*Research Article*

# Log-Gabor Energy Based Multimodal Medical Image Fusion in NSCT Domain

## Yong Yang,[1] Song Tong,[1] Shuying Huang,[2] and Pan Lin[3]

[1] School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China
[2] School of Software and Communication Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China
[3] Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Correspondence should be addressed to Shuying Huang; shuyinghuang2010@126.com

Multimodal medical image fusion is a powerful tool in clinical applications such as noninvasive diagnosis, image-guided radiotherapy, and treatment planning. In this paper, a novel nonsubsampled Contourlet transform (NSCT) based method for multimodal medical image fusion is presented, which is approximately shift invariant and can effectively suppress the pseudo-Gibbs phenomena. The source medical images are initially transformed by NSCT followed by fusing low- and high-frequency components. The phase congruency that can provide a contrast and brightness-invariant representation is applied to fuse low-frequency coefficients, whereas the Log-Gabor energy that can efficiently determine the frequency coefficients from the clear and detail parts is employed to fuse the high-frequency coefficients. The proposed fusion method has been compared with the discrete wavelet transform (DWT), the fast discrete curvelet transform (FDCT), and the dual tree complex wavelet transform (DTCWT) based image fusion methods and other NSCT-based methods. Visually and quantitatively experimental results indicate that the proposed fusion method can obtain more effective and accurate fusion results of multimodal medical images than other algorithms. Further, the applicability of the proposed method has been testified by carrying out a clinical example on a woman affected with recurrent tumor images.

## 1. Introduction

Medical imaging has attracted increasing attention in the recent years due to its vital component in medical diagnostics and treatment [1]. However, each imaging modality reports on a restricted domain and provides information in limited domains that some are common, and some are unique [2]. For example, computed tomography (CT) image can provide dense structures like bones and hard tissues with less distortion whereas magnetic resonance imaging (MRI) image is better visualized in the case of soft tissues [3]. Similarly, T1-MRI image provides the details of an anatomical structure of tissues while T2-MRI image provides information about normal and pathological tissues [4]. As a result, multimodal medical images which have relevant and complementary information are necessary to be combined for a compendious figure [5]. The multimodal medical image fusion is the possible way to integrate complementary information from multiple modality images [6]. The image fusion not only obtains a more accurate and complete description of the same target, but also reduces the randomness and redundancy to increase the clinical applicability of image-guided diagnosis and assessment of medical problems [7].

Generally image fusion techniques can be divided into spatial domain and frequency domain techniques [8]. Spatial domain techniques are carried out directly on the source images. Weighted average method is the simplest spatial domain approach. However, along with simplicity, this method leads to several undesirable side effects like reduced contrast [9]. Other spatial domain techniques have been developed, such as intensity-hue-saturation (IHS), principal component analysis (PCA), and the Brovey transform [10]. Although the fused images obtained by these methods have high spatial quality, they usually overlook the high quality

of spectral information and suffer from spectral degradation [10]. Li et al. [11] introduced the artificial neural network (ANN) to make image fusion. However, the performance of ANN depends on the sample images and this is not an appealing characteristic. Yang and Blum [12] used a statistical approach to fuse the images. In their method, the distortion is modeled as a mixture of Gaussian probability density functions which is a limiting assumption. Since the actual objects usually contain structures at many different scales or resolutions and multiscale techniques can provide the means to exploit this fact, the frequency domain techniques especially the multiscale techniques have attracted more and more interest in image fusion [13].

In frequency domain techniques, each source image is first decomposed into a sequence of multiscale coefficients. Various fusion rules are then employed in the selection of these coefficients, which are synthesized via inverse transforms to form the fused image. Recently, series of frequency domain methods have been explored by using multiscale transform, including Laplacian pyramid transform, gradient pyramid transform, filter-subtract-decimate pyramid transform, discrete wavelet transform (DWT), and complex wavelet transform (CWT) [14–20]. There is evidence that multiscale transform based signal decomposition is similar to the human visual system. As we know the wavelet analysis, with its upstanding localize peculiarity in both time and frequency domain, has become one of the most commonly used methods in the fields of multiscale transform based image fusion [16]. However, wavelet analysis cannot effectively represent the line singularities and plane singularities of the images and thus cannot represent the directions of the edges of images accurately. To overcome these shortcomings of the wavelet transform, Do and Vetterli [17] proposed Contourlet transform which can give the asymptotic optimal representation of contours and has been successfully used for image fusion. However, the up- and downsampling process of Contourlet decomposition and reconstruction results in the Contourlet transform lacking shift-invariance and having pseudo-Gibbs phenomena in the fused image [19]. Later, da Cunha et al. [20] proposed the nonsubsampled Contourlet transform (NSCT) based on Contourlet transform. This method inherits the advantages of Contourlet transform, while possessing shift-invariance and effectively suppressing pseudo-Gibbs phenomena.

Although quite good results have been reported by NSCT based method, there is still much room to improve the fusion performance in the coefficient selection as follows.

(a) The low-frequency coefficients of the fused image can be simply acquired by averaging the low-frequency coefficients of the input images. This rule decreased contrast in the fused images [21] and cannot give the fused subimage of high quality for medical images.

(b) The popularly used larger absolute rule is implemented in the value of a single pixel of the current high-frequency subband. The disadvantage of this method is that the coefficients only know the value of a single pixel but not any of the relationship between the corresponding coefficients in high-frequency subbands [22].

(c) Most fusion rules of the NSCT-based methods are implemented in multifocus images [23], remote sensing images [24], and infrared and visible images [25]. The results are not of the same quality as those of the multimodal medical images. For example, Chai et al. [22] proposed a NSCT method based on features contrast of multiscale products to fuse multifocus images. However, it has been proven that this algorithm is not able to utilize prominent information present in the subbands efficiently and results in the poor quality when it is used to fuse multimodal medical images [26].

In this paper, a novel fusion framework based on NSCT is proposed for multimodal medical images. The main contribution of the method lies in the proposed fusion rule, which can capture the best membership of source images' coefficients to the corresponding fused coefficient. The phase congruency and Log-Gabor energy are unified as the fusion rules for low- and high-frequency coefficients, respectively. The phase congruency provides a contrast and brightness-invariant representation of low-frequency coefficients whereas Log-Gabor energy efficiently determines the frequency coefficients from the clear and detail parts in the high frequency. The combinations of these two techniques can preserve more details from source images and thus improve the quality of the fused images. Experiments indicate that the proposed framework can provide a better fusion outcome when compared to series of traditional image fusion methods in terms of both subjective and objective evaluations.

The rest of the paper is organized as follows. NSCT and phase congruency are described in Section 2 followed by the proposed multimodal medical image fusion framework in Section 3. Experimental results and discussions are given in Section 4 and the concluding remarks are presented in Section 5.

## 2. Preliminaries

This section provides the description of concepts on which the proposed framework is based. These concepts, including NSCT and phase congruency, are described as follows.

