

## Research Article

# Image Fusion of CT and MR with Sparse Representation in NSST Domain

**Chenhui Qiu, Yuanyuan Wang, Huan Zhang, and Shunren Xia**

*Key Laboratory of Biomedical Engineering of Ministry of Education, Zhejiang University and the Zhejiang Provincial Key Laboratory of Cardio-Cerebral Vascular Detection Technology and Medicinal Effectiveness Appraisal, Hangzhou 310027, China*

Correspondence should be addressed to Shunren Xia; [shunren\\_xia@163.com](mailto:shunren_xia@163.com)

Received 22 June 2017; Revised 20 September 2017; Accepted 25 September 2017; Published 9 November 2017

Academic Editor: Po-Hsiang Tsui

Copyright © 2017 Chenhui Qiu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multimodal image fusion techniques can integrate the information from different medical images to get an informative image that is more suitable for joint diagnosis, preoperative planning, intraoperative guidance, and interventional treatment. Fusing images of CT and different MR modalities are studied in this paper. Firstly, the CT and MR images are both transformed to nonsubsampling shearlet transform (NSST) domain. So the low-frequency components and high-frequency components are obtained. Then the high-frequency components are merged using the absolute-maximum rule, while the low-frequency components are merged by a sparse representation- (SR-) based approach. And the dynamic group sparsity recovery (DGSR) algorithm is proposed to improve the performance of the SR-based approach. Finally, the fused image is obtained by performing the inverse NSST on the merged components. The proposed fusion method is tested on a number of clinical CT and MR images and compared with several popular image fusion methods. The experimental results demonstrate that the proposed fusion method can provide better fusion results in terms of subjective quality and objective evaluation.

## 1. Introduction

Image fusion techniques combine multiple input images of the same scene or object to get a composite image that is expected to be more informative and suitable for human visual perception or further image processing compared to any of the input images [1–3]. Multimodal medical image fusion, one of the major image fusion applications, draws increasing attention from medical academia [4, 5] and has been used for joint diagnosis, preoperative planning, intraoperative guidance, interventional treatment, and so on [6, 7]. CT and MRI techniques have been widely used in clinical diagnosis and treatment; nevertheless the accurate differences between lesion, bone, and peritumoral soft tissue sometimes are hardly obtained by the single CT or MR images [7]. Although CT images provide electron density map required for accurate radiation dose estimation and superior cortical bone contrast, they are limited in soft tissue contrast. MRI provides good soft tissue contrast that permits better visualization of tumors or tissue abnormalities. However, it lacks signal from cortical bone and has image

intensity values that have no relation to electron density [8]. Hence, for precise joint diagnoses and effective treatment procedures, the superior method for fusing CT and MR images should be explored urgently.

Multiscale transform (MST) theories are the most widely used tools in various image fusion situations [2, 3]. Classical MST tools include pyramid ones such as Laplacian pyramid (LP) [9] and gradient pyramid (GP) [10] and wavelet ones such as discrete wavelet transform (DWT) [11] and dual-tree complex wavelet transform (DTCWT) [12]. Since the pyramid and wavelet tools are originally proposed to process one-dimensional signals, even their improved versions cannot capture geometric structure information of the image very well. Current MST tools are multiscale geometric analysis (MGA) ones. MGA tools are able to extract geometric structure information including edges, curves, and textures of images much more effectively, since they are the “true” 2D transform. Curvelet transform (CVT) [13], contourlet transform (COT) [14], nonsubsampling contourlet transform (NSCT) [15], shearlet transform (ST) [16], and nonsubsampling shearlet transform (NSST) [17] are the most

representative tools of MGA. CVT is proposed to represent a curve as the superposition of bases of various lengths and widths obeying a scaling law. NSCT is the shift-invariant version of COT essentially, which is built upon nonsubsampling multiscale pyramids and nonsubsampling directional filter banks. ST is much more efficient than COT in computational consumption. Besides, there are no restrictions on the number of directions and the size of support for the shearing [17]. NSST is the shift-invariant version of ST essentially.

