of the image grayscale。The spatial frequency of the image is calculated from the row frequency (RF) and column frequency (CF) of the fused image. The higher the SF value, the higher the similarity between the fusion image and the input image, and the better the fusion quality. Its calculation formula is as follows:

$$SF=\sqrt{CF^2+RF^2} \tag{11}$$

Among this,

$$RF=\sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|F(i,j)-F(i,j-1)|^2} \tag{12}$$

$$CF=\sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|F(i,j)-F(i-1,j)|^2} \tag{13}$$

Among them, M and N represent the length and width of the image respectively, and F is the fused image.

(2) Average Gradient

Average Gradient is used to measure the clarity of the fused image. The larger the average gradient value of the fused image, the higher the definition and the better the quality of the fused image. Calculated as follows:

$$AG=\frac{1}{(M-1)(N-1)}\sum_{i=1}^{M-1}\sum_{i=1}^{N-1}\sqrt{\frac{\left(F(i+1,j)-F(i,j)\right)^2+\left(F(i,j+1)-F(i,j)\right)^2}{2}} \tag{14}$$

Parameters M, N, F have the same meaning as the parameters of the indicator SF.

(3) Entropy

ntropy[22] represents the amount of information in the fused image. The larger the entropy, the greater the amount of fusion image information. Its calculation formula is as follows:

$$E=\sum_{i=0}^{L-1}P_i\log_2 P_i \tag{15}$$

Among them, L represents the gray level of the image, and Pi represents the proportion of gray value i pixels to the total pixels.

(4) Quantization Parameter

Zhao et al.[23] proposed a method to calculate the availability of source image features in the fused image as an indicator to measure the efficiency of the fusion algorithm. The algorithm uses phase coherence to measure image features and uses local correlation calculations to quantify the availability of features in fusion results. The algorithm can be defined as:

$$P'_{blind}=\left(P_p\right)^{\alpha}\left(P_M\right)^{\beta}\left(P_m\right)^{\gamma} \tag{16}$$

$$C_{xy}^k=\frac{\sigma_{xy}^k+C_k}{\sigma_x^k\sigma_y^k+C_k} \tag{17}$$

From the source images A, B, the components Pp, PM, Pm can be obtained, which are defined by the maximum value of the correlation coefficient C_{xy}^k between the sets x and y. The exponent parameter can be adjusted according to the importance of the components α、β、γ. C_k is selected as a fixed constant 0.0001. Assuming that the image consists of K image blocks, the final result is:

$$P'_{blind} = \frac{1}{K}\sum_{k=1}^{K} P'_{blind}(k) \tag{18}$$

(5) Quality Assessment of Basic Functions

Quality Assessment of Basic Functions [22] is a novel fused image evaluation index, which uses local metrics to calculate the transfer amount of edge information injected into the fused image from the source image. The closer the value of QAB/F is to 1, the higher the quality of the fused image. Its calculation formula is as follows:

$$Q=\frac{1}{|w|}\sum_{\omega\in W}\left(\lambda(\omega)Q_0\left(A,F|\omega\right)+\left(1-\lambda(\omega)\right)Q_0\left(B,F|\omega\right)\right) \tag{19}$$

(6) Quality Evaluation (QE)

Based on the common image quality evaluation index (UIQI) [25], Piella and Heijman[24] proposed a UIQI-based fusion index Quality Evaluation[25],

$$Q=\frac{1}{|w|}\sum_{\omega\in W}\left(\lambda(\omega)Q_0\left(A,F|\omega\right)+\left(1-\lambda(\omega)\right)Q_0\left(B,F|\omega\right)\right) \tag{20}$$

$$Q_W=\sum_{\omega\in W}c(\omega)\left(\lambda(\omega)Q_0\left(A,F|\omega\right)+\left(1-\lambda(\omega)\right)Q_0\left(B,F|\omega\right)\right) \tag{21}$$

$$Q_E = Q_W\left(A,B,F\right)\cdot Q_W\left(A',B',F'\right)^{\alpha} \tag{22}$$

The central idea of Q is: first, the source image and the fused image are segmented using a sliding window; then, the structural similarity of each sub-block is calculated. Since each image block has different importance to the fusion quality, a QW weighted quality evaluation index is proposed to further improve the evaluation effect. Among them, $c(\omega)$ represents the importance of the image block at ω in the whole image. On the basis of QW, edge detection is performed on the fused image and the source image respectively, and then the edge image is brought into QW to obtain the weighted similarity, and the two QWs are combined to obtain the edge-based structural similarity index QE.