*2.1. Nonsubsampled Contourlet Transform (NSCT).* Contourlet transform can be divided into two stages [19]: Laplacian pyramid (LP) and directional filter bank (DFB) and offers an efficient directional multiresolution image representation. Among them, LP is first utilized to capture the point singularities and then followed by the DFB to link the singular point into linear structures. LP is used to decompose the original images into low-frequency and high-frequency subimages, and DFB divides the high-frequency subbands into directional subbands. The Contourlet decomposed schematic diagram is shown in Figure 1.
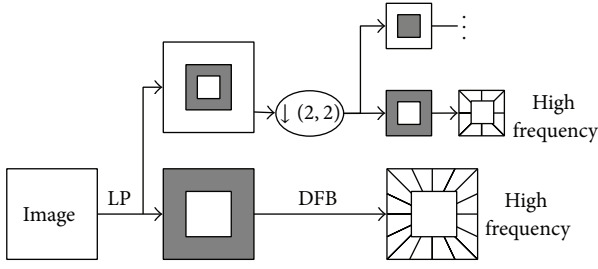
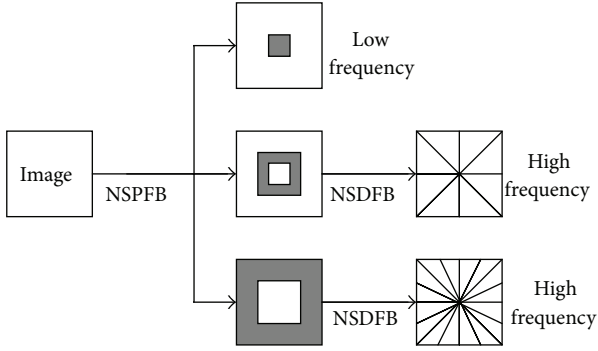FIGURE 1: Contourlet decomposed schematic diagram.



FIGURE 2: NSCT decomposed schematic diagram.

The NSCT is proposed based on the theory of Contourlet transform. NSCT inherits the advantage of Contourlet transform, enhances directional selectivity and shift-invariance, and effectively overcomes the pseudo-Gibbs phenomena. NSCT is built on nonsubsampled pyramid filter bank (NSPFB) and nonsubsampled directional filter bank (NSDFB) [21]. Figure 2 gives the NSCT decomposition framework with $k = 2$ levels.

The NSPFB ensures the multiscale performance by taking advantage of a two-channel nonsubsampled filter bank and one low-frequency subband and one high-frequency subband that can be produced at each decomposition level. The NSDFB is a two-channel nonsubsampled filter bank constructed by eliminating the downsamplers and upsamplers and combining the directional fan filter banks in the nonsubsampled directional filter [23]. NSDFB allows the direction decomposition with $l$ levels in each high-frequency subbands from NSPFB and then produces $2^l$ directional subbands with the same size as the source images. Thus, the NSDFB provides the NSCT the multidirection performance and offers more precise directional detail information to get more accurate results [23]. Therefore, NSCT leads to better frequency selectivity and an essential property of the shift-invariance on account of nonsubsampled operation. The size of different subimages decomposed by NSCT is identical. Additionally, NSCT-based image fusion can effectively mitigate the effects of misregistration on the results [27]. Therefore, NSCT is more suitable for image fusion.

*2.2. Phase Congruency.* Phase congruency is a feature perception approach which provides information that is invariant to image illumination and contrast [28]. This model is built on the Local Energy Model [29], which postulates that important features can be found at points where the Fourier components are maximally in phase. Furthermore, the angle at which the phase congruency occurs signifies the feature type. The phase congruency can be used for feature detection [30]. The model provides useful feature localization and noise compensation. The phase congruency at a point $(i, j)$ can be defined as follows [31]:

$$
\begin{aligned}
\mathrm{PC}^\theta & (i, j) \\
& = \sum_n W^\theta(i, j) \left[ A_n^\theta(i, j) \left( \cos \left( \phi_n^\theta(i, j) - \overline{\phi}_n^\theta(i, j) \right) \right) \right. \\
& \qquad \left. - \left| \sin \left( \phi_n^\theta(i, j) - \overline{\phi}_n^\theta(i, j) \right) \right| - T \right]_+ \\
& \quad \times \left( \sum_n A_n^\theta(i, j) + \varepsilon \right)^{-1},
\end{aligned}
\tag{1}
$$

where $\theta$ is the orientation, $W^\theta(i, j)$ is the weighting factor based on frequency spread, $A_n^\theta(i, j)$ and $\overline{\phi}_n^\theta(i, j)$ are the amplitude and phase for wavelet scale $n$, respectively, $\phi_n^\theta(i, j)$ is the weighted mean phase, $T$ is a noise threshold constant, and $\varepsilon$ is a small constant value to avoid division by zero. The notation $[ ]_+$ denotes that the enclosed quantity is equal to itself when the value is positive, and zero otherwise. For details of phase congruency measure see [29].

As we know the multimodal medical images have the following characteristics:

(a) the images of different modal have significantly different pixel mappings;

(b) the capturing environment of different modalities varies and resulted in the change of illumination and contrast;

(c) the edges and corners in the images are identified by collecting frequency components of the image that is in phase.

According to the literature [26, 32], it is easy to find that phase congruency is not only invariant to different pixel intensity mappings, illumination, and contrast changes, but also gives the Fourier components that are maximally in phase. These all will lead to efficient fusion. That is why we use phase congruency for multimodal medical fusion.

## 3. Proposed Multimodal Medical Image Fusion Framework

The framework of the proposed multimodal medical image fusion algorithm is depicted in Figure 3, but before describing it, the definition of local Log-Gabor energy in NSCT domain is first described as follows.

*3.1. Log-Gabor Energy in NSCT Domain.* The high-frequency coefficients in NSCT domain represent the detailed components of the source images, such as the edges, textures, and region boundaries [21]. In general, the coefficients with larger absolute values correspond to the sharper brightness
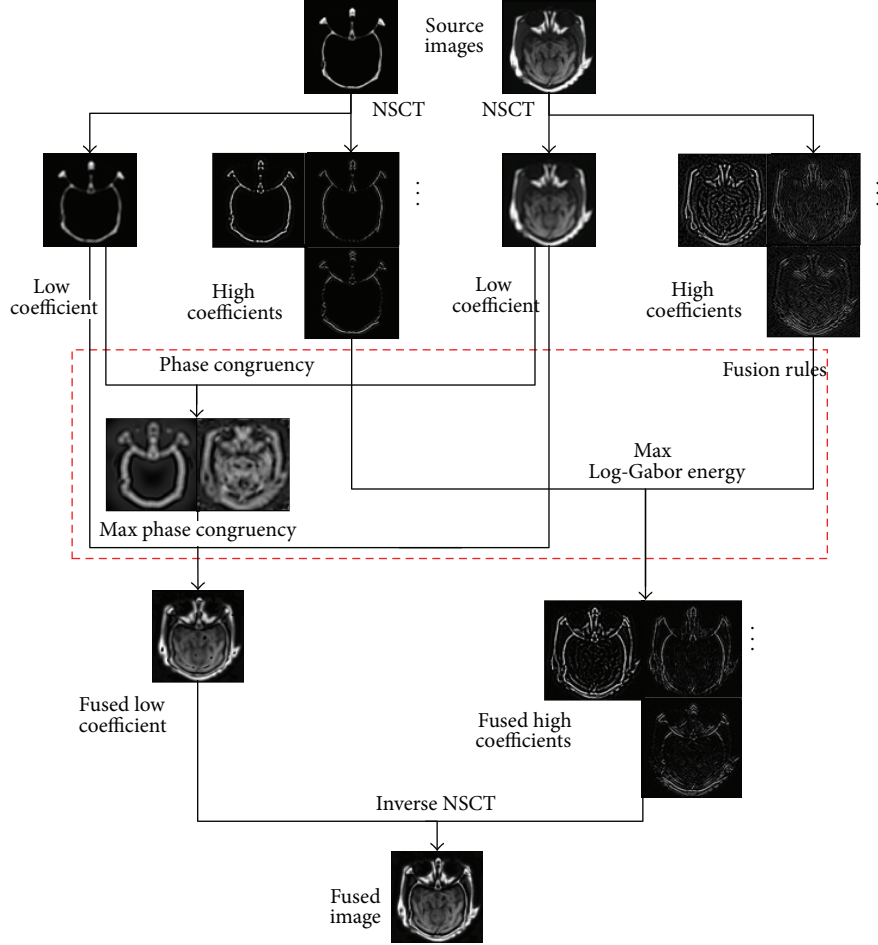
FIGURE 3: The framework of the proposed fusion algorithm.

