Medical Image Fusion With Parameter-Adaptive Pulse Coupled-Neural Network in Nonsubsampled Shearlet Transform Domain

Ming Yin, Xiaoning Liu, Yu Liu, *Member, IEEE*, and Xun Chen, *Member, IEEE*

***Abstract*— As an effective way to integrate the information contained in multiple medical images with different modalities, medical image fusion has emerged as a powerful technique in various clinical applications such as disease diagnosis and treatment planning. In this paper, a new multimodal medical image fusion method in nonsubsampled shearlet transform (NSST) domain is proposed. In the proposed method, the NSST decomposition is first performed on the source images to obtain their multiscale and multidirection representations. The high-frequency bands are fused by a parameter-adaptive pulse-coupled neural network (PA-PCNN) model, in which all the PCNN parameters can be adaptively estimated by the input band. The low-frequency bands are merged by a novel strategy that simultaneously addresses two crucial issues in medical image fusion, namely, energy preservation and detail extraction. Finally, the fused image is reconstructed by performing inverse NSST on the fused high-frequency and low-frequency bands. The effectiveness of the proposed method is verified by four different categories of medical image fusion problems [computed tomography (CT) and magnetic resonance (MR), MR-T1 and MR-T2, MR and positron emission tomography, and MR and single-photon emission CT] with more than 80 pairs of source images in total. Experimental results demonstrate that the proposed method can obtain more competitive performance in comparison to nine representative medical image fusion methods, leading to state-of-the-art results on both visual quality and objective assessment.**

***Index Terms*— Activity level measure, image fusion, medical imaging, nonsubsampled shearlet transform (NSST), pulse coupled neural network (PCNN).**

I. Introduction

As is well known, medical imaging is acting as an increasingly critical role in various clinical applications such as diagnosis, treatment planning, and surgical navigation. Due to the diversity in imaging mechanisms, medical images with different modalities focus on different categories of organ/tissue information. The computed tomography (CT) imaging can precisely detect dense structures such as bones and implants. The magnetic resonance (MR) imaging provides high-resolution anatomical information for soft tissues, but is less sensitive to the diagnosis of fractures than CT. In addition to these anatomical imaging techniques, the functional imaging techniques such as positron emission tomography (PET) and single-photon emission CT (SPECT) are often applied to reflect the metabolism information of organism, which is of great significance to many scenarios such as vascular disease diagnosis and tumor detection. Nevertheless, the spatial resolution of functional images is usually very low. To obtain sufficient information for accurate diagnosis, physicians often need to sequentially analyze medical images that are captured with different modalities, but this separating manner may still bring inconvenience in many cases. An effective way to solve this problem is known as medical image fusion technique [1], [2], which aims at generating a composite image to integrate the complementary information contained in multiple medical images with different modalities.

A variety of medical image fusion methods have been proposed over the past decades [3]–[13]. Since there is strong evidence that the human visual system (HVS) processes information in a multiresolution fashion [14], most medical image fusion methods are introduced under a multiscale transform (MST)-based framework to pursue perceptually good results. In general, the MST-based fusion methods consist of three basic steps. First, the source images are converted into an MST domain. Then, the transformed coefficients are merged using some predesigned fusion strategies. Finally, the fused image is reconstructed from the merged coefficients by performing the inverse transform. MST approaches that are commonly used in image fusion include pyramid-based ones (e.g., Laplacian pyramid (LP) [15] and morphological pyramid [16]), wavelet-based ones (e.g., discrete wavelet transform [17] and dual-tree complex wavelet transform [18]), and multiscale geometric analysis (MGA)-based ones (e.g., nonsubsampled contourlet transform (NSCT) [19] and nonsubsampled shearlet transform (NSST) [20]). Among them, the MGA-based methods, especially for the NSCT- and NSST-based methods, have exhibited significant advantages over other methods on account of their higher effectiveness in image representation. In addition to the selection of image

Manuscript received December 9, 2017; revised April 17, 2018; accepted May 5, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61701160, Grant 11601115, Grant 81571760, and Grant 61501164, in part by the Provincial Natural Science Foundation of Anhui under Grant 1808085QF186, and in part by the Fundamental Research Funds for the Central Universities under Grant JZ2018HGTB0228 and Grant JZ2016HGBZ1025. The Associate Editor coordinating the review process was Anirban Mukherjee. *(Corresponding author: Yu Liu.)*

M. Yin and X. Liu are with the School of Mathematics, Hefei University of Technology, Hefei 230009, China.

Y. Liu is with the Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China (e-mail: yuliu@hfut.edu.cn).

X. Chen is with the Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230026, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2018.2838778

transform, the design of fusion strategies for both the high-frequency and low-frequency coefficients is another crucial issue in MST-based fusion methods. Traditionally, the activity level of high-frequency coefficients is usually calculated based on their absolute values using a pixel-based or window-based manner, and then a simple fusion rule such as choose max or weighted average is applied to obtain the fused coefficients. The most popular low-frequency fusion strategy in the early days is just averaging the coefficients from different source images. Plenty of studies in the literature indicate that the performances of the MST-based methods could be significantly improved by designing more effective fusion strategies. In recent years, many novel MST-based medical image fusion methods have been proposed by developing more advanced image transforms and more complicated fusion strategies [21]–[30]. Here, we just give a few latest examples. Zhu *et al.* [28] proposed a medical image fusion method based on cartoon-texture decomposition (CTD) and applied a sparse representation (SR)-based fusion strategy [31] to merge the decomposed coefficients. Du *et al.* [29] proposed a multiscale decomposition approach based on local Laplacian filtering (LLF) for medical image fusion and introduced an information of interest (IOI)-based strategy to fuse the high-frequency components. Liu *et al.* [30] proposed a convolutional neural network (CNN)-based fusion strategy [32] under the LP-based framework for medical image fusion.

Pulse coupled neural network (PCNN) [33], a biologically inspired neural network derived from Echorn's [34] cortical model that owns characteristics including global coupling and pulse synchronization, has been verified a very suitable tool for image fusion. Specifically, PCNN is often used to extract the activity level of the decomposed coefficients obtained by a certain MST for image fusion. The firing times of each output neuron over a number of iterations are typically employed to measure the activity level of its corresponding coefficient. A popular PCNN-based fusion strategy for the NSCT-based image fusion was introduced by Qu *et al.* [21]. Das and Kundu [23] proposed a fuzzy-adaptive reduced PCNN (RPCNN)-based fusion strategy in NSCT domain for medical image fusion. Due to the strong association with the information processing mechanism of HVS, the PCNN-based fusion strategies have shown advantages over the conventional ones that are directly based on the decomposed coefficients. However, PCNN models generally contain a series of free parameters and the quality of fusion results relies heavily on these parameter settings. In most cases, the PCNN parameters are manually fixed as constants in an algorithm based on empirical or experimental results, which may limit the algorithm performance to a great extent.

In this paper, a multimodality medical image fusion method based on NSST is proposed. We mainly focus on the design of fusion strategies for both the high-frequency and low-frequency coefficients. The main contributions of this paper are outlined as follows.

1) We introduce a parameter-adaptive PCNN (PA-PCNN) model [35] into the field of image fusion. The PA-PCNN model is adopted to fuse high-frequency coefficients with all the PCNN parameters adaptively calculated based on the input bands, which can overcome the difficulty of setting free parameters in the conventional PCNN models. In addition, the PA-PCNN is experimentally verified to have a fast convergence speed with fewer iterations than some commonly used PCNN models in image fusion. To the best of our knowledge, this is the first time that the PA-PCNN model [35] is applied to image fusion.
2) We present a novel low-frequency fusion strategy that simultaneously addresses two crucial factors in medical image fusion, namely, energy preservation and detail extraction. To this end, two new activity level measures named weighted local energy (WLE) and weighted sum of eight-neighborhood-based modified Laplacian (WSEML) are defined in this paper, respectively.
3) We propose a new medical image fusion method in the NSST domain by applying the fusion strategies mentioned earlier. Extensive experiments are conducted to verify the effectiveness of our method on four different types of medical image fusion problems (CT and MR, MR-T1 and MR-T2, MR and PET, and MR and SPECT) with more than 80 pairs of source images. Nine representative medical image fusion methods are used for comparison and several of them were proposed very recently. Experimental results demonstrate that the proposed method can achieve state-of-the-art performances on both the visual quality and objective assessment.

The rest of this paper is organized as follows. In Section II, the theory of NSST is briefly introduced. Section III presents the adopted PA-PCNN model and discusses its feasibility for image fusion. The detailed NSST-based fusion scheme is presented in Section IV. Section V gives the experimental results and discussion. Finally, this paper is concluded in Section VI.

II. Nonsubsampled Shearlet Transform

Shearlet [36] is a relatively new member in the family of MGA. In comparison to some earlier multiscale approaches for image representation such as pyramid, wavelet, and curvelet, shearlet can capture the details/features of an image at diverse directions more effectively and is able to obtain a more optimal representation (often measured by the sparsity) for the targeting image. The implementation process of shearlet transform (ST) is similar to that of contourlet transform, but the directional filters in contourlet are replaced by the shearing filters. An important advantage of shearlet over contourlet is that there are no restrictions on the number of directions in shearlet. Moreover, the inverse ST only needs a summation of the shearing filters, instead of inverting a directional filter bank in contourlet, which improves the computational efficiency.

Despite of the above-mentioned advantages, the subsampling scheme in the standard ST causes its lack of shift invariance, which is a very critical property for image fusion to prevent undesirable Gibbs phenomenon. Having shift invariance, the related image fusion algorithm can be more robust

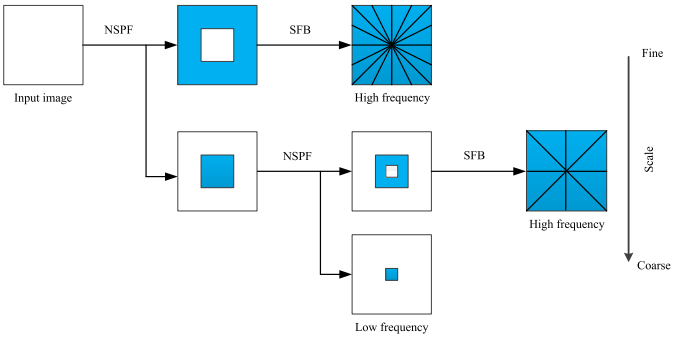


Fig. 1. Schematic of a two-level NSST decomposition.

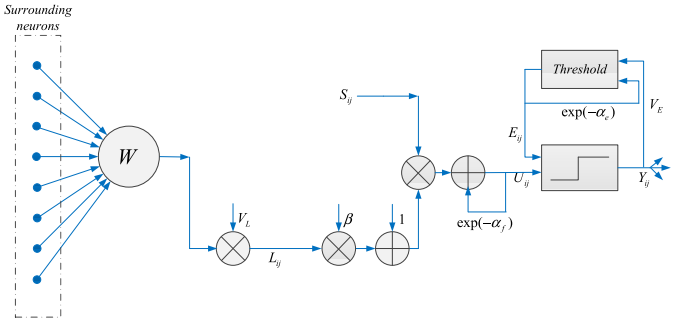


Fig. 2. Architecture of the PA-PCNN model used in the proposed image fusion method.

to misregistration[1] by developing region-based fusion strategies [13]. To overcome this defect, the nonsubsampled version of ST, known as NSST, has been introduced based on nonsubsampled pyramid filters (NSPFs) and shift-invariant shearlet filter banks (SFBs) [36]. Specifically, the NSPF is employed to obtain the multiscale decomposition of an input image from fine to coarse. If a L-level decomposition is applied, we can get $L+1$ subbands of the same size as the input image, which includes L high-frequency bands and one low-frequency band. For each decomposition level, an SFB is applied to obtain the multidirectional representations of the corresponding band. Fig. 1 shows the schematic of a two-level NSST decomposition. The inverse NSST can precisely reconstruct the original image with two steps. The first step is to generate the nonsubsampled pyramid, in which each scale is obtained by accumulating the filtered results of the decompositions at all directions using shearing filter banks. The second step is to reconstruct the image by the obtained nonsubsampled pyramid from coarse to fine using the reconstruction filters. More details about NSST could be found in [36]. Due to the properties such as multiscale, multidirection, and shift invariance, NSST has been recognized as a very effective approach for image fusion [37]–[39]. Also, it is selected as the MST for proposed fusion method in this paper.