4.2 Comparison Scheme

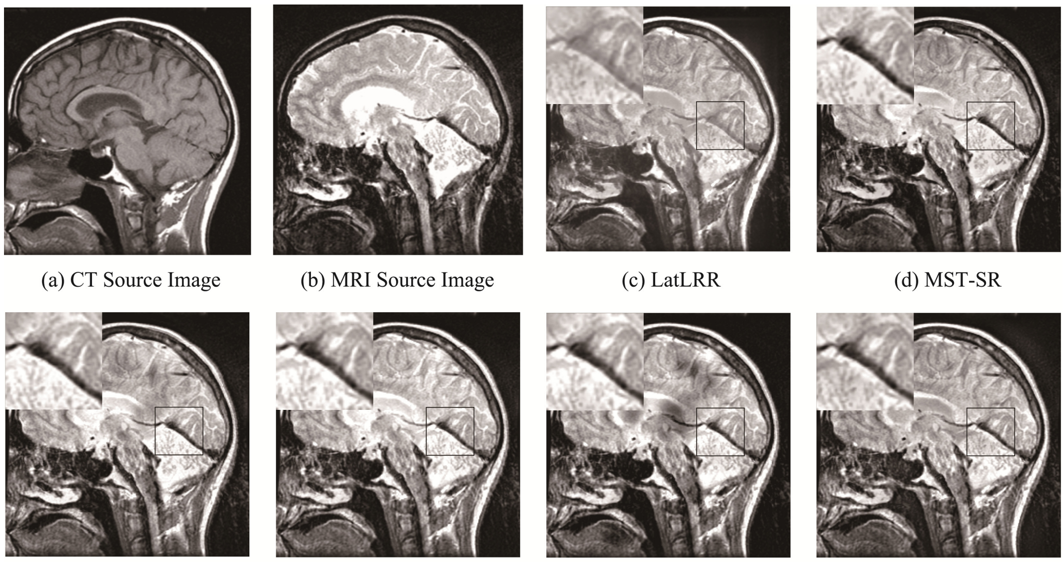
Compare the fusion method proposed in this paper with the existing five representative fusion methods, namely: LRR method[14], MST-SR method[26], NSCT-PCNNmethod[27], NSST-PAPCNN method [16], and NSST-PAPCNN-CSR method[28]. Among them, MST-SR, NSCT-PCNN, NSST-PAPCNN, and NSST-PAPCNN-CSR are all classic fusion schemes based on multi-scale transformation and have a similar fusion framework to the scheme in the paper. Other methods are also widely used in the field of image fusion. The parameter values used in these scenarios are the default values provided by the authors.

4.3 Image Fusion Experiment

The source images used in the experiment are all from the whole brain atlas library established by Harvard Medical School, and all the source images are obtained from the same slice with different imaging methods at the same angle. In order to verify the effectiveness of the algorithm in this paper, we selected 50 sets of brain source images for fusion testing. Among them, there are 20 pairs of CT/MR images and 30 pairs of MR-T1/MR-T2 images. All source images are of 256x256 pixel spatial resolution and were exactly matched before use. Select representative areas (such as image details, edges, etc.) of all fused images to zoom in as close-ups of fused images, to better evaluate the fused images visually. In this section, NSST-PAPCNN-LATLRR is compared comprehensively with five mainstream algorithms. Experiments prove that the algorithm in this paper is in the leading position in visual effect and objective index evaluation.