in the image. It is to be noted that the noise is also related to high-frequency coefficients and may cause miscalculation of sharpness values and, therefore, affect the fusion performance [26]. Furthermore, human visual system is generally more sensitive to texture detail features than the value of a single pixel.

To overcome the defect mentioned above, a novel high-frequency fusion rule based on local Log-Gabor energy is designed in this paper. Gabor wavelet is a popular technique that has been extensively used to extract texture features [33]. Log-Gabor filters are proposed based on Gabor filters. Compared with Gabor filters, Log-Gabor filters cover the shortage of the high frequency of Gabor filter component expression and more in accord with human visual system [34]. Therefore, Log-Gabor wavelet can achieve optimal spatial orientation and wider spectrum information at the same time and thus more truly reflect the frequency response of the natural images and improve the performance in terms of the accuracy [35].

Under polar coordinates, the Log-Gabor wavelet is expressed as follows [36]:

$$g\left(f,\theta\right) = \exp\left\{-\frac{\left[\ln\left(f/f_0\right)\right]^2}{2\left[\ln\left(\sigma/f_0\right)\right]^2}\right\} \times \exp\left\{-\frac{\left(\theta-\theta_0\right)^2}{2\sigma_\theta^2}\right\},\tag{2}$$

in which $f_0$ is the center frequency of the Log-Gabor filter, $\theta_0$ is the direction of the filter, $\sigma$ is used to determine the bandwidth of the radial filter, and $\sigma_\theta$ is used to determine the bandwidth of the orientation. If $g_{kl}^{uv}(i,j)$ correspond to Log-Gabor wavelets in scale $u$ and direction $v$, the signal response is expressed as follows:

$$G_{kl}^{uv}\left(i,j\right) = H_{kl}\left(i,j\right) * g_{kl}^{uv}\left(i,j\right),\tag{3}$$

where $H_{kl}(i,j)$ is the coefficient located at $(i,j)$ in high-frequency subimages of the source image $A$ or $B$ at the $k$th scale, $l$th direction, and $*$ denotes convolution operation. The Log-Gabor energy of high-frequency subimages at the $k$th scale, $l$th direction, is expressed as follows:

$$E_{kl}\left(i,j\right) = \sum_{u=1}^{U}\sum_{v=1}^{V}\sqrt{\mathrm{real}(G_{kl}^{uv}\left(i,j\right))^2 + \mathrm{imag}(G_{kl}^{uv}\left(i,j\right))^2},\tag{4}$$

in which, $\mathrm{real}(G_{kl}^{uv}(i,j))$ is the real part of $G_{kl}^{uv}(i,j)$ and $\mathrm{imag}(G_{kl}^{uv}(i,j))$ is the imaginary part of $G_{kl}^{uv}(i,j)$. The Log-Gabor energy in NSCT domain at the local area around the pixel $(i,j)$ is given as

$$\mathrm{LE}_{kl}\left(i,j\right) = \frac{1}{\left(2M+1\right)\left(2N+1\right)}\sum_{m=-M}^{M}\sum_{n=-N}^{N}E_{kl}\left(i+m,j+n\right),\tag{5}$$

in which $(2M + 1)(2N + 1)$ is the window size. The proposed definition of the local Log-Gabor energy not only extracts more useful features from high-frequency coefficients, but also keeps a well performance in noisy environment.

*3.2. Proposed Fusion Framework.* The proposed NSCT-based image fusion framework is discussed in this subsection. Considering the input multimodal medical images ($A$ and $B$) are perfectly registered. The framework of the proposed fusion method is shown Figure 3 and described as the following three steps.

*Step 1.* Perform $\theta$-level NSCT on $A$ and $B$ to obtain one low-frequency subimage and series of high-frequency subimages at each level and direction $l$; that is, $A : \{L^A, H^A_{k,l}\}$ and $B : \{L^B, H^B_{k,l}\}$, where $L^A, L^B$ are the low-frequency subimages and $H^A_{k,l}, H^A_{k,l}$ represent the high-frequency subimages at level $k \in [1, \theta]$ in the orientation $l$.

*Step 2.* Fuse low- and high-frequency subbands via the following novel fusion rule to obtain composite low- and high-frequency subbands.

The low-frequency coefficients represent the approximation component of the source images. The popular widely used approach is to apply the averaging methods to produce the fused coefficients. However, this rule reduced contrast in the fused images and cannot give the fused subimage of high quality for medical image. Therefore, the criterion based on the phase congruency that is introduced in Section 2.2 is employed to fuse the low-frequency coefficients. The fusion rule for the low-frequency subbands is defined as

$$
\begin{aligned}
&L^F (i, j) \\
&= \begin{cases} L^A (i, j), & \text{if } \mathrm{PC}^{A\theta} (i, j) > \mathrm{PC}^{B\theta} (i, j) \\ L^B (i, j), & \text{if } \mathrm{PC}^{A\theta} (i, j) < \mathrm{PC}^{B\theta} (i, j) \\ 0.5 \left( L^A (i, j) + L^B (i, j) \right), & \text{if } \mathrm{PC}^{A\theta} (i, j) = \mathrm{PC}^{B\theta} (i, j), \end{cases}
\end{aligned}
\tag{6}
$$

where $\mathrm{PC}^{A\theta} (i, j), \mathrm{PC}^{B\theta} (i, j)$ is the phase congruency extracted from low-frequency subimages of the source images $A$ and $B$, respectively.

For the high-frequency coefficients, the most common fusion rule is selecting the coefficient with larger absolute values. This rule does not take any consideration of the surrounding pixels and cannot give the fused components of high quality for medical image. Especially when the source images contain noise, the noise could be mistaken for fused coefficients and cause miscalculation of the sharpness value. Therefore, the criterion based on Log-Gabor energy is introduced to fuse high-frequency coefficients. The fusion rule for the high-frequency subbands is defined as

$$
H^F_{kl} (i, j) = \begin{cases} H^A_{kl} (i, j), & \text{if } \mathrm{LE}^A_{kl} (i, j) \geq \mathrm{LE}^B_{kl} (i, j) \\ H^B_{kl} (i, j), & \text{otherwise,} \end{cases}
\tag{7}
$$

where $\mathrm{LE}^A_{kl}(i, j)$ and $\mathrm{LE}^B_{kl}(i, j)$ are the local Log-Gabor energy extracted from high-frequency subimages at the $k$th scale and $l$th direction of source images $A$ and $B$, respectively.