III. Parameter-Adaptive PCNN

Unlike the most existing artificial neural networks, PCNN is based on iterative calculation and does not require any training process. The PCNN model applied in image processing tasks is generally a single-layer network with a 2-D array input. There is a one-to-one correspondence between input image pixels and PCNN neurons, so the number of neurons is equal to that of pixels. Each neuron is linked with its neighboring neurons for information transmission and coupling. As mentioned earlier, one key challenge that exists in conventional PCNN models is the setting of several free parameters such as linking strength, various amplitudes, and decay coefficients [33], [40]. In order to avoid the difficulty in manually setting these parameters, Chen *et al.* [35] recently proposed a simplified PCNN (SPCNN) model along with an automatic parameter setting approach for image segmentation. The SPCNN model is described as follows:

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

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

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

$$Y_{ij}[n] = \begin{cases} 1, & \text{if } U_{ij}[n] > E_{ij}[n-1] \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

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

In the above-mentioned SPCNN model, $F_{ij}[n]$ and $L_{ij}[n]$ are feeding input and linking input of the neuron at position (i, j) in iteration n, respectively. $F_{ij}[n]$ is fixed to the intensity of input image S_{ij} during the whole iteration. $L_{ij}[n]$ is associated with the previous firing status of eight neighboring neurons through the synaptic weights

$$W_{ijkl} = \begin{bmatrix} 0.5 & 1 & 0.5 \\ 1 & 0 & 1 \\ 0.5 & 1 & 0.5 \end{bmatrix}.$$

The parameter V_L is the amplitude of linking input. The internal activity $U_{ij}[n]$ consists of two terms. The first term $e^{-\alpha_f} U_{ij}[n-1]$ is a decay of its previous value, where the parameter α_f is an exponential decay coefficient. The second term $F_{ij}[n](1+\beta L_{ij}[n])$ is the nonlinear modulation of $L_{ij}[n]$ and $F_{ij}[n]$, where the parameter β is the linking strength. The output module of the PCNN is known as *pulse generator* [33], which determines the firing events of the model. In particular, its output $Y_{ij}[n]$ has two status: fired ($Y_{ij}[n]=1$) and unfired ($Y_{ij}[n]=0$). The status depends on its two inputs, namely, current internal activity $U_{ij}[n]$ and previous dynamic threshold $E_{ij}[n-1]$. As given in (4), the firing condition is that $U_{ij}[n]$ is larger than $E_{ij}[n-1]$. Please refer to [33] for more details about firing. The final step in the iteration is updating the dynamic threshold according to (5), where α_e and V_E are the exponential decay coefficient and amplitude of $E_{ij}[n]$, respectively. The SPCNN model is initialized as $Y_{ij}[0]=0$, $U_{ij}[0]=0$ and $E_{ij}[0]=0$. As a result, all the nonzero intensity neurons will fire in the first iteration, because the firing condition is always valid ($U_{ij}[1]=S_{ij}>0$). Fig. 2 shows the architecture of the SPCNN model.

[1]In the field of pixel-level image fusion, a common assumption is that the source images have been accurately preregistered by some image registration approaches. The term *misregistration* here indicates some slight misalignments between source images due to factors such as slight motion of objects or the performance limitation of the applied image registration method. Large degree of misregistration goes outside the scope of most image fusion papers.

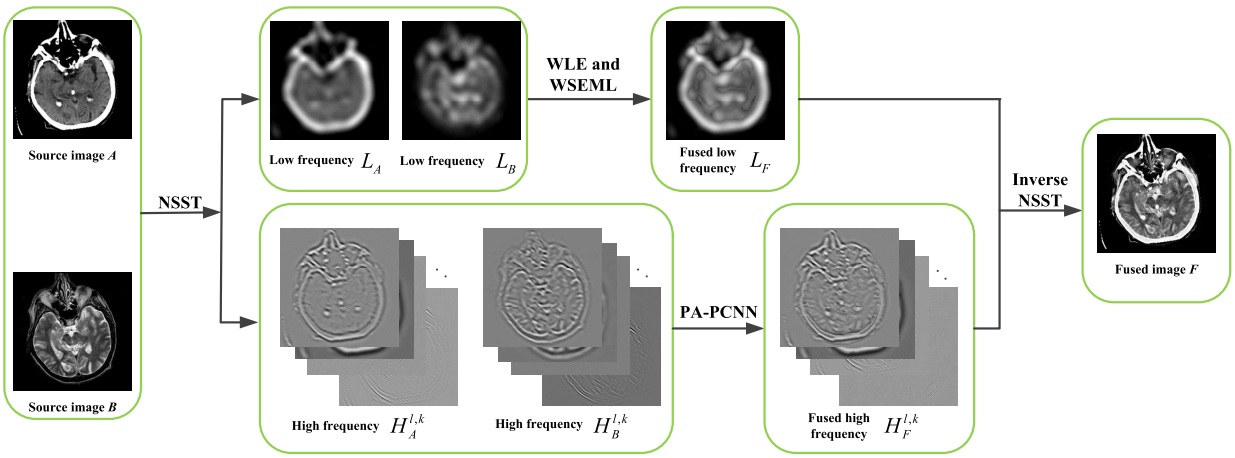


Fig. 3. Schematic of the proposed medical image fusion method.

There exist five free parameters in the SPCNN model: α_f, β, V_L, α_e, and V_E. Moreover, it can be observed from (1) to (5) that either β or V_L just acts as the weight of $\sum_{kl} W_{ijkl} Y_{kl}[n-1]$, so they can be tackled as a whole βV_L in the SPCNN model. Let $\lambda = \beta V_L$ denote the weighted linking strength, thus there are actually four parameters in SPCNN. According to the analysis in [35], all of these parameters can be adaptively calculated as

$$\alpha_f = \log(1/\sigma\ (S)) \tag{6}$$

$$\lambda = \frac{(S_{\max}/S') - 1}{6} \tag{7}$$

$$V_E = e^{-\alpha_f} + 1 + 6\lambda \tag{8}$$

$$\alpha_e = \ln\left(\frac{\frac{V_E}{S'}}{\frac{1-e^{-3\alpha_f}}{1-e^{-\alpha_f}} + 6\lambda e^{-\alpha_f}}\right) \tag{9}$$

where $\sigma\ (S)$ indicates the standard deviation of the input image S of range [0, 1]. S' and $S_{\max}$ denote the normalized Otsu threshold and the maximum intensity of the input image, respectively. Please refer to [35] for more details about the parameter settings.

The above-mentioned PA-PCNN model is primarily proposed for image segmentation in [35]. In this paper, we argue that it is also effective for image fusion. In particular, it is reasonable to apply the PA-PCNN model to fuse the high-frequency coefficients obtained by an MST. Actually, in the SPCNN-based image segmentation method [35], the target image is used as the PCNN input and the segmentation principle is essentially based on a pixel intensity. The role of PCNN in image segmentation is to transmit the intensity information from one pixel to its surrounding pixels, so that the information of neighboring pixels can be coupled and comprehensively utilized. For the fusion of high-frequency MST coefficients, the absolute values of coefficients are widely recognized as containing the activity level information, and a higher value is more likely to indicate a higher activity level among multiple source images. Considering that the PCNN model can tackle the pixel intensity for segmentation, it is also able to distinguish the absolute values of high-frequency coefficients from different source images. The advantage of PCNN in image fusion is similar to that in segmentation, namely, the activity level information of neighboring pixels can be transmitted to each other to achieve a more robust activity measure. Based on the above-mentioned considerations, it can be found that the problem of PCNN-based image fusion has a strong correlation with the PCNN-based image segmentation, which motivates us to introduce the above-mentioned PA-PCNN model into the fusion of high-frequency MST coefficients.

IV. Proposed Fusion Method

A. Overview

Fig. 3 shows the schematic of the proposed medical image fusion method. For the sake of simplicity, the fusion scheme presented in this section is based on the assumption that there are two preregistered source images to be fused. The proposed fusion method can be straightforwardly extended to fuse more than two source images. The detailed fusion scheme consists of four steps: NSST decomposition, fusion of high-frequency bands, fusion of low-frequency bands, and NSST reconstruction.

B. Detailed Fusion Scheme

1) NSST Decomposition: A L-level NSST decomposition is performed on two source images A and B to obtain their decomposed bands $\{H_A^{l,k}, L_A\}$ and $\{H_B^{l,k}, L_B\}$, respectively. The notation $H_A^{l,k}$ denotes a high-frequency band of A at decomposition level l and direction k, while the notation L_A denotes the low-frequency band of A. The meanings of $H_B^{l,k}$ and L_B are the same with respect to B.

2) Fusion of High-Frequency Bands: The PA-PCNN model presented in Section III is applied to fuse the high-frequency bands. Based on the discussion in Section III, the absolute value map of a high-frequency band is employed as the network input, namely, the feeding input is $F_{ij}[n] = |H_S^{l,k}|, S \in \{A, B\}$. The activity level of a high-frequency coefficient is measured by the total firing times during the whole iteration. According to the PA-PCNN model described

from (1) to (5), the firing times can be accumulated by adding the following step at the end of each iteration:

$$T_{ij}[n] = T_{ij}[n-1] + Y_{ij}[n]. \tag{10}$$

Thus, the firing times of each neuron is $T_{ij}[N]$, where N is the total number of iterations. For the corresponding high-frequency bands $H_A^{l,k}$ and $H_B^{l,k}$, their PA-PCNN firing times can be calculated and denoted by $T_{A,ij}^{l,k}[N]$ and $T_{B,ij}^{l,k}[N]$, respectively. The fused band is obtained using the following rule:

$$H_F^{l,k}(i,j) = \begin{cases} H_A^{l,k}(i,j), & \text{if } T_{A,ij}^{l,k}[N] \geq T_{B,ij}^{l,k}[N] \\ H_B^{l,k}(i,j), & \text{otherwise} \end{cases} \tag{11}$$

which indicates that the coefficient with larger firing times is selected as the fused coefficient.

3) Fusion of Low-Frequency Bands: The fusion strategy for low-frequency bands also has a significant impact on the final fusion quality. In our method, a strategy that simultaneously addresses two crucial factors (energy preservation and detail extraction) in medical image fusion is designed.