A. Visual Quality

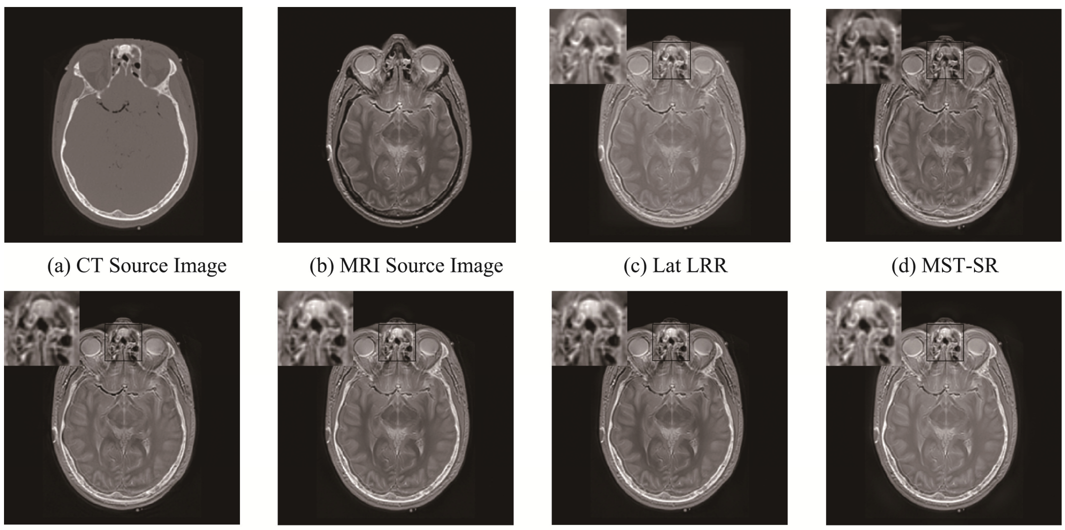
Fig.5 and Fig.6 show the fusion results of CT and MRI images, and the upper left corner is the enlarged image of local details. Compared with the source image, the LatLRR algorithm (Figure c) has insufficient ability to protect the detail information of the MRI source image, and the overall brightness of the fused image is higher. This is because the salient structure of the LatLRR algorithm adopts a simple "maximum value" fusion rule, and there is no suitable Reasonable selection of the weight of energy preservation; MST-SR algorithm (Figure d) adopts the mean fusion rule in low-frequency fusion, and has poor performance on details such as edges and contours, resulting in poor feature information fusion effect of the source image; NSCT -The high-frequency fusion part of the PCNN



(a) CT Source Image (b) MRI Source Image (c) LatLRR (d) MST-SR

(e) NSCT-PCNN (f) NSST-PAPCNN (g) NSST-PAPCNN-CSR (h) Proposed

Fig.5 CT and MR (1) Medical Image Fusion Results



(a) CT Source Image (b) MRI Source Image (c) Lat LRR (d) MST-SR

(e) NSCT-PCNN (f) NSST-PAPCNN (g) NSST-PAPCNN-CSR (h) Proposed

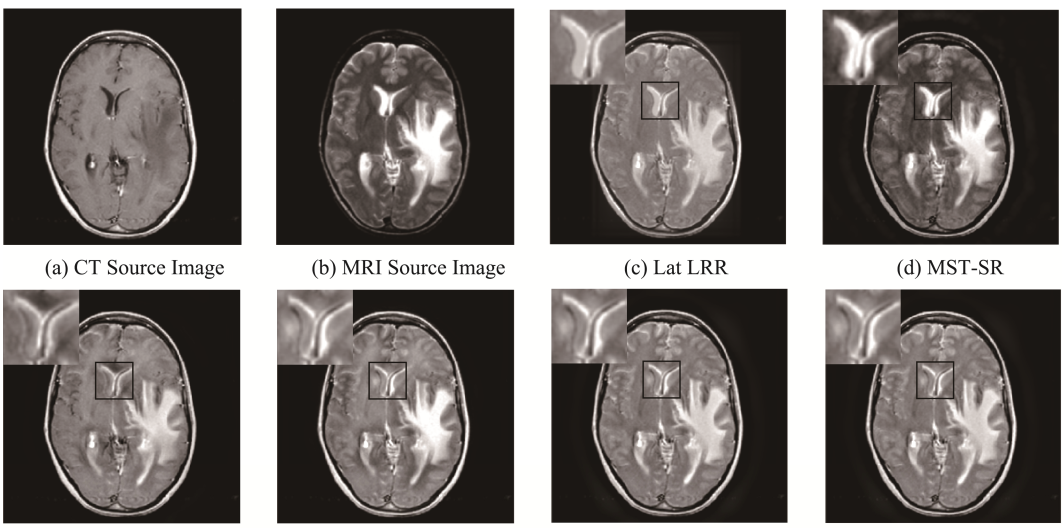
Fig.6 CT and MR (2) Medical Image Fusion Results.