*Step 3.* Perform $\theta$-level by the inverse NSCT on the fused low- and high-frequency subimages. The fused image is obtained ultimately in this way.

# 4. The Experimental Results and Analysis

It is well known that different image quality metrics imply the visual quality of images from different aspects, but none of them can directly imply the quality. In this paper, we consider both the visual representation and quantitative assessment of the fused images. For evaluation of the proposed fusion method, we have considered five separate fusion performance metrics defined as below.

## 4.1. Evaluation Index System

*4.1.1. Standard Deviation.* The standard deviation (STD) of an image with size of $M_0 \times N_0$ is defined as [37]:

$$
\mathrm{Std} = \left( \frac{1}{M_0 \times N_0} \sum_{i=1}^{M_0} \sum_{j=1}^{N_0} (F (i, j) - \widehat{\mu})^2 \right)^{1/2},
\tag{8}
$$

where $F(i, j)$ is the pixel value of the fused image at the position $(i, j)$ and $\widehat{\mu}$ is the mean value of the image. The STD can be used to estimate how widely the gray values spread in an image. The larger the STD, the better the result.

*4.1.2. Edge Based Similarity Measure.* The edge based similarity measure ($Q^{AB/F}$) is proposed by Xydeas and Petrović [38]. The definition is given as

$$
\begin{aligned}
&Q^{AB/F} \\
&= \frac{\sum_{i=1}^{M_0} \sum_{j=1}^{N_0} \left[ Q^{AF} (i, j) w^A (i, j) + Q^{BF} (i, j) w^B (i, j) \right]}{\sum_{i=1}^{M_0} \sum_{j=1}^{N_0} \left[ w^A (i, j) + w^B (i, j) \right]},
\end{aligned}
\tag{9}
$$

where $w^A(i, j)$ and $w^B(i, j)$ are the corresponding gradient strength for images $A$ and $B$, respectively. The definition of $Q^{AF}(i, j)$ and $Q^{BF}(i, j)$ is given as

$$
\begin{aligned}
Q^{AF} (i, j) &= Q^{AF}_a (i, j) Q^{AF}_g (i, j), \\
Q^{BF} (i, j) &= Q^{BF}_a (i, j) Q^{BF}_g (i, j),
\end{aligned}
\tag{10}
$$

where $Q^{xF}_a(i, j)$ and $Q^{xF}_g(i, j)$ are the edge strength and orientation preservation values at location for image $x$ ($A$ or $B$), respectively. The edge based similarity measure gives the similarity between the edges transferred from the input images to the fused image [26]. The larger the value, the better the fusion result.

*4.1.3. Mutual Information.* Mutual information (MI) [39] between the fusion image $F$ and the source images $A$ and $B$ is defined as follows:

$$\text{MI} = \text{MI}^{AF} + \text{MI}^{BF},$$

$$\text{MI}^{AF} = \sum_{f=0}^{L}\sum_{a=0}^{L} p^{AF}(a,f) \log_2\left(\frac{p^{AF}(a,f)}{p^A(a)\, p^F(f)}\right),$$ (11)

$$\text{MI}^{BF} = \sum_{f=0}^{L}\sum_{b=0}^{L} p^{BF}(b,f) \log_2\left(\frac{p^{BF}(b,f)}{p^B(b)\, p^F(f)}\right),$$

where $\text{MI}^{xF}$ denotes the normalized mutual information between the fused image and the input image $A$, $B$; $a$, $b$ and $f \in [0, L]$, and $L$ is the number of bins. $p^A(a)$, $p^B(b)$, and $p^F(f)$ are the normalized gray level histograms of source images and fused image. $p^{xF}(a,f)$ is the joint gray level histograms between fused image and each source image.

MI can indicate how much information the fused image $F$ conveys about the source images $A$ and $B$ [22]. Therefore, the greater the value of MI, the better the fusion effect.

*4.1.4. Cross Entropy.* The cross entropy is defined as [8]:

$$\text{CE} = \sum_{l=0}^{L-1} P_l \log_2 \frac{P_l}{Q_l},$$ (12)

where $P_l$ and $Q_l$ denote the gray level histogram of the source image and the fused image, respectively. The cross entropy is used to evaluate the difference between the source images and the fused image. The lower value corresponds to the better fusion result.

*4.1.5. Spatial Frequency.* Spatial frequency is defined as [40]:

$$\text{SF} = \sqrt{\text{RF}^2 + \text{CF}^2},$$ (13)

where RF and CF are the row frequency and column frequency, respectively, and are defined as

$$\text{RF} = \sqrt{\frac{1}{M_0 N_0} \sum_{i=0}^{M_0-1}\sum_{j=0}^{N_0-1} [F(i,j) - F(i,j-1)]^2},$$

$$\text{CF} = \sqrt{\frac{1}{M_0 N_0} \sum_{i=0}^{M_0-1}\sum_{j=0}^{N_0-1} [F(i,j) - F(i-1,j)]^2}.$$ (14)

The spatial frequency reflects the edge information of the fused image. Larger spatial frequency values indicate better image quality.

*4.2. Experiments on Multimodal Medical Image Fusion.* To evaluate the performance of the proposed image fusion approach, the experiments are performed on three groups of multimodal medical images. These images are characterized in two distinct pairs: (1) CT and MRI; (2) MR-T1 and MR-T2. The images in Figures 4(a1)-4(b1) and 4(a2)-4(b2) are

CT and MRI images, whereas Figures 4(a3)-4(b3) are T1-weighted MR image (MR-T1) and T2-weighted MR image (MR-T2). All images have the same size of $256 \times 256$ pixel, with 256-level gray scale. For all these image groups, the results of the proposed fusion framework are compared with those of the traditional discrete wavelet transform (DWT) [13, 16], the second generation curvelet transform (fast discrete curvelet transform, FDCT) [41, 42], the dual tree complex wavelet transform (DTCWT) [4], and the nonsubsampled Contourlet transform (NSCT-1 and NSCT-2) based methods. The high-frequency coefficients and low-frequency coefficients of DWT, FDCT, DTCWT, and NSCT-1 based methods are merged by the popular widely used fusion rule of selecting the coefficient with larger absolute values and the averaging rule (average-maximum rule), respectively. NSCT-2 based method is merged by the fusion rules proposed by Bhatnagar, et al. in [26]. In order to perform a fair comparison, the source images are all decomposed into the same levels with 3 for those methods except FDCT method. For DWT method, the images are decomposed using the DBSS (2, 2) wavelet. For implementing NSCT, "9-7" filters and "pkva" filters (how to set the filters can be seen in [43]) are used as the pyramidal and directional filters, respectively.

*4.2.1. Subjective Evaluation.* The first pair of medical images are two groups of brain CT and MRI images on different aspects, shown in Figures 4(a1), 4(b1) and 4(a2), 4(b2), respectively. It can be easily seen that the CT image shows the dense structure while MRI provides information about soft tissues. The obtained fused images from DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 are shown in Figures 4(c1)–4(g1) and 4(c2)–4(g2), respectively. The results for the proposed fusion method have been shown in Figures 4(h1) and 4(h2). On comparing these results, it can be easily observed that the proposed method outperforms those fusion methods and has good visual representation of fused image.