Since an image can be generally viewed as a 2-D piecewise smooth signal [36], its energy is mostly contained in its low-frequency component.[2] In medical image fusion, the intensities of different source images at the same location may vary significantly, because the source images are captured with different imaging mechanisms. Therefore, the conventional averaging-based low-frequency fusion rule tends to cause the loss of energy in the fused image [8]. As a result, the brightness of some regions may have a sharp decrease, leading to inferior visual perception. To address this issue, we define an activity level measure called WLE as

$$\begin{aligned} \mathrm{WLE}_S(i,j) = &\sum_{m=-r}^{r}\sum_{n=-r}^{r} W \\ &\times(m+r+1, n+r+1)\ L_S(i+m, j+n)^2 \end{aligned} \tag{12}$$

where $S \in \{A, B\}$ and W is a $(2r+1) \times (2r+1)$ weighting matrix with radius r. For each element in W, its value is set to 2^{2r-d}, where d is its four-neighborhood distance to the center. As an example, the 3×3 normalized version of W is

$$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

Due to some factors (e.g., computational efficiency), there always exists a limitation on the number of NSST decomposition level. Accordingly, the low-frequency band still contains some detail information. To fully extract the details from source images, an activity level measure named weighted sum of WSEML is defined as

$$\begin{aligned} \mathrm{WSEML}_S(i,j) = &\sum_{m=-r}^{r}\sum_{n=-r}^{r} W(m+r+1, n+r+1) \\ &\times \mathrm{EML}_S(i+m, j+n) \end{aligned} \tag{13}$$

[2]The term *energy* here originates from Parseval's Theorem (also known as Rayleigh Energy Theorem), although the theorem is not strictly valid for NSST decomposition because of its redundancy. Specifically, the energy of a band is defined as the sum of squares of all the band coefficients.

where $S \in \{A, B\}$, W is the weighting matrix that has the same definition as that in (12), and the EML is the WSEML

$$\begin{aligned} &\mathrm{EML}_S(i,j) \\ &= |2S(i,j) - S(i-1,j) - S(i+1,j)| \\ &\quad + |2S(i,j) - S(i,j-1) - S(i,j+1)| \\ &\quad + \frac{1}{\sqrt{2}}|2S(i,j) - S(i-1,j-1) - S(i+1,j+1)| \\ &\quad + \frac{1}{\sqrt{2}}|2S(i,j) - S(i-1,j+1) - S(i+1,j-1)|. \end{aligned} \tag{14}$$

In comparison to the original modified Laplacian (ML) [41], the EML takes the effect of diagonal coefficients into account to make full use of neighboring information. The weights of two diagonal terms are set to $1/\sqrt{2}$ according to the Euclidean distance.

The final activity level measure of low-frequency band is defined as the multiplication of WLE and WSEML, so the low-frequency fused band is calculated by the following rule:

$$L_F(i,j) = \begin{cases} L_A(i,j), & \text{if } \mathrm{WLE}_A(i,j) \cdot \mathrm{WSEML}_A(i,j) \\ & \quad \geq \mathrm{WLE}_B(i,j) \cdot \mathrm{WSEML}_B(i,j) \\ L_B(i,j), & \text{otherwise.} \end{cases} \tag{15}$$

4) NSST Reconstruction: The fused image F is finally reconstructed by performing the inverse NSST over the fused bands $\{H_F^{l,k}, L_F\}$.

The main steps of the proposed medical image fusion method are summarized in Algorithm 1.

C. Extension to Grayscale and Color Image Fusion

This section extends the proposed method to fuse a grayscale image and a color image, which is the typical situation of anatomical and functional image fusion. In medical imaging, the functional images such as PET and SPECT images are usually shown as pseudocolor images, and they can be treated as color images with RGB channels in the fusion process [6], [27], [30], [42]. One possible way for grayscale and color image fusion is merging the grayscale image with each channel of the color image independently and then combining the three fused channels to construct an RGB image, but this may cause serious color distortion. A more effective approach is to separate the brightness or luminance component from the color image for fusion via some color space transform methods. In this paper, the YUV color space is applied to accomplish the grayscale and color image fusion issues. The YUV space encodes a color image into one luminance component (Y) and two chrominance components (U and V) taking human perception into account, which makes it a popular approach in color image pipeline. In particular, it has been verified a very effective tool for anatomical and functional image fusion [27], [30]. Specifically, the fusion scheme contains the following three steps. First, the RGB color image is converted into YUV color space to obtain the Y, U, and V channels. Then, the grayscale image and the Y channel are fused using the proposed fusion scheme described in Section IV-B. Finally, the fused color image is obtained by

Algorithm 1 Proposed Medical Image Fusion Algorithm

Input: the source images: A and B.
Parameters: the number of NSST decomposition levels: L, the number of directions at each decomposition level: $K(l), l \in [1, L]$, the number of PA-PCNN iterations: N, the radius of weighting kernel: r
Part 1: NSST decomposition
For each source image $S = [A, B]$
    Perform NSST decomposition on S to obtain $\{H_S^{l,k}, L_S\}$, $l \in [1, L]$, $k \in [1, K(l)]$;
End
Part 2: Fusion of high-frequency bands
For each level $l = 1 : L$
    For each direction $k = 1 : K(l)$
        For each source image $S = [A, B]$
            Initialize the PA-PCNN model: $Y_{ij}[0] = 0$, $U_{ij}[0] = 0$, $E_{ij}[0] = 0$, $T_{ij}[0] = 0$ and $F_{ij}[n] = \left|H_S^{l,k}\right|$, $n \in [1, N]$;
            Estimate the PA-PCNN parameters using Eq. (6)–(9);
            For each iteration $n = 1 : N$
                Calculate the PA-PCNN model by Eq. (2)–(5) and Eq. (10);
            End
        End
        Merge $H_A^{l,k}$ and $H_B^{l,k}$ using Eq. (11) to obtain $H_F^{l,k}$;
    End
End
Part 3: Fusion of low-frequency bands
For each source image $S = [A, B]$
    Calculate the WLE for L_S using Eq. (12);
    Calculate the WSEML for L_S using Eq. (13)-(14);
End
Merge L_A and L_B using Eq. (15) to obtain L_F;
Part 4: NSST reconstruction
Perform inverse NSST on $\{H_F^{l,k}, L_F\}$ to obtain F;
Output: the fused image F.

performing inverse YUV conversion (YUV to RGB) over the fused Y channel, the original U channel, and the original V channel. Fig. 4 shows the schematic of the grayscale and color image fusion method.

V. Experiments

A. Experimental Settings

1) Source Images: To verify the effectiveness of the proposed method, 83 pairs of multimodal medical images, which include 10 pairs of MR and CT images, 13 pairs of T1-weighted MR (MR-T1) and T2-weighted (MR-T2) images, 30 pairs of MR and PET images, and 30 pairs of MR and SPECT images, are used in our experiments. All of these source images are collected from the database of Whole Brain Atlas [43] created by Harvard Medical School and have been widely adopted in previous publications related to medical image fusion. All the source images have the same spatial resolution of 256×256 pixels. The source images in each pair have been accurately registered.

2) Methods for Comparison: The proposed fusion method is compared to nine representative methods that are: the NSCT-SF-PCNN method [21], the SR with simultaneous orthogonal matching pursuit (SR-SOMP) method [4], the guided filtering (GF) method [11], the phase congruency and directive contrast in NSCT domain (NSCT-PCDC) method [22], the NSCT-RPCNN method [23], the LP-SR method [8], the CTD-SR method [28], the LLF-IOI method [29], and the LP-CNN method [30]. Among them, the CTD-SR, LLF-IOI, and LP-CNN methods were just recently proposed within 1 year. The NSCT-SF-PCNN and NSCT-RPCNN are two well-known MST-based methods using PCNN-based fusion strategies, which have the similar framework of the proposed method. The other methods are also of high popularity in the past few years. The source codes of the NSCT-SF-PCNN, GF, NSCT-PCDC, NSCT-RPCNN, LP-SR, and LLF-IOI methods are publicly available online at websites [44]–[49], respectively. The implementations of the other three methods are provided by their respective authors. All the parameters in these compared methods are set to the default values given by their authors.

3) Objective Evaluation Metrics: To quantitatively assess the performances of different methods, five widely recognized objective fusion metrics are applied in our experiments. They are the standard deviation (SD), entropy (EN), localized mutual information (LMI) [50], Piella's structure similarity-based metric Q_W [51], and the human visual perception-based metric visual information fidelity fusion (VIFF) [52]. SD measures the overall contrast of the fused image while the amount of information contained in the fused image can be reflected by the EN. LMI is a localized variation of

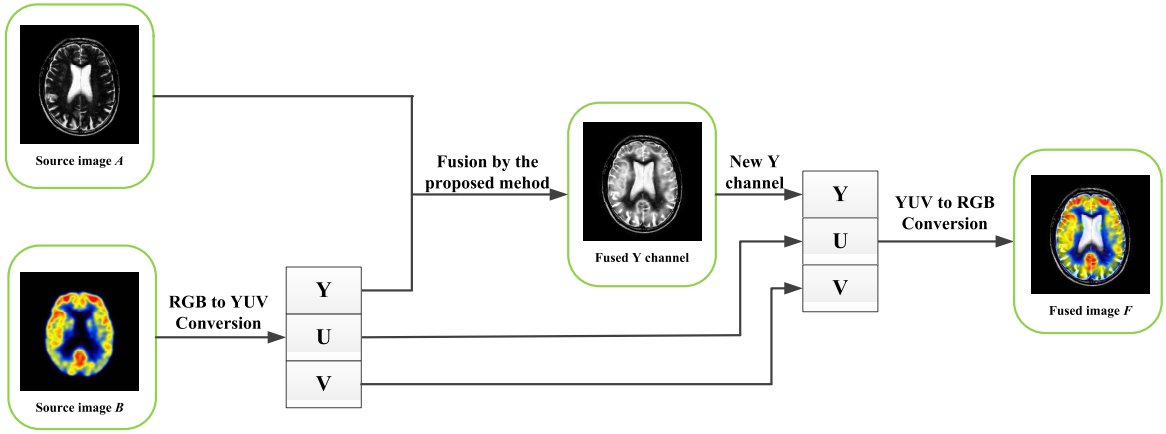


Fig. 4. Schematic of the grayscale and color image fusion method.

conventional mutual information (MI) index based on quadtree decomposition, and it can overcome several defects of MI when evaluating fusion performance. The metric Q_W mainly values the structure similarity between the fused images and source images by simultaneously addressing coefficient correlation, illumination, and contrast. It also takes image salience into account during the calculation. VIFF is a newly proposed metric by measuring the visual information fidelity between the fused image and each source image based on the Gaussian scale mixture model, the distortion model, and the HVS model. For all the above five metrics, a larger score indicates a better performance. Please refer to the related references for more details about these metrics.

4) Experimental Environment: All the experiments are conducted on a PC that equipped with an Intel(R) Core(TM) i7-6700 K CPU (4.00 GHz) and 32-GB RAM. The software environment is MATLAB R2013b installed on Win 7 64-bit operating system.

B. Analysis of Algorithm Parameters

The proposed fusion method has three main free parameters: 1) the number of NSST decomposition levels[3] L; 2) the number of PA-PCNN iterations N; and 3) the radius r of weighting kernel used in WLE and WSEML. In this section, we analyze the impacts of these parameters on fusion performance via the above-mentioned five objective fusion metrics. For each metric, the average score of all the testing examples in each category of fusion problem is used for evaluation. A popular approach for investigating the impacts of multiple parameters is known as *controlling for a variable*, which has also been widely adopted in the study of image fusion [9], [11], [29]. Considering the limitation of paper length, it is practically impossible to list all the results as it has a very large scale (a "cube" of results and too many combinations). For the sake of clarity, we just show one set of results for each parameter, in which the other two parameters are set to well-performed values (this is a commonly used manner in the literature [9], [11], [29]).