algorithm (Figure e) adopts the traditional PCNN model, and there is a problem that the experimental results are difficult to achieve optimal results due to manual setting of parameters, so that the fusion effect on the image details is not good, and the energy conservation ability is limited, resulting in the overall brightness of the image Low problem; the NSST-PCNN algorithm (Figure f) can decompose the source image in multiple scales and directions, but the parameter adaptive PAPCNN model is used to process high-frequency coefficients, which usually performs well, but the function of preserving the source image There will be some deficiencies in information and image details; the NSST-PAPCNN-CSR algorithm (Figure g) chooses the global convolutional sparse model (CSR) model for its low-frequency algorithm, and lacks the details of low-frequency information, resulting in obvious defects in the fusion image. Artifacts and low overall brightness. The NSST-PAPCNN-LatLRR algorithm (Figure h) proposed in this paper fully retains the detailed information of the source image, and the low-frequency band also contains part of the detailed information. The LatLRR algorithm in the paper decomposes the low-frequency sub-band information twice, fully retaining the global image of the low-frequency segment. structure and salient information structure. Therefore, low-frequency information can well retain detailed information. The high-frequency sub-band information is fused through the improved PAPCNN model. In this paper, the optimal value of the target parameter is obtained through the improved parameter αe algorithm. The fused image can fully retain the significant details of the source image and there is no problem that the overall brightness of other algorithms is not suitable. It can also clearly express the outline, texture and edge features of the source image, without the problems of blurred or lost details and obvious noise artifacts.

Fig.7 and Fig.8 show the fusion results of two sets of MR-T1 and MR-T2 images. The overall brightness of LatLRR is low and there is an undesirable effect of blurring key information. In addition, some salient information (contour, etc.) in the MR-T1 source image is not well preserved in the fused image. The MST-SR algorithm retains too much information in the MR-T2 source image, and the detail information in MR-T1 is blurred or even lost, resulting in too low overall brightness of the fused image. The main disadvantages of the NSCT-PCNN algorithm are low computational efficiency and limited ability to extract details, resulting in obvious artifacts (see the local enlarged area in Figure 7e), and the overall fusion effect is not good. NSST-PAPCNN and NSST-PAPCNN-CSR perform well in detail extraction, but the NSST-PAPCNN-CSR algorithm over-enhances the information in the MR-T2 source image, and there is a problem of inconsistent intensity in some areas. The scheme proposed in this paper is more competitive than the comparison scheme, the overall visual quality is higher, and the details of the image can be well preserved. In the partially enlarged picture, the texture of the effect picture of this scheme is clearer, the salient features are more obvious, and there is no noise interference such as artifacts.

B. Objective Evaluation

Table 1 objectively evaluates the performance of different fusion algorithms for medical image fusion. Six index data of six algorithms corresponding to fusion images are given. For each metric, the highest scores of the six algorithms are indicated in bold, with the second and third ranked figures underlined. The scores of the top three ranks are indicated by numbers in parentheses. In order to more intuitively compare the objective indicators of different fusion schemes, we visualize the data content in Table 1 as shown in Fig.9.



(a) CT Source Image (b) MRI Source Image (c) Lat LRR (d) MST-SR

(e) NSCT-PCNN (f) NSST-PAPCNN (g) NSST-PAPCNN-CSR (h) Proposed

Fig.7 MR-T1 and MR-T2
(1) Medical Image Fusion Results.