The second pair of medical images are MR-T1 and MR-T2 images, shown in Figures 4(a3) and 4(b3). The comparison of DWT, FDCT, DTCWT, NSCT-1, NSCT-2, and proposed method, shown in Figures 4(c3)–4(h3), clearly implies that the fusion result of the proposed method has better quality and contrast in comparison to other methods.

Similarly, on observing the noticeable improvement has been emphasized in Figure 4 by the red arrows and the analysis above, one can easily verify the fact that again the proposed method has been found superior in terms of visual representation over DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 fusion methods.

*4.2.2. Objective Evaluation.* For objective evaluation of the fusion results, shown in Figure 4, we have used five fusion metrics: cross entropy, spatial frequency, STD, $Q^{AB/F}$, and MI. The quantitative comparison of cross entropy and spatial frequency for these fused images is visually given by Figures 5-6 and other metrics are given by Table 1.

On observing Figure 5, one can easily observe all the three results of the proposed scheme have lower values of cross entropy than any of the DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 fusion methods. The cross entropy is used to
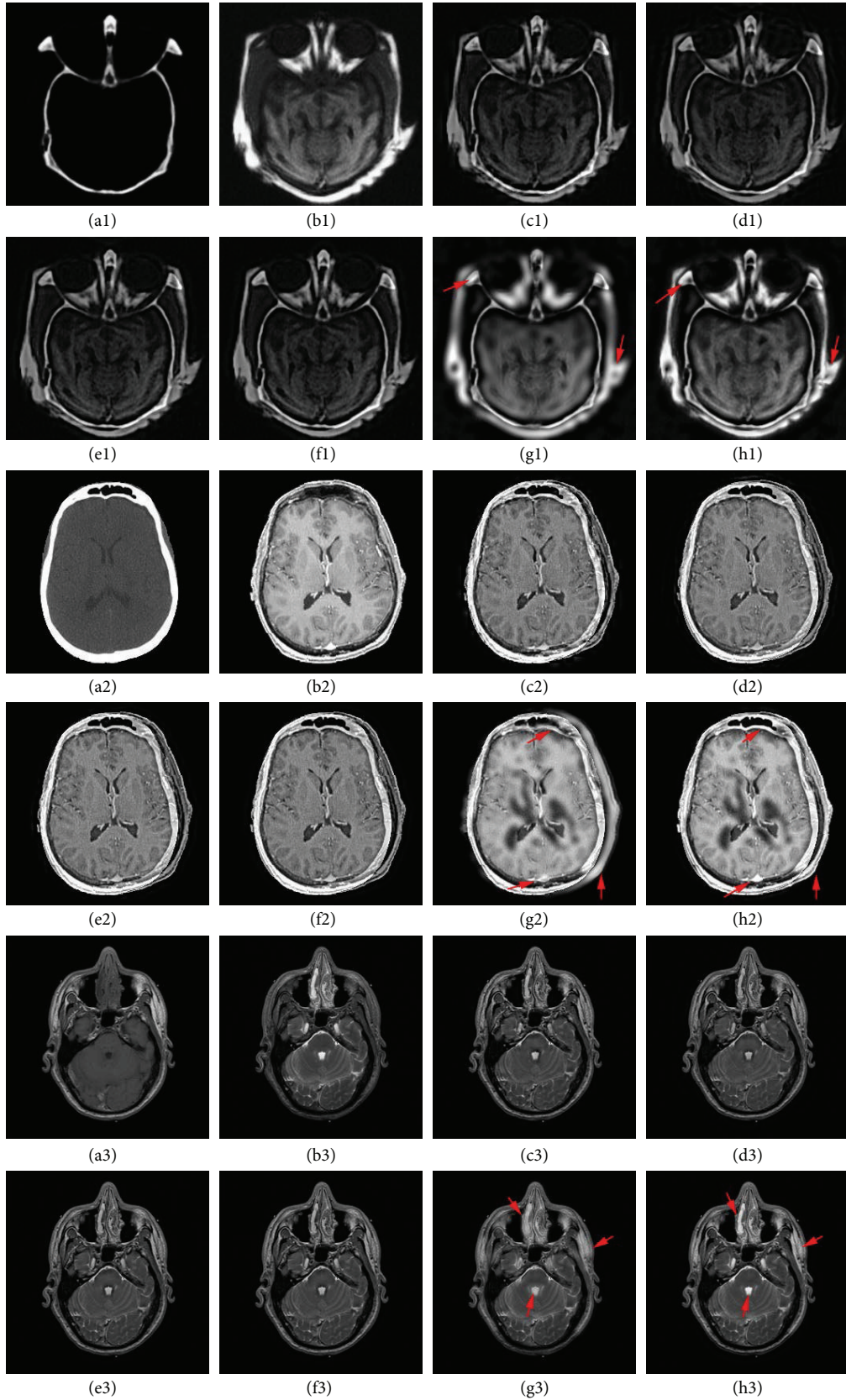
Figure 4: Source multimodal medical images: (a1), (b1) image group 1 (CT and MRI); (a2), (b2) image group 2 (CT and MRI); (a3), (b3) image group 3 (MR-T1 and MR-T2); fused images from (c1), (c2), (c3) DWT based method; (d1), (d2), (d3) FDCT based method; (e1), (e2), (e3) DTCWT based method; (f1), (f2), (f3) NSCT-1 based method; (g1), (g2), (g3) NSCT-2 based method; (h1), (h2), (h3) our proposed method.

TABLE 1: Comparison on quantitative evaluation of different methods for the first set medical images.

| Images | Evaluation | DWT | FDCT | DTCWT | NSCT-1 | NSCT-2 | Proposed method |
|---|---|---|---|---|---|---|---|
| Image group 1 (CT and MRI) | STD | 41.160 | 39.170 | 40.132 | 40.581 | 54.641 | **58.476** |
| | $Q^{AB/F}$ | 0.618 | 0.592 | 0.652 | 0.683 | 0.427 | **0.716** |
| | MI | 2.371 | 2.029 | 2.393 | 2.438 | 1.886 | **2.580** |
| Image group 2 (CT and MRI) | STD | 72.260 | 70.572 | 71.407 | 71.733 | 80.378 | **83.334** |
| | $Q^{AB/F}$ | 0.567 | 0.584 | 0.606 | **0.622** | 0.450 | 0.618 |
| | MI | 3.137 | 3.138 | 3.198 | 3.235 | 2.915 | **3.302** |
| Image group 3 (MR-T1 and MR-T2) | STD | 37.358 | 36.856 | 37.358 | 37.277 | 39.023 | **40.630** |
| | $Q^{AB/F}$ | 0.635 | 0.647 | 0.668 | 0.682 | 0.589 | **0.689** |
| | MI | 3.209 | 3.142 | 3.302 | 3.347 | 3.349 | **3.469** |



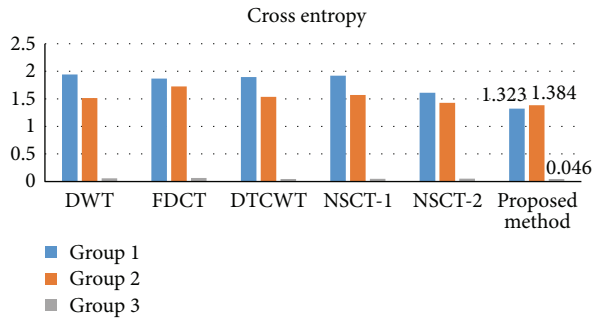FIGURE 5: Comparison on cross entropy of different methods and images.