[3]In this paper, the number of directions at each decomposition level is set according to the number of decomposition level, as discussed later.

TABLE I

NUMBER OF NSST DECOMPOSITION LEVELS AND THE CORRESPONDING DIRECTION SETTINGS FOR PARAMETER ANALYSIS

Number of decomposition levels	Number of directions at each level (from fine to coarse)
1	16
2	16, 16
3	16, 16, 8
4	16, 16, 8, 8
5	16, 16, 8, 8, 4
6	16, 16, 8, 8, 4, 4

1) Number of NSST Decomposition Levels: In this set of experiments, the number of NSST decomposition levels L is set to 1–6, respectively. Another important issue to be determined is the number of directions at each decomposition level. In multiscale geometric transform (e.g., NSCT and NSST)-based image processing applications including image fusion [8], [21]–[23], [27], the number of directions (usually set to small powers of 2, such as 4, 8, and 16) is generally under the trend of successively decrease from a finer scale to a coarser scale (refer to Fig. 1 and note that the finest scale is first generated in NSST decomposition). In this paper, the numbers of directions are empirically set to 16, 16, 8, 8, 4, and 4 for the first six scales from fine to coarse. Therefore, the detailed direction settings for different values of L can be listed in Table I. In this situation, it is not difficult to observe that a smaller number of levels (e.g., three levels) decomposition is exactly a subset of a larger number of levels (e.g., four levels) decomposition in terms of high-frequency fusion. The first row of Fig. 5 shows the results indicating the influence of parameter L. The other two parameters are fixed as $N = 110$ and $r = 1$. It can be seen that the scores of all the metrics generally rise with the increase in L, but tend to be stable when $L \geq 4$. In some cases (e.g., Q_W and VIFF), the scores even have a tendency to decrease when $L \geq 4$. Therefore, and also considering that the computational cost will increase when L becomes larger, it is appropriate to apply a four-level decomposition in our method.

2) Number of PA-PCNN Iterations: The number of PA-PCNN iterations is, respectively, set to 50, 70, 90, 110, 130, and 150 to study its influence on the fusion performance.

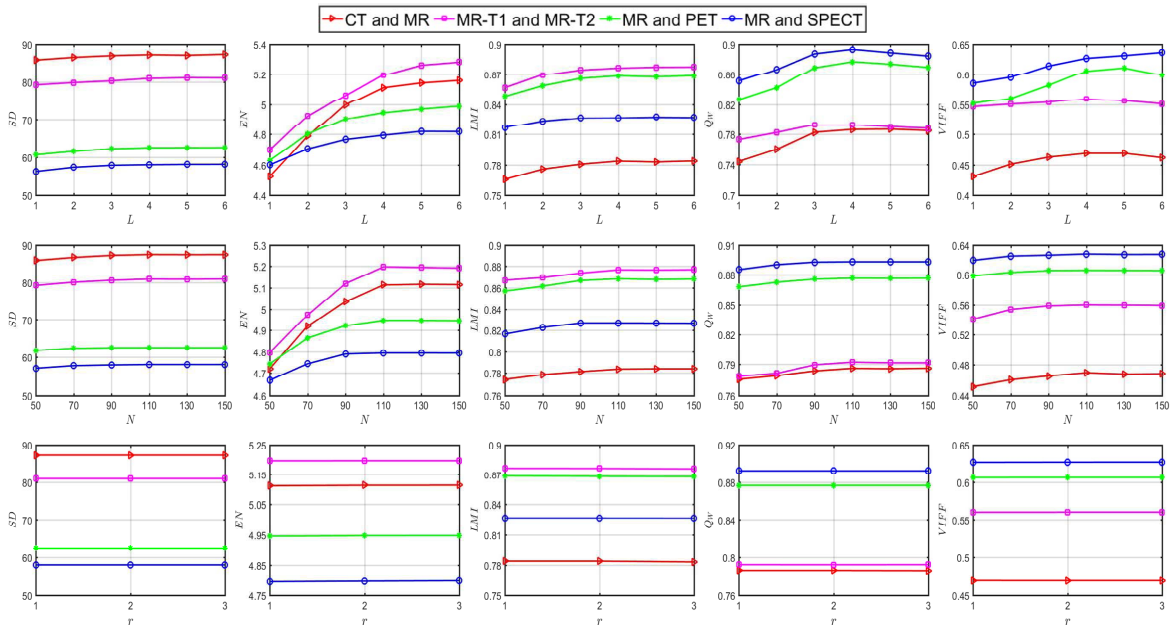


Fig. 5. Objective performance of the proposed method with different parameters including the number of NSST decomposition levels L, the number of PA-PCNN iterations N, and the radius r of weighting kernel. First row: impact of L with $N = 110$ and $r = 1$. Second row: impact of N with $L = 4$ and $r = 1$. Third row: impact of r with $L = 4$ and $N = 110$.

The results are shown in the second row of Fig. 5, where the other two parameters are fixed as $L = 4$ and $r = 1$. It is clear that the metric scores in most cases are convergent when $N = 110$, while some scores become stable just after about 90 iterations. Thus, $N = 110$ is a reasonable choice if a uniform setting is required. It is worthwhile to note that image fusion methods using conventional PCNN models usually need more iterations to reach convergence. For instance, 200 iterations are adopted in [21] and [23]. A possible explanation for this issue is that the parameters of PA-PCNN are adaptively calculated, rather than fixed to predetermined values, leading to a faster convergence speed. An obvious advantage of fewer iterations is its potential to achieve a higher computational efficiency.

3) Radius of Weighting Kernel: To investigate the impact of this parameter, three different values ($r = \{1, 2, 3\}$) are tested in this set of experiments, so the corresponding kernel sizes are 3×3, 5×5, and 7×7, respectively. In addition, the sizes of weighting kernels used in WLE and WSEML are set to the same. The third row of Fig. 5 exhibits the results with $L = 4$ and $N = 110$. In contrast to the above-mentioned two parameters, we can see that the effects of r on all the five metrics are so slight that can be ignorable. Considering the factor of computational cost, the 3×3 kernel ($r = 1$) is surely the best solution.

Based on the earlier discussions, the setting $\{L = 4, N = 110, r = 1\}$ is adopted in all the following experiments.

C. Comparison to Other Image Fusion Methods

In this section, the proposed method is compared with other approaches on three aspects: visual quality, objective assessment, and computational efficiency.

1) Visual Quality: For each category of the above-mentioned four medical image fusion problems, three sets of fusion results are given. Furthermore, two representative regions (mainly from the viewpoint of image processing community such as energy preservation, detail extraction, color fidelity, and so on) are enlarged as close-ups in each image to make better comparisons.

Fig. 6 shows three sets of CT and MR image fusion results. It can be seen that the SR-SOMP, GF, NSCT-PCDC, and CTD-SR methods lose a large amount of energy, leading to a significant decrease in the intensity and contrast in many regions [see the bone regions in Fig. 6(a4)–(a6), (a9), (b4)–(b6), (b9), (c4)–(c6), and (c9)]. The LP-SR and LP-CNN methods perform better on this issue, but some regions still tend to lose energy [see the second close-up in Fig. 6(a8) and the bone regions in Fig. 6(b11) and (c11)]. The NSCT-SF-PCNN method can preserve the image energy, but fails in extracting the structural details from the MR image [see the first close-up in Fig. 6(a3), (b3), and (c3)]. The NSCT-RPCNN performs better than the NSCT-SF-PCNN method, but some details are still not successfully extracted [see the first close-up in Fig. 6(a7), (b7), and (c7)]. The fused images of the LLF-IOI method suffers from serious noiselike artifacts [see the soft-tissue regions in Fig. 6(a10), (b10), and (c10)]. The proposed method performs well on both energy preservation and detail extraction in all these three examples [see Fig. 6(a12), (b12), and (c12)].

Fig. 7 shows three sets of MR-T1 and MR-T2 image fusion results. The SR-SOMP, GF, NSCT-PCDC, and CTD-SR methods still suffer from the undesirable effects caused by loss of energy. In addition, some important information (e.g., edge) in the MR-T1 source image is not well preserved in the fused images of these methods [see the second