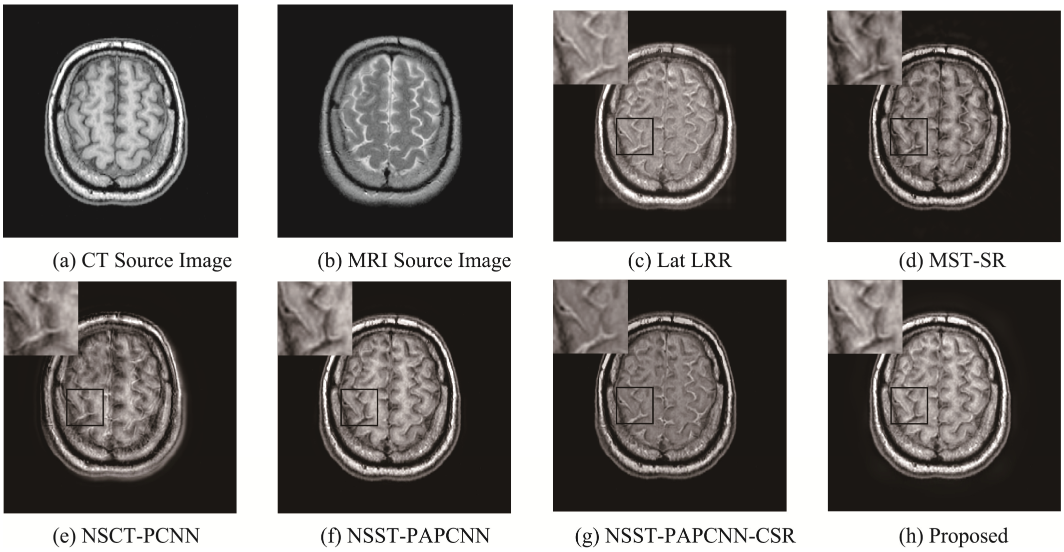
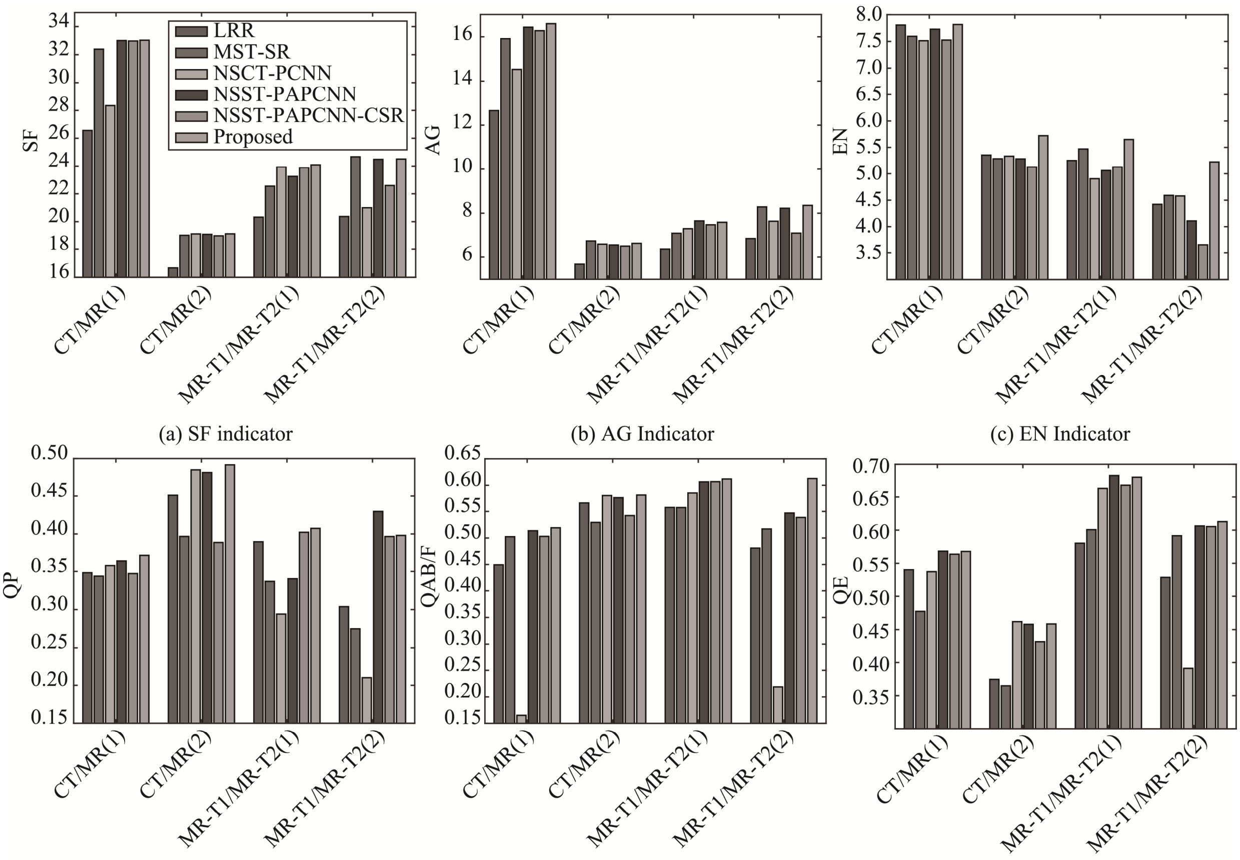