FIGURE 6: Comparison on spatial frequencies of different methods and images.

evaluate the difference between the source images and the fused image. Therefore, the lower value corresponds to the better fusion result.

On observing Figure 6, two values of the spatial frequency of the fused image obtained by the proposed method are the highest, and the other one is 6.447 which is close to the highest value 6.581. Observation of Table 1 yields that all the three results of the proposed fusion scheme have higher values of STD, $Q^{AB/F}$, and MI than any of other methods except one value of $Q^{AB/F}$ (image group 2) is the second best. However, an overall comparison shows the superiority of the proposed fusion scheme.

*4.2.3. Combined Evaluation.* Since the subjective and objective evaluations separately are not able to examine fusion results, we have combined them. From these figures (Figures 4–6) and table (Table 1), it is clearly to find that the proposed method not only preserves most of the source images characteristics and information, but also improves the definition and the spatial quality better than the existing methods, which can be justified by the optimum values of objective criteria except one value of spatial frequency (image group 3) and one value of $Q^{AB/F}$ (image group 2). Consider the example of the first set of images: the five criteria values of the proposed method are 1.323 (cross entropy), 7.050 (spatial frequency), 58.476 (STD), 0.716 ($Q^{AB/F}$), and 2.580 (MI), respectively. Each of them is the optimal one in the first set of experiments.
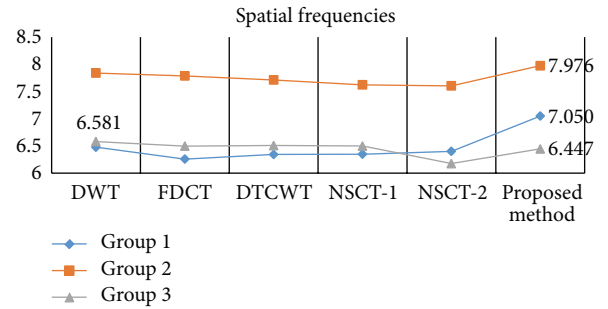
Among these methods, the result of NSCT-2 based method also gives poor results when comparing to the proposed NSCT-based method. This stems from the fact that high-frequency fusion rule of NSCT-2 based method is not able to extract the detail information in the high frequency effectively. Also, by carefully looking at the outputs of the proposed NSCT-based method (Figures 4(h1), 4(h2), and 4(h3)), we can find that they get more contrast and more spatial resolution than the outputs of NSCT-2 based method (highlighted by the red arrows) and other methods. The main reason behind the better performance is that the proposed fusion rules for low- and high-frequency coefficients can effectively extract prominent and detail information from the source images. Therefore, it can be possible to conclude that the proposed method is better than the existing methods.

*4.3. Fusion of Multimodal Medical Noisy Images and a Clinical Example.* To evaluate the performance of the proposed method in noisy environment, the input image group 1 has been additionally corrupted with Gaussian noise, with a standard deviation of 5% (shown in Figures 7(a) and 7(b)). In addition, a clinical applicability on noninvasive diagnosis of neoplastic disease is given in the last subsection.

*4.3.1. Fusion of Multimodal Medical Noisy Images.* For comparison, apart from visual observation, objective criteria on STD, MI, and $Q^{AB/F}$ are used to evaluate how much clear or detail information of the source images is transferred to the fused images. However, maybe these criteria cannot

(a)　　　　　　(b)　　　　　　(c)　　　　　　(d)
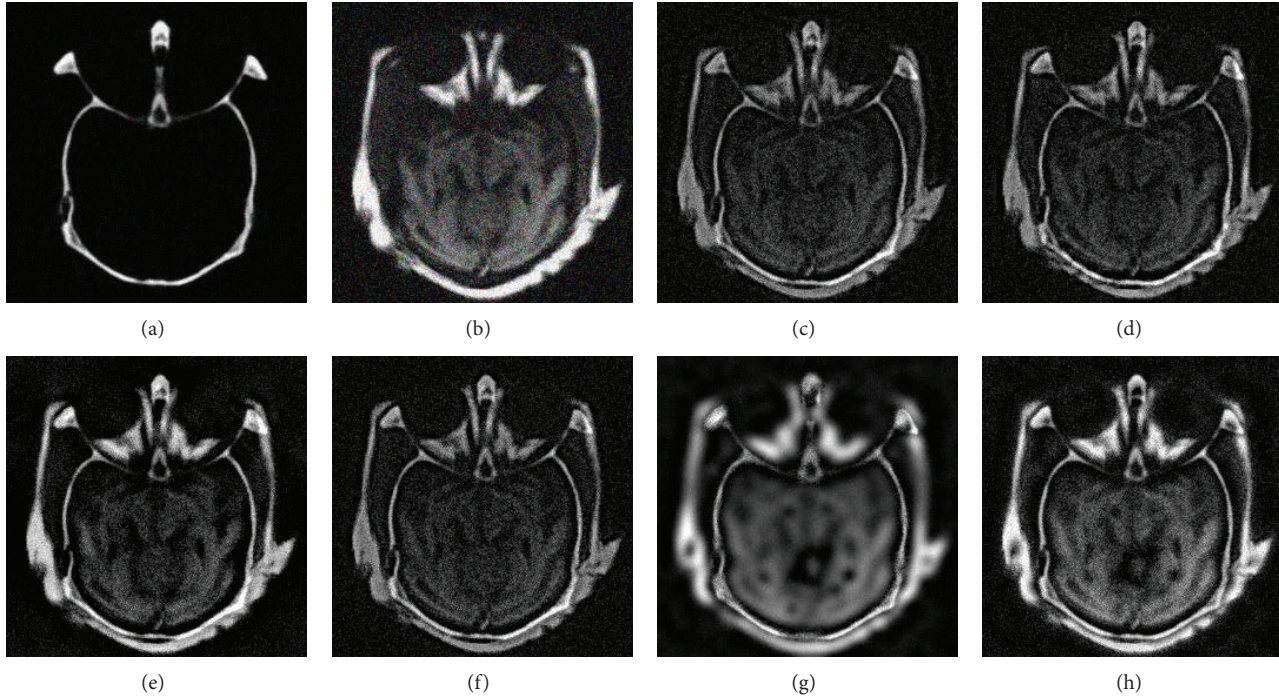
(e)　　　　　　(f)　　　　　　(g)　　　　　　(h)

FIGURE 7: The multimodal medical images with noise: (a) CT image with 5% noise, (b) MRI image with 5% noise; fused images by (c) DWT, (d) FDCT, (e) DTCWT, (f) NSCT-1, (g) NSCT-2, and (h) proposed method.

TABLE 2: Comparison on quantitative evaluation of different methods for the noise medical images.

| Evaluation | DWT | FDCT | DTCWT | NSCT-1 | NSCT-2 | Proposed method |
|---|---|---|---|---|---|---|
| STD | 41.893 | 40.044 | 40.995 | 41.338 | 52.533 | **57.787** |
| $Q^{AB/F}$ | 0.592 | 0.579 | 0.619 | 0.637 | 0.269 | **0.640** |
| MI | 2.785 | 2.539 | 2.826 | 2.882 | 1.816 | **2.914** |

effectively evaluate the performance of the fusion methods in terms of the noise transmission. For further comparison, Peak Signal to Noise Ratio (PSNR), a ratio between the maximum possible power of a signal and the power of noise that affects the fidelity [44], is used. The larger the value of PSNR, the less the image distortion [45]. PSNR is formulated as

$$\text{PSNR} = 10\lg\left|\frac{255^2}{\text{RMSE}^2}\right|, \tag{15}$$

where RMSE denotes the Root Mean Square Error between the fused image and the reference image. The reference image in the following experiment is selecting from Figure 4(h1), which is proven to be the best performance compared to other images.