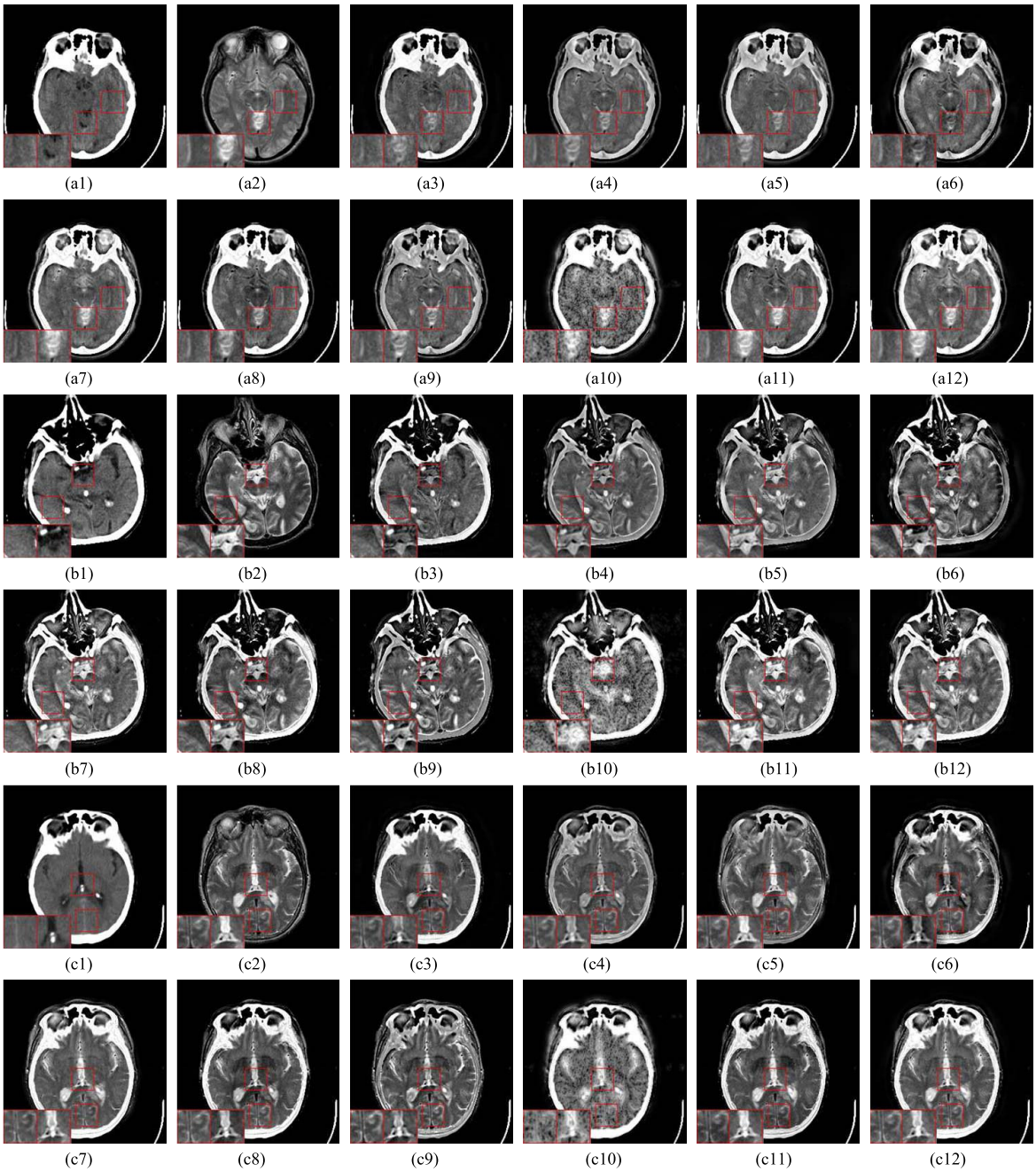


Fig. 6. Three sets of CT and MR image fusion results. Two close-ups are provided in each set for better comparison. The SR-SOMP, GF, NSCT-PCDC, CTD-SR, LP-SR, and LP-CNN methods suffer from loss of energy to varying degrees. The NSCT-SF-PCNN and NSCT-RPCNN methods have some defects in extracting anatomical details from the MR image. The LLF-IOI method introduces serious noiselike artifacts into the fused image. The proposed method performs well on both energy preservation and detail extraction. Please refer to the main text for more detailed descriptions. (a1) CT. (a2) MR. (a3) NSCT-SF-PCNN. (a4) SR-SOMP. (a5) GF. (a6) NSCT-PCDC. (a7) NSCT-RPCNN. (a8) LP-SR. (a9) CTD-SR. (a10) LLF-IOI. (a11) LP-CNN. (a12) Proposed. (b1) CT. (b2) MR. (b3) NSCT-SF-PCNN. (b4) SR-SOMP. (b5) GF. (b6) NSCT-PCDC. (b7) NSCT-RPCNN. (b8) LP-SR. (b9) CTD-SR. (b10) LLF-IOI. (b11) LP-CNN. (b12) Proposed. (c1) CT. (c2) MR. (c3) NSCT-SF-PCNN. (c4) SR-SOMP. (c5) GF. (c6) NSCT-PCDC. (c7) NSCT-RPCNN. (c8) LP-SR. (c9) CTD-SR. (c10) LLF-IOI. (c11) LP-CNN. (c12) Proposed.

close-up in Fig. 7(b4)–(b6) and (b9)]. The main defect of the NSCT-SF-PCNN and NSCT-RPCNN methods is their lower ability in detail extraction. Many details are blurred or even lost in the fused images of these two methods [see the first close-up in Fig. 7(a3), (a7), (b3), and (b7)]. Noiselike artifacts still exist in the fused images of the LLF-IOI method [see Fig. 7(a10), (b10), and (c10)]. The LP-SR and LP-CNN methods generally perform well, but intensity inconsistency exists in some regions [see Fig. 7(c8) and (c11)]. The proposed method obtains more competitive results [see Fig. 7(a12), (b12), and (c12)] than the other methods.

Fig. 8 shows three sets of MR and PET image fusion results. As structural details are mostly contained in the MR image, almost all of these methods perform well in detail extraction

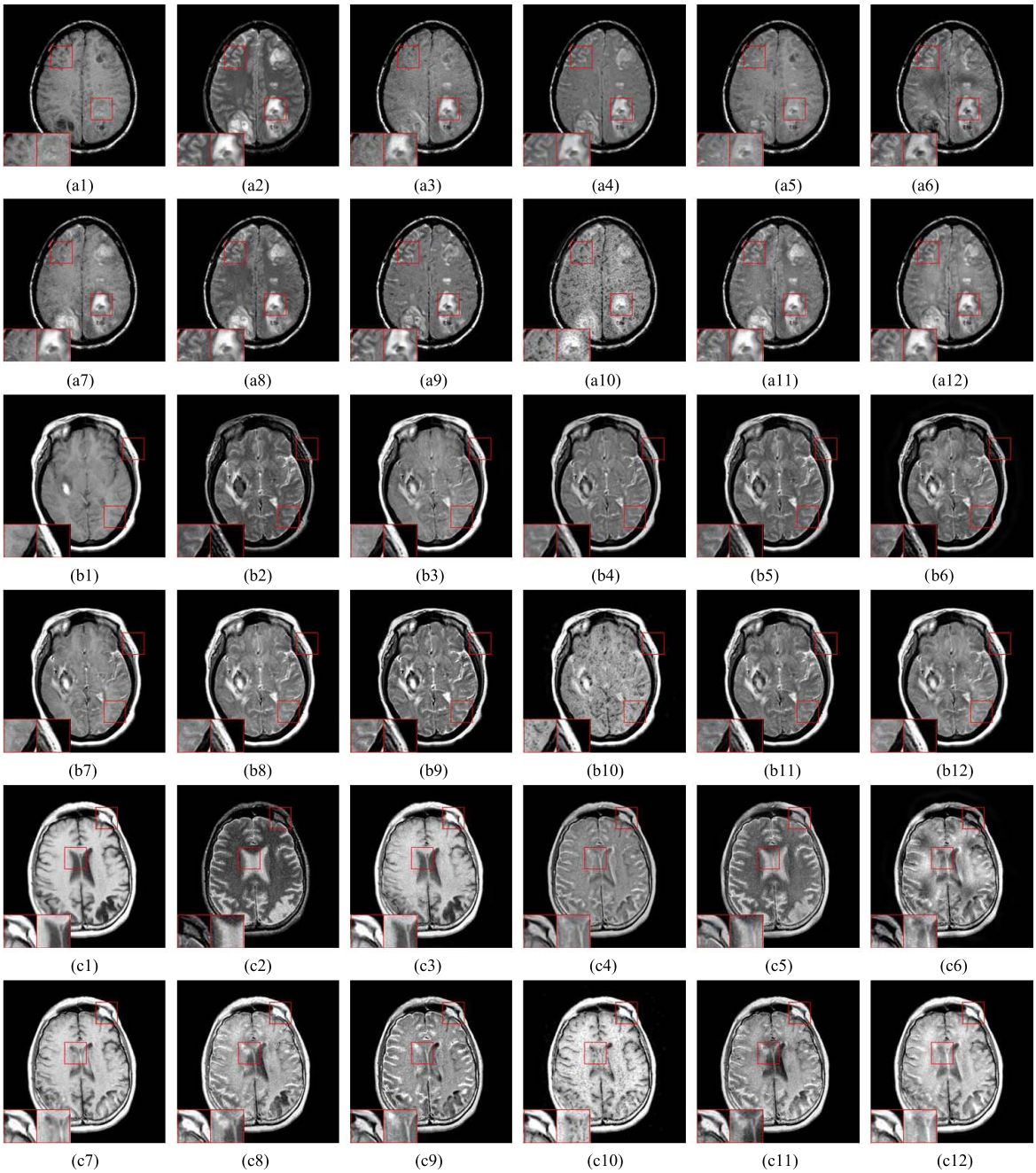


Fig. 7. Three sets of MR-T1 and MR-T2 image fusion results. Two close-ups are provided in each set for better comparison. The SR-SOMP, GF, NSCT-PCDC, and CTD-SR methods tend to lose energy and some important information in the MR-T1 source image are not well preserved. The NSCT-SF-PCNN and NSCT-RPCNN methods do not extract sufficient details. The LLF-IOI method suffers from serious noiselike artifacts. The LP-SR and LP-CNN methods introduce intensity inconsistency in some regions of the fused images. The proposed method obtains more competitive results than other methods. Please refer to the main text for more detailed descriptions. (a1) MR-T1. (a2) MR-T2. (a3) NSCT-SF-PCNN. (a4) SR-SOMP. (a5) GF. (a6) NSCT-PCDC. (a7) NSCT-RPCNN. (a8) LP-SR. (a9) CTD-SR. (a10) LLF-IOI. (a11) LP-CNN. (a12) Proposed. (b1) MR-T1. (b2) MR-T2. (b3) NSCT-SF-PCNN. (b4) SR-SOMP. (b5) GF. (b6) NSCT-PCDC. (b7) NSCT-RPCNN. (b8) LP-SR. (b9) CTD-SR. (b10) LLF-IOI. (b11) LP-CNN. (b12) Proposed. (c1) MR-T1. (c2) MR-T2. (c3) NSCT-SF-PCNN. (c4) SR-SOMP. (c5) GF. (c6) NSCT-PCDC. (c7) NSCT-RPCNN. (c8) LP-SR. (c9) CTD-SR. (c10) LLF-IOI. (c11) LP-CNN. (c12) Proposed.

while the main difference lies in the color fidelity. It can be clearly seen that the fusion results of the NSCT-SF-PCNN, SR-SOMP, GF, and NSCT-PCDC methods suffer from severe color distortion [see Fig. 8(a3)–(a6), (b3)–(b6), and (c3)–(c6)]. The LP-SR, CTD-SR, and LP-CNN perform better on this issue, but color distortion still exists more or less in the fused images of these three methods [see Fig. 8(a8), (a9), (a11), (b9), (c8), and (c9)]. The LLF-IOI method tackles the color information well, but the spatial information in the MR source image is over enhanced such that the original structural information is destroyed [see Fig. 8(a10), (b10), and (c10)]. The NSCT-RPCNN and proposed methods can achieve higher visual quality in terms of color preservation than other methods, and our method performs better than the