Fig.8 MR-T1 and MR-T2
(2) Medical Image Fusion Results.

Table 1 Objective Evaluation of Different Medical Image Fusion Methods

Images	Metrics	LatLRR	MST-SR	NSCT-PCNN	NSST-PAPCNN	NSST-PAPCNN-CSR	Proposed
CT/MR (1)	SF	26.562	32.3869	28.3453	33.0035②	32.9781③	33.0266①
	AG	12.6594	15.9146	14.5297	16.4285②	16.2706③	16.5895①
	EN	7.8073②	7.5956	7.5137	7.7279③	7.5252	7.8177①
	QP	0.3483	0.3439	0.3584③	0.3644②	0.3472	0.3718①
	$Q^{AB/F}$	0.4495	0.5023	0.1653	0.5136②	0.503③	0.5192①
	QE	0.5403	0.4773	0.5367	0.5683①	0.5637③	0.568②
CT/MR (2)	SF	16.6706	19.0223	19.1176②	19.0896③	18.9806	19.1285①
	AG	5.6809	6.7412①	6.594 ③	6.5648	6.5098	6.636②
	EN	5.3535②	5.2849	5.333 ③	5.2823	5.1292	5.721①
	QP	0.4511	0.3967	0.4851②	0.4816③	0.3888	0.4918①
	$Q^{AB/F}$	0.5661	0.5293	0.5799②	0.5759③	0.5421	0.5808①
	QE	0.3745	0.3648	0.462 ①	0.4578③	0.4318	0.4584②
MR-T1/ MR-T2 (1)	SF	20.3221	22.5456	23.9221	23.246	23.8766③	24.0841①
	AG	6.3508	7.0882	7.2972	7.6503①	7.4741③	7.5868②
	EN	5.2515③	5.4707	4.9045	5.0599	5.1294	5.6494①
	QP	0.3897③	0.337	0.2941	0.3405	0.4022②	0.4074①
	$Q^{AB/F}$	0.5575	0.5573	0.5848	0.6065③	0.6068②	0.612①
	QE	0.5803	0.6007	0.6635	0.6829①	0.6684③	0.6803②
MR-T1/ MR-T2 (2)	SF	20.3689	24.668①	20.9992	24.4838③	22.5898	24.5038②
	AG	6.8512	8.2826②	7.6368	8.2234③	7.095	8.3514①
	EN	4.4209	4.5898②	4.5801③	4.1088	3.657	5.2241①
	QP	0.3038	0.2748	0.21	0.4297①	0.3966③	0.3981②
	$Q^{AB/F}$	0.4809	0.5171	0.2188	0.5469②	0.5387③	0.5474①
	QE	0.5283	0.5915	0.3912	0.6063②	0.6053③	0.6128①