Figure 7 illustrates the fusion results obtained by the different methods. The comparison of the images fused by DWT, FDCT, DTCWT, NSCT-1, NSCT-2, and proposed method, shown in Figures 7(c)–7(h), clearly implies that the fused image by proposed method has better quality and contrast than other methods. Figure 8 shows the values of PSNR of different methods in fusing noisy images. One can observe that the proposed method has higher values of
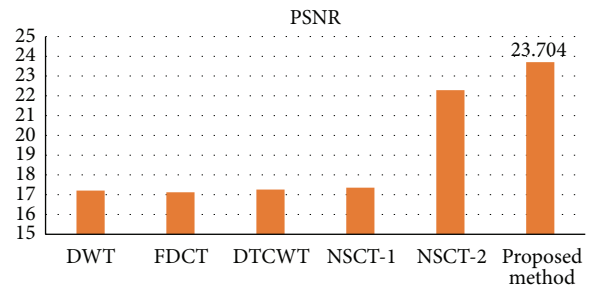


FIGURE 8: Comparison on PSNR of different methods for the noise medical images.

PSNR than any of the DWT, FDCT, DTCWT, NSCT-1, and NSCT-2 fusion methods. Table 2 gives the quantitative results of fused images and shows that the values of STD, $Q^{AB/F}$, and MI are also the highest of all the six methods. From the analysis above, we can also observe that the proposed scheme provides the best performance and outperforms the other algorithms. In addition, compared with the result of the NSCT-1 method using the average-maximum rule, it
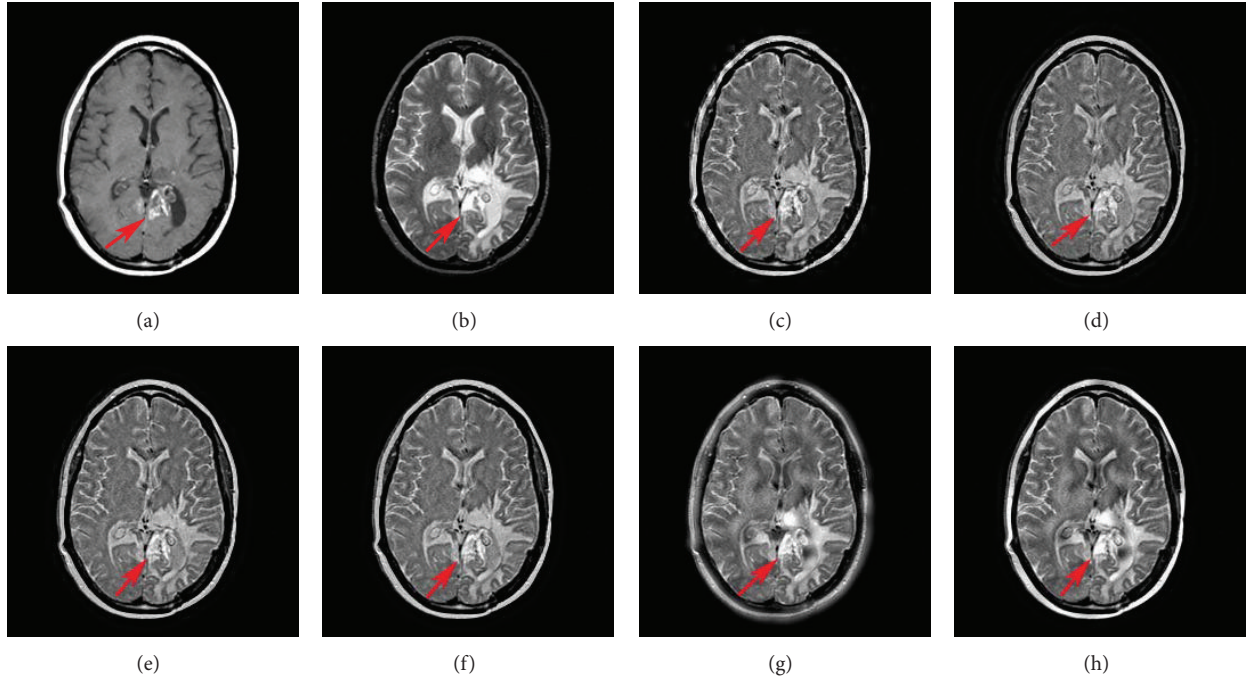
FIGURE 9: Brain images of the man with recurrent tumor: (a) MR-T1 image, (b) MR-T2 image; fused images by (c) DWT, (d) FDCT, (e) DTCWT, (f) NSCT-1, (g) NSCT-2, and (h) proposed method.

TABLE 3: Comparison on quantitative evaluation of different methods for the clinical medical images.

| Evaluation | DWT | FDCT | DTCWT | NSCT-1 | NSCT-2 | Proposed method |
|---|---|---|---|---|---|---|
| STD | 64.433 | 62.487 | 63.194 | 63.652 | 60.722 | **67.155** |
| $Q^{AB/F}$ | 0.541 | 0.565 | 0.578 | 0.597 | 0.411 | **0.624** |
| MI | 2.507 | 2.455 | 2.540 | 2.584 | 2.392 | **2.624** |

demonstrated the validities of the proposed fusion rule in noisy environment.

*4.3.2. A Clinical Example on Noninvasive Diagnosis.* In order to demonstrate the practical value of the proposed scheme in medical imaging, one clinical case on neoplastic diagnosis is considered where MR-T1/MR-T2 medical modalities are used. The images have been downloaded from the Harvard University site (http://www.med.harvard.edu/AANLIB/home .html). Figures 9(a)-9(b) show the recurrent tumor case of a 51-year-old woman who sought medical attention because of gradually increasing right hemiparesis (weakness) and hemi-anopia (visual loss). At craniotomy, left parietal anaplastic astrocytoma was found. A right frontal lesion was biopsied. A large region of mixed signal on MR-T1 and MR-T2 images gives the signs of the possibility of active tumor (highlighted by the red arrows).

Figures 9(c)–9(h) show the fused images by DWT, FDCT, DTCWT, NSCT-1, NSCT-2, and proposed method. It is obvious that the fused image by proposed method has better contrast and sharpness of active tumor (highlighted by the red arrows) than other methods. Table 3 shows the quantitative evaluation of different methods for the clinical medical images. The values of the proposed method are optimum in

terms of STD, MI, and $Q^{AB/F}$. From Figure 9 and Table 3, we can obtain the same conclusion that the proposed scheme provides the best performance and outperforms the other algorithms.