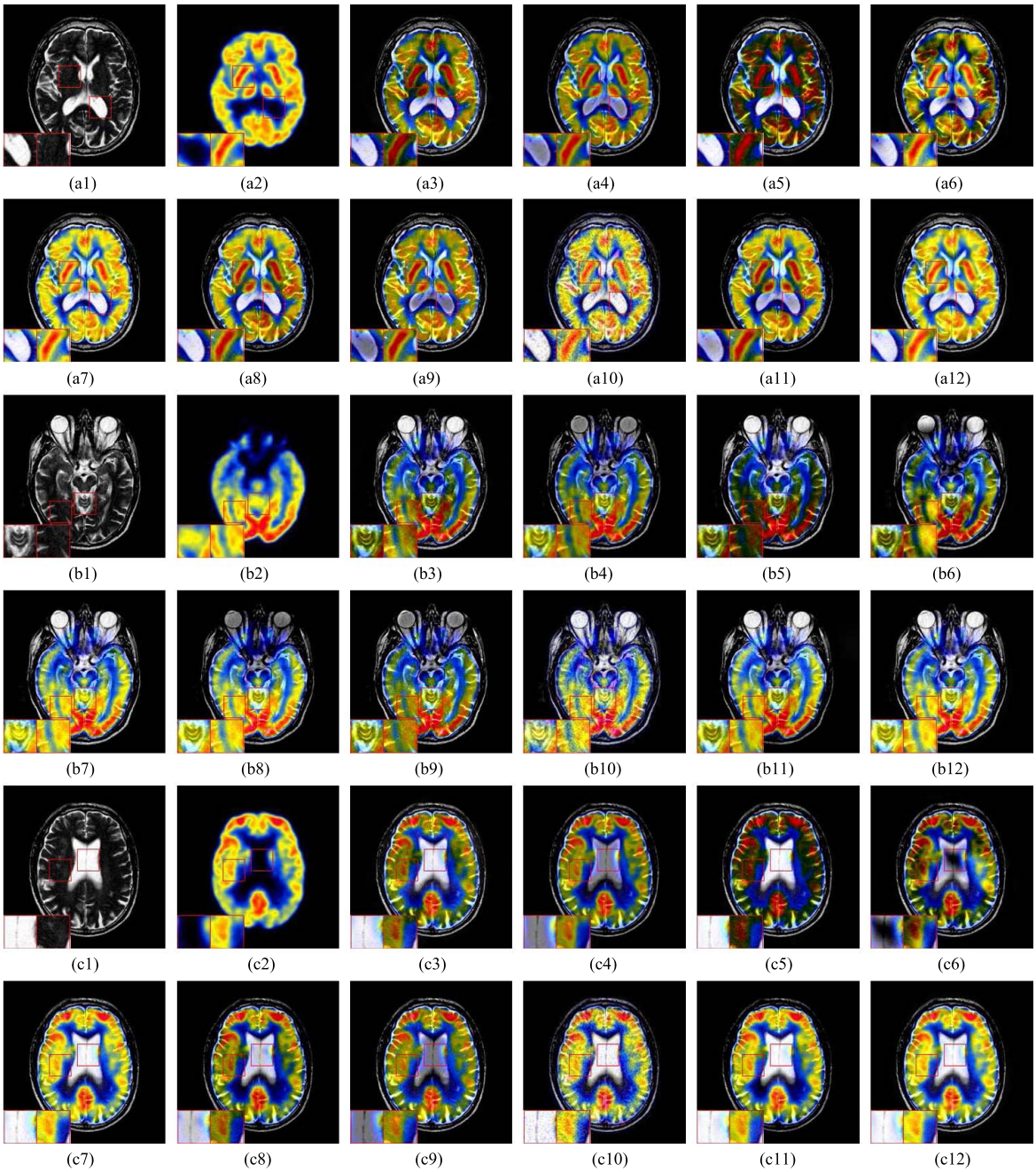


Fig. 8. Three sets of MR and PET image fusion results. Two close-ups are provided in each set for better comparison. The NSCT-SF-PCNN, SR-SOMP, GF, NSCT-PCDC, LP-SR, CTD-SR, and LP-CNN methods suffer from color distortion at varying degrees. The LLF-IOI method over enhances the spatial information in the MR source image, leading to the damage of original structural information. Both the NSCT-RPCNN and proposed methods perform well in terms of color preservation, but the NSCT-RPCNN method tends to lose a few structural details. Please refer to the main text for more detailed descriptions. (a1) MR. (a2) PET. (a3) NSCT-SF-PCNN. (a4) SR-SOMP. (a5) GF. (a6) NSCT-PCDC. (a7) NSCT-RPCNN. (a8) LP-SR. (a9) CTD-SR. (a10) LLF-IOI. (a11) LP-CNN. (a12) Proposed. (b1) MR. (b2) PET. (b3) NSCT-SF-PCNN. (b4) SR-SOMP. (b5) GF. (b6) NSCT-PCDC. (b7) NSCT-RPCNN. (b8) LP-SR. (b9) CTD-SR. (b10) LLF-IOI. (b11) LP-CNN. (b12) Proposed. (c1) MR. (c2) PET. (c3) NSCT-SF-PCNN. (c4) SR-SOMP. (c5) GF. (c6) NSCT-PCDC. (c7) NSCT-RPCNN. (c8) LP-SR. (c9) CTD-SR. (c10) LLF-IOI. (c11) LP-CNN. (c12) Proposed.

NSCT-RPCNN method on detail extraction in some regions [see the first close-up in Fig. 8(b7) and (b12)].

Fig. 9 shows three sets of MR and SPECT image fusion results. The performances of the SR-SOMP, GF, NSCT-PCDC, LP-SR, and CTD-SR methods on preserving color fidelity are relatively low [see the serious visual inconsistency in the white regions in Fig. 9(a4)–(a6), (a8), and (a9)]. The LLF-IOI method over enhances the anatomical details in the MR source images. As a consequence, some important functional information contained in the SPECT source image is lost [see Fig. 9(a10), (b10), and (c10)]. The NSCT-SF-PCNN and NSCT-RPCNN generally perform well, but some defects still exist in terms of the preserving functional information from the SPECT source image [see the second close-up

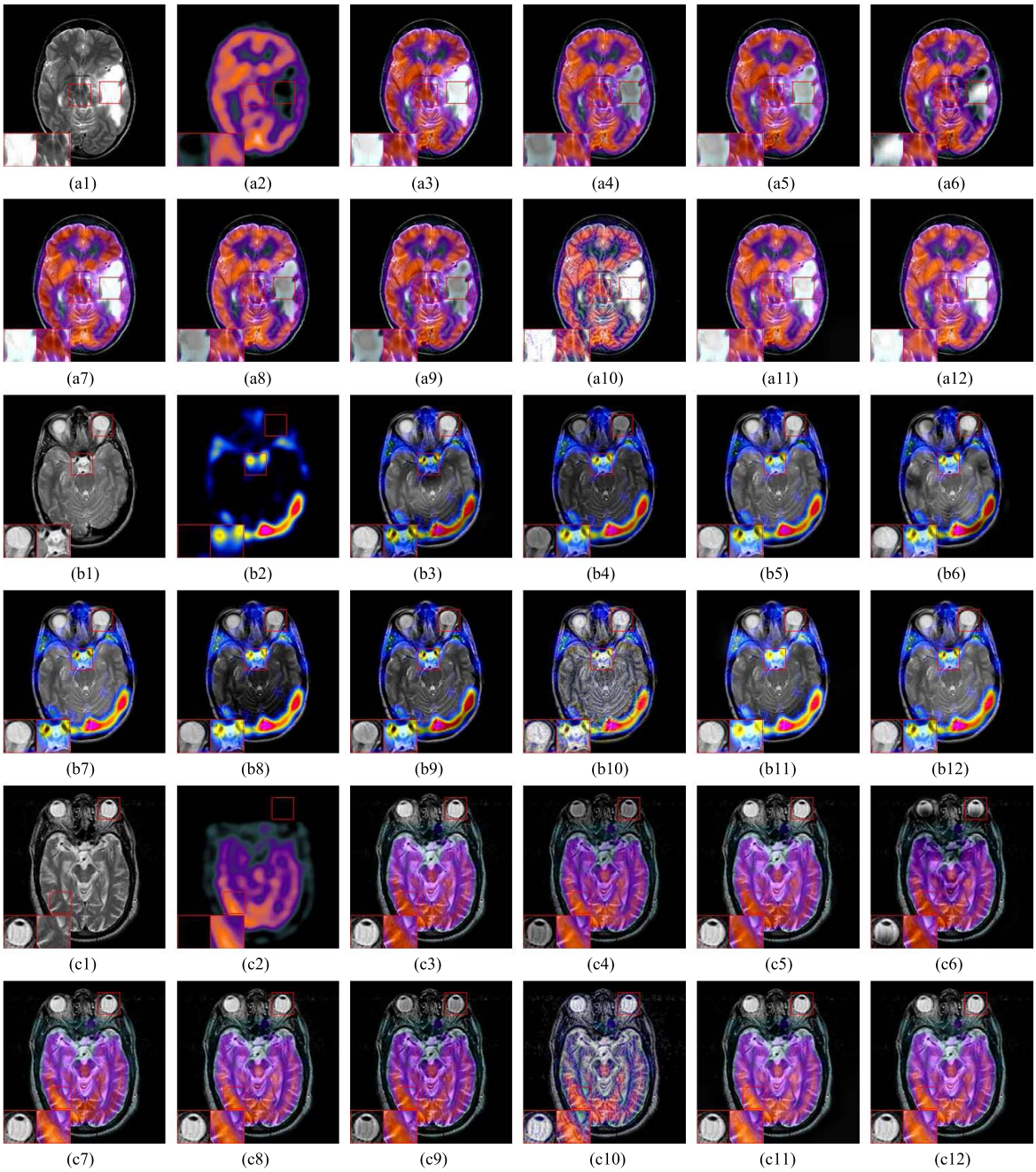


Fig. 9. Three sets of MR and SPECT image fusion results. Two close-ups are provided in each set for better comparison. The SR-SOMP, GF, NSCT-PCDC, LP-SR, and CTD-SR methods are not effective in preserving color fidelity. The LLF-IOI method tends to lose some important functional information contained in the SPECT source image as some structural details in the MR source images are over enhanced. The NSCT-SF-PCNN and NSCT-RPCNN generally perform well, but some defects still exist in terms of the preserving functional information. The LP-CNN and proposed methods have the best performance on color preservation, but some details are not well extracted by the LP-CNN method. Please refer to the main text for more detailed descriptions. (a1) MR. (a2) SPECT. (a3) NSCT-SF-PCNN. (a4) SR-SOMP. (a5) GF. (a6) NSCT-PCDC. (a7) NSCT-RPCNN. (a8) LP-SR. (a9) CTD-SR. (a10) LLF-IOI. (a11) LP-CNN. (a12) Proposed. (b1) MR. (b2) SPECT. (b3) NSCT-SF-PCNN. (b4) SR-SOMP. (b5) GF. (b6) NSCT-PCDC. (b7) NSCT-RPCNN. (b8) LP-SR. (b9) CTD-SR. (b10) LLF-IOI. (b11) LP-CNN. (b12) Proposed. (c1) MR. (c2) SPECT. (c3) NSCT-SF-PCNN. (c4) SR-SOMP. (c5) GF. (c6) NSCT-PCDC. (c7) NSCT-RPCNN. (c8) LP-SR. (c9) CTD-SR. (c10) LLF-IOI. (c11) LP-CNN. (c12) Proposed.

in Fig. 9(a2), (a3), (a7), and (a12)]. The LP-CNN method and the proposed method have the best performance on color preservation, but some structural details are blanketed by the functional information in the fusion results of the LP-CNN method [see the second close-up in Fig. 9(b11)].

2) Objective Assessment: Table II lists the objective assessment of different fusion methods on four categories of medical image fusion problems. The average score of each method over all the testing images in each fusion problem is reported. For each metric, the highest value shown in bold indicates the best score among all the 10 methods, and we also underline the scores at the second place and the third place. For each of the scores at the first three places, a digit within a parenthesis is given to indicate its rank. To have a more intuitive grasp

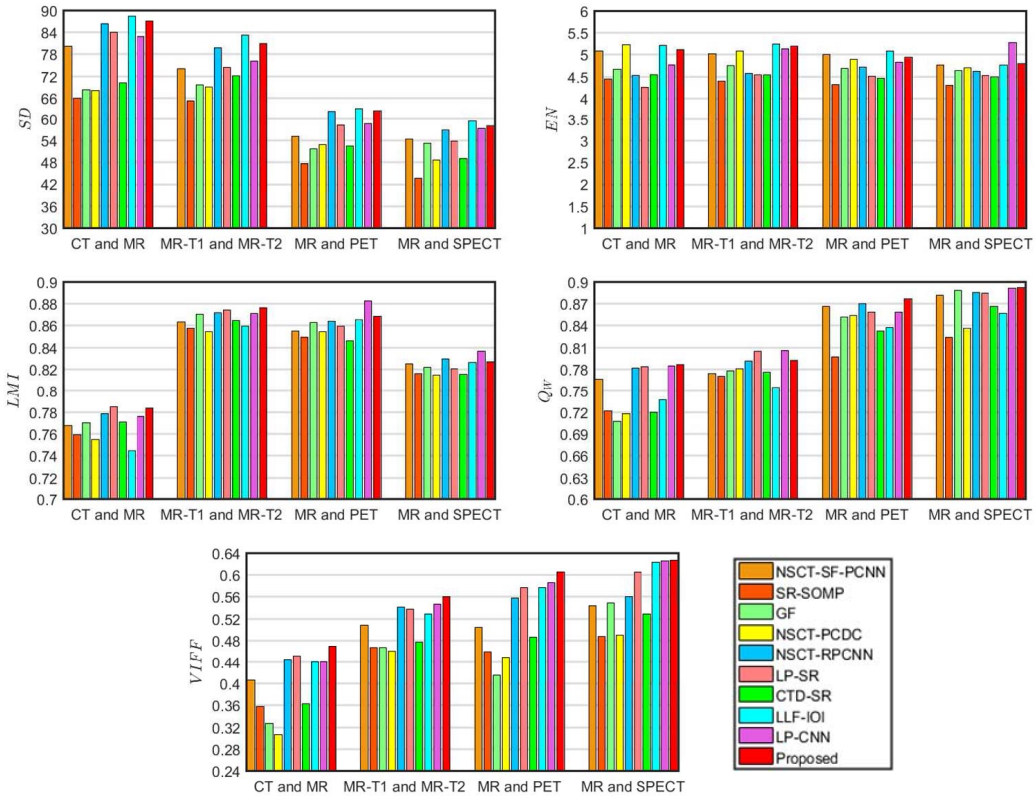


Fig. 10. Objective assessment of different image fusion methods. Among all the 10 methods, the proposed method is the only one that always ranks at the first three places for all the five metrics and all the four types of fusion problems.

about the objective performances of different fusion methods, the contents given in Table II are visualized in Fig. 10.