(a) SF indicator (b) AG Indicator (c) EN Indicator

(d) QP Indicator (e) $Q^{AB/F}$ Indicator (f) QE Indicator

Fig.9 Visualization of Objective Evaluation of Different Image Fusion Methods

Fig.9 shows, among the four groups of fused images, the six index parameters of the fusion scheme proposed in this paper are all in the top two. Among all the 6 fusion indicators, our scheme is the only one that always ranks in the top two in all indicators. Especially for EN and QAB/F metrics, our method achieves the highest scores. The proposed algorithm has the same fusion framework as the NSST-PAPCNN and NSST-PAPCNN-CSR algorithms, but the proposed method outperforms both of them in the four sets of fused images, which shows the clear advantage of our method. The MST-SR algorithm obtained a high score on the AG index of the two groups of photos, which is in line with the advantages of the sparse representation model, but the algorithm in this paper is in a leading position in other indexes. Based on the above comparison, it can be concluded that the algorithm in this paper can gain greater competitiveness in advanced image fusion schemes.

5 Conclusion

A new multi-scale transformation domain medical image fusion scheme is proposed in this paper. The innovation of this scheme mainly lies in two aspects. First, we introduce a parameter-adaptive PCNN model into the fusion of high-frequency subband coefficients. And optimize the parameter algorithm of the time decay factor αe to better coordinate the decay speed of the dynamic threshold. On the other hand, for the problem that low-frequency coefficients also contain some detailed information, the latent low-rank representation (LatLRR) algorithm is introduced as a fusion scheme for low-frequency sub-band coefficients. Solve the problem of difficult parameter setting and insufficient protection of detail ability in PCNN model. 50 groups of medical images were selected for a large number of experiments to verify the effectiveness of the method. Compared with five advanced medical image fusion schemes, the experimental results show that the fusion scheme in this paper is in the leading position in terms of visual perception and objective indicators. In the future, we will devote ourselves to developing more efficient image fusion strategies to further enhance the practical value of image fusion in clinical applications. In addition, we will continue to explore the potential of the NSST-PAPCNN-LatLRR algorithm in the fields of remote sensing images, infrared and visible light images.

References

[1] James, A.P. and Dasarathy, B.V. (2014) Medical image fusion: A survey of the state of the art, *Information Fusion*, 19, pp. 4-19.

[2] Du, J. et al. (2016) An overview of multi-modal medical image fusion, *Neurocomputing*, 215, pp. 3-20.

[3] Yang, B. and Li, S. (2012) Pixel-level image fusion with simultaneous orthogonal matching pursuit, *Information Fusion*, 13(1), pp. 10–19.

[4] Singh, R. and Khare, A. (2014) Of multimodal medical images using Daubechies complex wavelet transform—Amultiresolution approach. *Information Fusion*, 19, pp. 49–60.

[5] Bhatnagar, G., Wu, Q. M. J., and Liu, Z. (2015) A new contrast based multimodal medical image fusion framework, *Neurocomputing*, 157, pp. 143–152.

[6] Wang, Q. et al, (2015) Robust multi-modal medical image fusion via anisotropic heat diffusion guided low-rank structural analysis, *Information Fusion*, 26, pp. 103–121.

[7] Ding, W., Bi, D. and He, L. (2018) Infrared and visible image fusion based on sparse features, *Acta Photonica Sinica*, 47(9), pp. 1-10.

[8] Johnson, J. L. and Ritter, D. (1993) Observation of periodic waves in a pulse coupled neural network, *Opt. Lett.* 18(15), pp. 1253–1255.

[9] Johnson, J. L. (1994) Pulse-coupled neural nets: Translation, rotation, scale, distortion, and intensity signal invariance for images, *Appl. Opt.*, 33(26), pp. 6239–6253.

[10] Johnson, J. L., Padgett, M. L., and Omidvar, O. (1999) Guest editorial overview of pulse coupled neural network (PCNN) special issue, *IEEE Trans. Neural Netw*, 10(3), pp. 461–463.

[11] Zambotti-Villela S.C. et al (2008) Prospective evaluation of aminopeptidase activities in plasma and peripheral organs of streptozotocin-induced diabetic rats *Endocrinol. Invest*, 31, pp. 492-498.

[12] Chen, Y., Park, S.-K. and Ma, Y. (2011) A New Automatic Parameter Setting Method of a Simplified PCNN for Image Segmentation, *IEEE Transactions on Neural Networks*, 22(6), pp. 880-892.