## 5. Conclusions

Multimodal medical image fusion plays an important role in clinical applications. But the real challenge is to obtain a visu-ally enhanced image through fusion process. In this paper, a novel and effective image fusion framework based on NSCT and Log-Gabor energy is proposed. The potential advantages include (1) NSCT is more suitable for image fusion because of its advantages such as multiresolution, multidirection, and shift-invariance; (2) a new couple of fusion rules based on phase congruency and Log-Gabor energy are used to preserve more useful information in the fused image to improve the quality of the fused images and overcome the limitations of the traditional fusion rules; and (3) the proposed method can provide a better performance than the current fusion methods whatever the source images are clean or noisy. In the experiments, five groups of multimodal medical images, including one group with noise and one group clinical example of a woman affected with recurrent tumor, are

fused by using traditional fusion methods and the proposed framework. The subjective and objective comparisons clearly demonstrate that the proposed algorithm can enhance the details of the fused image and can improve the visual effect with less information distortion than other fusion methods. In the future, we plan to design a pure C++ platform to reduce the time cost and extend our method for 3D or 4D medical image fusion.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

## References

[1] R. Shen, I. Cheng, and A. Basu, "Cross-scale coefficient selection for volumetric medical image fusion," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 1069–1079, 2013.

[2] V. D. Calhoun and T. Adali, "Feature-based fusion of medical imaging data," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 711–720, 2009.

[3] Y. Yang, D. S. Park, S. Huang, and J. Yang, "Fusion of CT and MR images using an improved wavelet based method," *Journal of X-Ray Science and Technology*, vol. 18, no. 2, pp. 157–170, 2010.

[4] R. Singh, R. Srivastava, O. Prakash, and A. Khare, "Multimodal medical image fusion in dual tree complex wavelet transform domain using maximum and average fusion rules," *Journal of Medical Imaging and Health Informatics*, vol. 2, no. 2, pp. 168–173, 2012.

[5] Z. Wang and Y. Ma, "Medical image fusion using $m$-PCNN," *Information Fusion*, vol. 9, no. 2, pp. 176–185, 2008.

[6] V. Barra and J. Boire, "A general framework for the fusion of anatomical and functional medical images," *NeuroImage*, vol. 13, no. 3, pp. 410–424, 2001.

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

[8] P. Balasubramaniam and V. P. Ananthi, "Image fusion using intuitionistic fuzzy sets," *Information Fusion*, vol. 20, pp. 21–30, 2014.

[9] Y. Yang, S. Y. Huang, F. J. Gao et al., "Multi-focus image fusion using an effective discrete wavelet transform based algorithm," *Measurement Science Review*, vol. 14, no. 2, pp. 102–108, 2014.

[10] P. S. Pradhan, R. L. King, N. H. Younan, and D. W. Holcomb, "Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 12, pp. 3674–3686, 2006.

[11] S. Li, J. T. Kwok, and Y. Wang, "Multifocus image fusion using artificial neural networks," *Pattern Recognition Letters*, vol. 23, no. 8, pp. 985–997, 2002.

[12] J. Yang and R. S. Blum, "A statistical signal processing approach to image fusion for concealed weapon detection," in *Proceedings of the International Conference on Image Processing (ICIP '02)*, vol. 1, pp. 513–516, September 2002.

[13] R. Singh and A. Khare, "Multiscale medical image fusion in wavelet domain," *The Scientific World Journal*, vol. 2013, Article ID 521034, 10 pages, 2013.

[14] B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, and I. Pippi, "Multispectral fusion of multisensor image data by the generalized Laplacian pyramid," in *Proceedings of the 1999 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '99)*, vol. 2, pp. 1183–1185, Hamburg, Germany, July 1999.

[15] V. S. Petrović and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 228–237, 2004.

[16] Y. Yang, D. S. Park, S. Huang, and N. Rao, "Medical image fusion via an effective wavelet-based approach," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, Article ID 579341, 2010.

[17] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.

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

[19] X. B. Qu, J. W. Yan, and G. D. Yang, "Multifocus image fusion method of sharp frequency localized Contourlet transform domain based on sum-modified-Laplacian," *Optics and Precision Engineering*, vol. 17, no. 5, pp. 1203–1212, 2009.

[20] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.

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

[22] Y. Chai, H. Li, and X. Zhang, "Multifocus image fusion based on features contrast of multiscale products in nonsubsampled contourlet transform domain," *Optik*, vol. 123, no. 7, pp. 569–581, 2012.

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

[24] Y. Wu, C. Wu, and S. Wu, "Fusion of remote sensing images based on nonsubsampled contourlet transform and region segmentation," *Journal of Shanghai Jiaotong University (Science)*, vol. 16, no. 6, pp. 722–727, 2011.

[25] J. Adu, J. Gan, and Y. Wang, "Image fusion based on nonsubsampled contourlet transform for infrared and visible light image," *Infrared Physics & Technology*, vol. 61, pp. 94–100, 2013.

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

[27] S. Das and M. K. Kundu, "NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency," *Medical and Biological Engineering and Computing*, vol. 50, no. 10, pp. 1105–1114, 2012.

[28] P. Kovesi, "Image features from phase congruency," *Videre*, vol. 1, no. 3, pp. 1–26, 1999.

[29] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, vol. 6, no. 5, pp. 303–313, 1987.

[30] P. Kovesi, "Phase congruency: a low-level image invariant," *Psychological Research*, vol. 64, no. 2, pp. 136–148, 2000.

[31] L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[32] A. Wong and W. Bishop, "Efficient least squares fusion of MRI and CT images using a phase congruency model," *Pattern Recognition Letters*, vol. 29, no. 3, pp. 173–180, 2008.

[33] L. Yu, Z. He, and Q. Cao, "Gabor texture representation method for face recognition using the Gamma and generalized Gaussian models," *Image and Vision Computing*, vol. 28, no. 1, pp. 177–187, 2010.

[34] P. Yao, J. Li, X. Ye, Z. Zhuang, and B. Li, "Iris recognition algorithm using modified Log-Gabor filters," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, pp. 461–464, August 2006.

[35] R. Raghavendra and C. Busch, "Novel image fusion scheme based on dependency measure for robust multispectral palm-print recognition," *Pattern Recognition*, vol. 47, no. 6, pp. 2205–2221, 2014.

[36] R. Redondo, F. Šroubek, S. Fischer, and G. Cristóbal, "Multifocus image fusion using the log-Gabor transform and a Multisize Windows technique," *Information Fusion*, vol. 10, no. 2, pp. 163–171, 2009.

[37] X. Zhang, X. Li, and Y. Feng, "A new image fusion performance measure using Riesz transforms," *Optik*, vol. 125, no. 3, pp. 1427–1433, 2014.

[38] C. S. Xydeas and V. Petrović, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.

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

[40] Q. Miao, C. Shi, P. Xu, M. Yang, and Y. Shi, "A novel algorithm of image fusion using shearlets," *Optics Communications*, vol. 284, no. 6, pp. 1540–1547, 2011.

[41] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861–899, 2006.

[42] M. Fu and C. Zhao, "Fusion of infrared and visible images based on the second generation curvelet transform," *Journal of Infrared and Millimeter Waves*, vol. 28, no. 4, pp. 254–258, 2009.

[43] S. Li, B. Yang, and J. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Information Fusion*, vol. 12, no. 2, pp. 74–84, 2011.

[44] Y. Phamila and R. Amutha, "Discrete Cosine Transform based fusion of multi-focus images for visual sensor networks," *Signal Processing*, vol. 95, pp. 161–170, 2014.

[45] L. Wang, B. Li, and L. Tian, "Multi-modal medical volumetric data fusion using 3D discrete shearlet transform and global-to-local rule," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 1, pp. 197–206, 2014.