Overall, among all the 10 methods, our method is the only one that always ranks at the first three places for all the five metrics and all the four types of fusion problems, which reflects the high robustness of the proposed method. In particular, our method wins the first place on the metric VIFF for all the four problems.

In comparison to the NSCT-SF-PCNN and NSCT-RPCNN methods, it can be seen from Table II and Fig. 10 that the proposed method outperforms those two methods on almost all the metrics over all the four image fusion problems, which indicates the clear advantages of our method.

We can see from Table II and Fig. 10 that the LLF-IOI and LP-CNN methods also have high objective performances. The LLF-IOI always obtains the highest scores on metric SD, but this is just in accord with its noiselike visual perception mentioned earlier. Our method owns obvious advantages over the LLF-IOI method on metrics LMI, Q_W, and VIFF. The LP-CNN method obtains higher scores on metric LMI in the last two types of fusion problems, but is beaten by the proposed method in most of the other cases.

3) Computational Efficiency: In this section, the computational costs of different fusion methods are compared. Without loss of generality, this part of experiment is conducted on the set of CT and MR source images while the results on other three image sets are very similar.[4] Specifically, for each method, we first measure the total running time on fusing all the 10 pairs of source images, and then divide it by 10 to obtain the mean value T indicating the average running time to fuse one pair of source images. The above-mentioned steps are repeated eight times to calculate the statistical average value and SD (unbiased) of T. The results of different fusion methods are listed in Table III. It can be seen that the computational efficiency of the proposed method is just lower than the LP-SR and GF methods, while higher than the other seven methods. In particular, the proposed method is more efficient than the other two PCNN-based methods, which is mainly owing to the fewer PCNN iterations required.[5] With a more efficient programming language such as C++, it can be predicted that the computational efficiency of our method is able to satisfy the demand of practical applications.

D. Further Discussion

Considering the above-mentioned comparisons comprehensively, it can be concluded that the proposed method can

[4]All the source images are of the same size and conversion between the RGB and YUV color spaces are very efficient. More importantly, as the results are relevant to the hardware/software platforms, the relative differences among different methods are of more significance than the absolute values.

[5]The efficiency of the NSCT-SF-PCNN method is lower than that of the NSCT-RPCNN method and our method to a considerable extent, which is mainly because the directional filters used in [21] have much larger spatial sizes. In this paper, for the sake of fair comparison, all the parameters in the compared methods are set to their default values used in the original implementations.

TABLE II

OBJECTIVE ASSESSMENT OF DIFFERENT METHODS ON FOUR CATEGORIES OF MEDICAL IMAGE FUSION PROBLEMS

Images	Metrics	NSCT-SF-PCNN	SR-SOMP	GF	NSCT-PCDC	NSCT-RPCNN	LP-SR	CTD-SR	LLF-IOI	LP-CNN	Proposed
CT and MR	SD	80.2229	65.9633	68.2632	68.1021	86.4484(3)	84.0814	70.1822	**88.5294(1)**	82.8770	87.1653(2)
	EN	5.0921	4.4392	4.6712	**5.2250(1)**	4.5307	4.2613	4.5501	5.2089(2)	4.7745	5.1139(3)
	LMI	0.7679	0.7594	0.7704	0.7548	0.7786(3)	**0.7853(1)**	0.7713	0.7446	0.7765	0.7837(2)
	Q_W	0.7660	0.7221	0.7078	0.7183	0.7816	0.7837(3)	0.7199	0.7376	0.7849(2)	**0.7862(1)**
	$VIFF$	0.4062	0.3582	0.3274	0.3064	0.4448(3)	0.4514(2)	0.3631	0.4406	0.4404	**0.4691(1)**
MR-T1 and MR-T2	SD	74.1095	65.2850	69.6947	69.0816	79.7867(3)	74.4300	72.0650	**83.2814(1)**	76.2079	80.9326(2)
	EN	5.0170	4.3975	4.7547	5.0796	4.5771	4.5401	4.5353	**5.2446(1)**	5.1420(3)	5.1961(2)
	LMI	0.8632	0.8577	0.8704	0.8546	0.8720(3)	0.8742(2)	0.8646	0.8598	0.8709	**0.8766(1)**
	Q_W	0.7741	0.7704	0.7779	0.7805	0.7914	0.8047(2)	0.7759	0.7549	**0.8062(1)**	0.7923(3)
	$VIFF$	0.5084	0.4670	0.4670	0.4610	0.5407(3)	0.5376	0.4771	0.5291	0.5471(2)	**0.5600(1)**
MR and PET	SD	55.1429	47.6831	51.7561	52.9358	62.3292(3)	58.3257	52.5482	**62.9966(1)**	58.6606	62.4796(2)
	EN	5.0049(2)	4.3179	4.6909	4.8931	4.7210	4.5065	4.4625	**5.0854(1)**	4.8230	4.9461(3)
	LMI	0.8550	0.8497	0.8630	0.8542	0.8639	0.8595	0.8459	0.8652(3)	**0.8828(1)**	0.8685(2)
	Q_W	0.8668(3)	0.7970	0.8520	0.8540	0.8702(2)	0.8583	0.8326	0.8376	0.8583	**0.8769(1)**
	$VIFF$	0.5040	0.4594	0.4153	0.4495	0.5575	0.5768(3)	0.4864	0.5767	0.5862(2)	**0.6060(1)**
MR and SPECT	SD	54.4103	43.6319	53.3418	48.6938	56.8244	53.8736	49.1180	**59.3942(1)**	57.3672(3)	58.0182(2)
	EN	4.7694(3)	4.2969	4.6320	4.7001	4.6151	4.5258	4.4930	4.7658	**5.2794(1)**	4.7940(2)
	LMI	0.8252	0.8159	0.8219	0.8148	0.8297(2)	0.8204	0.8152	0.8260	**0.8368(1)**	0.8268(3)
	Q_W	0.8816	0.8237	0.8883(3)	0.8362	0.8857	0.8844	0.8667	0.8572	0.8914(2)	**0.8925(1)**
	$VIFF$	0.5435	0.4874	0.5488	0.4902	0.5602	0.6058	0.5289	0.6237(3)	0.6259(2)	**0.6265(1)**

TABLE III

RUNNING TIME OF DIFFERENT METHODS WHEN FUSING TWO SOURCE IMAGES OF SIZE 256 × 256 PIXELS (UNIT: SECONDS)

	NSCT-SF-PCNN	SR-SOMP	GF	NSCT-PCDC	NSCT-RPCNN	LP-SR	CTD-SR	LLF-IOI	LP-CNN	Proposed
Average value	52.68	42.17	0.047	14.93	11.23	0.036	23.76	124.25	12.07	7.68
Standard deviation	0.22	0.39	0.002	0.04	0.03	0.001	0.18	0.37	0.07	0.04

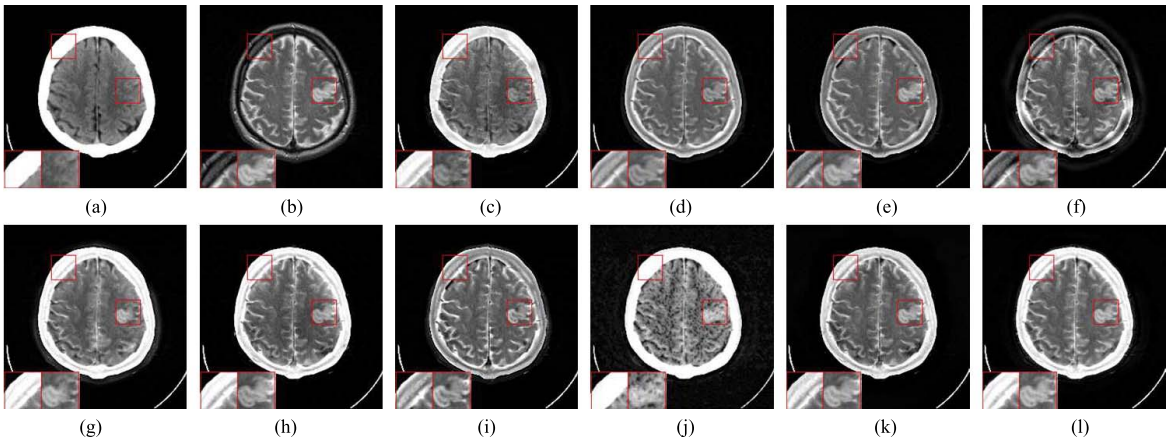


Fig. 11. CT and MR image fusion examples in our experiments. The clinical information in this case is of a 63-year-old male getting acute stroke. The proposed method performs better than most of the other methods in the abnormal region (right close-up) when referring to the MR source image. However, the fusion quality of the bone regions (an example is given by the left close-up) from the CT image is not satisfactory in almost all the fusion results except for the one achieved by the LLF-IOI method. (a) CT. (b) MR. (c) NSCT-SF-PCNN. (d) SR-SOMP. (e) GF. (f) NSCT-PCDC. (g) NSCT-RPCNN. (h) LP-SR. (i) CTD-SR. (j) LLF-IOI. (k) LP-CNN. (l) Proposed.

obtain very competitive performance among those state-of-the-art fusion methods. In spite of this, there still exist some defects that are worth further studying. Fig. 11 shows a CT and MR image fusion example in which our method does not perform very well in some regions. In this example, the CT image mainly provides accurate location of the bones, as illustrated by the first close-up. The MR image captures the information of brain tissues and the second close-up focuses

on an abnormal region (the clinical information in this case is of a 63-year-old male getting acute stroke). It can be seen that our method outperforms most of the other methods (e.g., NSCT-SF-PCNN, SR-SOMP, NSCT-RPCNN, CTD-SR, and LLF-IOI) in the abnormal region when referring to the MR source image. Another noticeable issue in this example is that the fusion quality of bone regions from the CT image is not satisfactory in almost all the results except for the one achieved by the LLF-IOI method. This is mainly because there are no salient details in the bone regions, while most image fusion methods including the proposed one concentrate on the extraction of structural details. As a result, the meaningless details in the same regions of the MR image are introduced into the fused image more or less, leading to undesirable visual artifacts. To overcome this defect, in our opinion, some prior information of the source images maybe helpful to design more effective fusion strategies for further improvement.