[13] Yin, M., Liu, X. and Liu, Y. (2019) Medical Image Fusion with Parameter-Adaptive Pulse Coupled Neural Network in Nonsubsampled Shearlet Transform Domain, *IEEE Transactions on Instrumentation and Measurement*, 68(1), pp. 49-64.

[14] Liu, G. et al (2013) Robust Recovery of Subspace Structures by Low-Rank Representation, *IEEE Transactions on Pattern Analysis & Machine Intelligence.* 35(1), pp. 171-184.

[15] Li, H. and Wu, X.J. (2017) Multi-focus Image Fusion Using Dictionary Learning and Low-Rank Representation, *Lecture Notes in Computer Science*, 10666, pp. 675-686.

[16] Li, H. and Wu, X.J., Infrared and visible image fusion using Latent Low-Rank Representation, *arxiv*, 24 April 2018. Available at: https://arxiv.org/abs/1804.08992. (Accessed: 4 May 2023).

[17] Easley, G., Labate, D. and Lim, W.-Q. (2008) Sparse directional image representations using the discrete shearlet transform, *Appl. Comput. Harmon. Anal*, 25(1), pp. 25-46.

[18] Singh, S., Gupta, D. and Anand, R.S. (2015) Nonsubsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network, *Biomedical Signal Processing and Control*, 18, pp. 91-101.

[19] Chen, Y., Park, S.-K. and Ma, Y. (2011) A New Automatic Parameter Setting Method of a Simplified PCNN for Image Segmentation, *IEEE Transactions on Neural Networks*, 22(6), pp. 880-892.

[20] Guo Z. et al (2020) An adaptive infrared image segmentation method based on fusion SPCNN, *Signal Processing: Image Communication*, pp. 87(11), 115905.

[21] Liu, G., Yan, S. (2011) Latent low-rank representation for subspace segmentation and feature extraction, *International Conference on Computer Vision*, pp.

1615–1622.doi: 10.1016/j.image.2020.115905

[22] Jagalingam, P. and Hegde, A.V. (2015) A review of quality metrics for fused image, *Aquatic Procedia* ,4, pp. 133-142. doi: 10.1016/j.aqpro.2015.02.019

[23] Zhao, J., Laganiere, R. and Liu, Z. (2000) Performance Assessment of Combinative Pixel-Level Image Fusion Based on an Absolute Feature Measurement, *Innovative Computing, Information and Control,vol*. 3(6), pp. 1433-1447.

[24] Piella, G. and Heijmans, H. (2003) A new quality metric for image fusion, in: *Proceedings of 10th IEEE International Conference on Image Processing (ICIP)*, Barcelona, Spain, pp. 173–176.

[25] Wang, Z. and Bovik, A. (2002) A universal image quality index, *IEEE Signal Processing Letters*, 9(3), pp. 81-84.

[26] Liu, Y., Liu, S., and Wang, Z. (2015) A General Framework for Image Fusion Based on Multi-scale Transform and Sparse Representation, *Information Fusion*, 24(1), pp. 147-164.

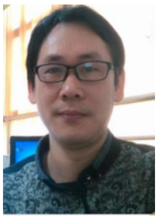
[27] Jia, Y.et al (2016) A novel image fusion algorithm using PCNN In NSCT domain, *International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, *(CISP-BMEI)*, Yichang, China, 9, pp. 751-755.

[28] Jingming Xia, Yi Lu, Ling Tan. (2020) Research of Multimodal Medical Image Fusion Based on Parameter-Adaptive Pulse-Coupled Neural Network and Convolutional Sparse Representation *Computational and Mathematical Methods in Medicine*, 3290136.

Author Biographies

WANG Wenyan is now a B.Sc. candidate at Nanjing University of Information Science and Technology. Her main research interest includes image processing.

E-mail: 107535193@qq.com



ZHOU Xianchun received Ph.D. from Nanjing University of Information Science and Technology (NUIST) in 2003. He is now a NUIST professor, master supervisor and IEEE Member. His main research interests include image processing, pattern recognition, and intelligent signal processing.

E-mail: zhouxc2008@163.com