VI. Conclusion

In this paper, a new medical image fusion method in the NSST domain is presented. The main novelty of the proposed method is twofold. For one thing, we introduce a PA-PCNN model into the fusion of high-frequency coefficients. All the free parameters in the PCNN model can be adaptively calculated according to the input band and the model has a fast convergence speed. For another, we propose a low-frequency fusion strategy that simultaneously addresses two crucial issues in medical image fusion, namely, energy preservation and detail extraction. Two new activity level measures based on local energy and ML are designed to achieve this goal. Extensive experiments are conducted using 83 pairs of source images over four categories of medical image fusion problems to verify the effectiveness of the proposed method. Nine representative fusion methods are used for comparison and the results demonstrate that the proposed method can achieve state-of-the-art performance in terms of both the visual perception and objective assessment. In the future, we will devote to develop more effective fusion strategies such as region-adaptive-based ones to further improve the algorithm performance. Another goal is to promote the practical value of the image fusion method in clinical applications by simultaneously considering some related critical issues such as data preprocessing and image registration for multimodality medial images. Moreover, we will also explore the potential of the PA-PCNN model for other image fusion issues such as multifocus image fusion, infrared and visible image fusion, and so on.

References

[1] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 19, pp. 4–19, Sep. 2014.

[2] J. Du, W. Li, K. Lu, and B. Xiao, "An overview of multi-modal medical image fusion," *Neurocomputing*, vol. 215, pp. 3–20, Nov. 2016.

[3] Z. Wang and Y. Ma, "Medical image fusion using m-PCNN," *Inf. Fusion*, vol. 9, pp. 176–185, Apr. 2008.

[4] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Inf. Fusion*, vol. 13, no. 1, pp. 10–19, 2012.

[5] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 4, pp. 691–703, Apr. 2017.

[6] S. Daneshvar and H. Ghassemian, "MRI and PET image fusion by combining IHS and retina-inspired models," *Inf. Fusion*, vol. 11, no. 2, pp. 114–123, 2010.

[7] R. Singh and A. Khare, "Fusion of multimodal medical images using Daubechies complex wavelet transform—A multiresolution approach," *Inf. Fusion*, vol. 19, pp. 49–60, Sep. 2014.

[8] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Inf. Fusion*, vol. 24, pp. 147–164, Jul. 2015.

[9] W. Zhao and H. Lu, "Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2283–2294, Sep. 2017.

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

[11] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.

[12] Q. Wang, S. Li, H. Qin, and A. Hao, "Robust multi-modal medical image fusion via anisotropic heat diffusion guided low-rank structural analysis," *Inf. Fusion*, vol. 26, pp. 103–121, Nov. 2015.

[13] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1882–1886, Dec. 2016.

[14] G. Piella, "A general framework for multiresolution image fusion: From pixels to regions," *Inf. Fusion*, vol. 4, no. 4, pp. 259–280, 2003.

[15] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. TCOMM-31, no. 4, pp. 532–540, Apr. 1983.

[16] A. Toet, "A morphological pyramidal image decomposition," *Pattern Recognit. Lett.*, vol. 9, no. 4, pp. 255–261, 1989.

[17] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.

[18] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Inf. Fusion*, vol. 8, no. 2, pp. 119–130, 2007.

[19] Q. Zhang and B.-L. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Process.*, vol. 89, no. 7, pp. 1334–1346, 2009.

[20] G. Guorong, X. Luping, and F. Dongzhu, "Multi-focus image fusion based on non-subsampled shearlet transform," *IET Image Process.*, vol. 7, no. 6, pp. 633–639, 2013.

[21] X.-B. Qu, J.-W. Yan, H.-Z. Xiao, and Z.-Q. Zhu, "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain," *Acta Autom. Sin.*, vol. 34, no. 12, pp. 1508–1514, 2008.

[22] G. Bhatnagar, Q. M. J. Wu, and Z. Liu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1014–1024, Aug. 2013.

[23] S. Das and M. K. Kundu, "A neuro-fuzzy approach for medical image fusion," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3347–3353, Dec. 2013.

[24] Z. Liu, H. Yin, Y. Chai, and S. X. Yang, "A novel approach for multimodal medical image fusion," *Expert Syst. Appl.*, vol. 41, no. 16, pp. 7425–7435, 2014.

[25] L. Wang, B. Li, and L. Tian, "Multimodal medical volumetric data fusion using 3-D discrete shearlet transform and global-to-local rule," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 1, pp. 197–206, Jan. 2014.

[26] J. Du, W. Li, B. Xiao, and Q. Nawaz, "Union Laplacian pyramid with multiple features for medical image fusion," *Neurocomputing*, vol. 194, pp. 326–339, Jun. 2016.

[27] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multimodal sensor medical image fusion based on type-2 fuzzy logic in NSCT domain," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3735–3745, May 2016.

[28] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Inf. Sci.*, vol. 432, pp. 516–529, Mar. 2018.

[29] J. Du, W. Li, and B. Xiao, "Anatomical-functional image fusion by information of interest in local laplacian filtering domain," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5855–5866, Dec. 2017.

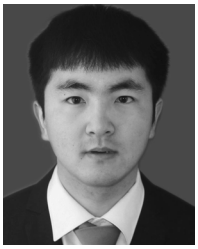
[30] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," in *Proc. 20th Int. Conf. Inf. Fusion*, Jul. 2017, pp. 1–7.

[31] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 884–892, Apr. 2010.

[32] Y. Liu, X. Chen, H. Peng, and Z. F. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Inf. Fusion*, vol. 36 pp. 191–207, Jul. 2017.
[33] J. L. Johnson, "Pulse-coupled neural nets: Translation, rotation, scale, distortion, and intensity signal invariance for images," *Appl. Opt.*, vol. 33, no. 26, pp. 6239–6253, 1994.
[34] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, "Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex," *Neural Comput.*, vol. 2, no. 3, pp. 293–307, 1990.
[35] Y. Chen, S. K. Park, Y. Ma, and R. Ala, "A new automatic parameter setting method of a simplified PCNN for image segmentation," *IEEE Trans. Neural Netw.*, vol. 22, no. 6, pp. 880–892, Jun. 2011.
[36] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 25–46, Jul. 2008.
[37] S. Singh, D. Gupta, R. S. Anand, and V. Kumar, "Nonsubsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network," *Biomed. Signal Process. Control*, vol. 18, pp. 91–101, Apr. 2015.
[38] W. Kong, B. Wang, and Y. Lei, "Technique for infrared and visible image fusion based on non-subsampled shearlet transform and spiking cortical model," *Infr. Phys. Technol.*, vol. 71, pp. 87–98, Jul. 2015.
[39] M. Yin, W. Liu, X. Zhao, Y. Yin, and Y. Guo, "A novel image fusion algorithm based on nonsubsampled shearlet transform," *Optik-Int. J. Light Electron Opt.*, vol. 125, no. 10, pp. 2274–2282, 2014.
[40] K. Zhan, J. Shi, H. Wang, Y. Xie, and Q. Li, "Computational mechanisms of pulse-coupled neural networks: A comprehensive review," *Arch. Comput. Methods Eng.*, vol. 24, pp. 573–588, Jul. 2017.
[41] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognit. Lett.*, vol. 28, no. 4, pp. 493–500, 2007.
[42] G. Bhatnagar, Q. M. J. Wu, and Z. Liu, "Human visual system inspired multi-modal medical image fusion framework," *Expert Syst. Appl.*, vol. 40, pp. 1708–1720, Apr. 2013.
[43] *The Whole Brain Atlas of Harvard Medical School*. Accessed: Nov. 2, 2015. [Online]. Available: http://www.med.harvard.edu/AANLIB/
[44] *The Source Code of the NSCT-SF-PCNN Method From Xiaobo Qu's Website*. Accessed: Dec. 24, 2015. [Online]. Available: http://csrc.xmu.edu.cn/
[45] *The Source Code of the GF Method From Xudong Kang's Website*. Accessed: Sep. 10, 2013. [Online]. Available: http://xudongkang.weebly.com/index.html
[46] *The Source Code of the NSCT-PCDC Method From Gaurav Bhatnagar's Website*. Accessed: Mar. 4, 2017. [Online]. Available: https://sites.google.com/site/goravdma
[47] *The Source Code of the NSCT-RPCNN Method From Dudeb Das's Website*. Accessed: Oct. 16, 2017. [Online]. Available: https://sites.google.com/site/wodrsdas/
[48] *The Source Code of the LP-SR Method From Yu Liu's Website*. Accessed: Feb. 22, 2017. [Online]. Available: http://www.escience.cn/people/liuyu1/englishing.html
[49] *The Source Code of the LLF-IOI Method From Jiao Du's Website*. Accessed: Oct. 12, 2017. [Online]. Available: http://www.escience.cn/people/xiaomi
[50] M. Hossny, S. Nahavandi, D. Creighton, and A. Bhatti, "Image fusion performance metric based on mutual information and entropy driven quadtree decomposition," *Electron. Lett.*, vol. 46, no. 18, pp. 1266–1268, 2010.
[51] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. 10th Int. Conf. Image Process.*, Sep. 2003, pp. 173–176.
[52] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Inf. Fusion*, vol. 14, pp. 127–135, Apr. 2013.

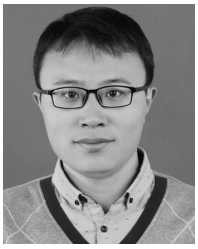
Ming Yin received the B.E. degree from Anhui Normal University, Wuhu, China, in 1985, and the M.S. and Ph.D. degrees from the Hefei University of Technology, Hefei, China, in 1991 and 2012, respectively.

He is currently a Professor with the School of Mathematics, Hefei University of Technology. His current research interests include wavelet transform, image processing, and compressed sensing.



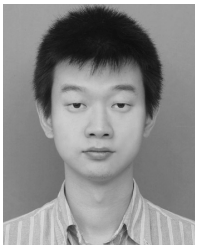
Xiaoning Liu received the B.E. degree from Jiaxing University, Jiaxing, China, in 2015. He is currently pursuing the M.S. degree in computational mathematics with the Hefei University of Technology, Hefei, China.

His current research interests include image fusion and sparse representation.



Yu Liu (M'17) received the B.S. and Ph. D degrees from the Department of Automation, University of Science and Technology of China, Hefei, China, in 2011 and 2016, respectively.

He is currently an Assistant Professor with the Department of Biomedical Engineering, Hefei University of Technology, Hefei. His current research interests include image processing, computer vision, information fusion, and machine learning, in particular image fusion, image restoration, visual recognition, and deep learning.



Xun Chen (M'14) received the B.S. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2009, and the Ph.D. degree in biomedical engineering from the University of British Columbia, Vancouver, BC, Canada, in 2014.

He was a Research Scientist with the Department of Electrical and Computer Engineering, University of British Columbia. He is currently a Full Professor with the Department of Electronic Science and Technology, University of Science and Technology of China. His current research interests include statistical signal processing and machine learning in biomedical applications.