The geometric structure information in different scales of the image can be extracted effectively by these MGA tools. The tools above are widely studied to fuse general multimodal images [8, 18–21].

Sparse representation (SR) is another popular and powerful theory used to solve the image fusion problem. SR addresses the signals' natural sparsity by simulating the sparse coding mechanism of human vision systems [22]. Yang and Li first apply SR theory to multifocus image fusion, and they use orthogonal matching pursuit (OMP) to implement sparse coding [23]. They also study the influence of dictionary and simultaneous OMP algorithm on the fusion of multimodal images [24].

Although MST-based and SR-based image fusion methods have achieved giant success, they have their special defects, respectively. We will discuss them in Section 4. In order to combine the advantages of MST and SR and avoid their defects, Wang et al. propose a fusion method that combines NSCT and SR, and OMP algorithm is used in their method to calculate the SR coefficients of the low-frequency components [25]. A general framework for image fusion combining MST (NSST is not included) and SR is proposed and discussed in detail by Liu et al. in [26]. They make comparisons and find that the method with combination of NSCT and SR which utilizes OMP algorithm for sparse coding obtains excellent fusion performance.

In this paper, we propose a novel fusion method to fuse clinical CT and MR images by combining NSST and SR. Firstly, the source images are both transformed to NSST domain. Then the high-frequency components are merged by the absolute-maximum rule, while the low-frequency components are merged with the SR-based approach. Wang et al. and Liu et al. utilize OMP algorithm to implement sparse coding, while the dynamic group sparsity recovery (DGSR) algorithm is first explored in our proposed method to improve the performance of SR-based approach. Finally, the fused image is obtained by performing the inverse NSST on the merged low-frequency and high-frequency components.

Our proposed fusion method is tested on forty pairs of clinical brain CT and MR anatomical images and compared with seven popular image fusion methods. MR imaging techniques can provide four kinds of anatomical images which are all considered in our research. The experimental results demonstrate that the proposed fusion method can provide superior fusion results in terms of subjective quality and objective evaluation.

The remainder of this paper is organized as follows. The related work including NSST and SR is presented in Section 2. The proposed fusion method is described in Section 3. Experiments and discussions are shown in Section 4. Finally,

we draw the conclusions and depict some future work in Section 5.

## 2. Related Work

*2.1. Nonsubsampling Shearlet Transform (NSST).* Shearlet has been regarded as the best sparse directional image representation frame among MST theories so far [5]. In dimension  $n = 2$ , the affine systems with composite dilations are described as follows:

$$\psi_{j,l,k}(x) = |\det A|^{j/2} \psi(S^l A^j x - k), \quad (1)$$

where  $\psi \in L^2(\mathbb{R}^2)$ ,  $A$  refers to anisotropy matrix which is associated with scale transformations, and  $S$  refers to shear matrix which is associated with area-preserving geometrical transformations, such as rotations and shear.  $j$ ,  $l$ , and  $k$  are scale, direction, and shift parameter, respectively. For each  $a > 0$  and  $s \in \mathbb{R}$ , the matrices of  $A$  and  $S$  which play vitally important roles in the process of ST are given as follows [17]:

$$A = \begin{bmatrix} a & 0 \\ 0 & \sqrt{a} \end{bmatrix}, \quad (2)$$

$$S = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}.$$

Assume that  $a = 4$  and  $s = 1$ ; then we obtain

$$A_0 = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, \quad (3)$$

$$S_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Let  $A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$  and  $S_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ , for any  $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$  and  $\xi_1 \neq 0$ ; also let  $\psi^{(0)}$  and  $\psi^{(1)}$  be given by

$$\widehat{\psi}^{(0)}(\xi) = \widehat{\psi}^{(0)}(\xi_1, \xi_2) = \widehat{\psi}_1(\xi_1) \widehat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right) \quad (4)$$

$$\widehat{\psi}^{(1)}(\xi) = \widehat{\psi}^{(1)}(\xi_1, \xi_2) = \widehat{\psi}_1(\xi_2) \widehat{\psi}_2\left(\frac{\xi_1}{\xi_2}\right),$$

where  $\widehat{\psi}_1, \widehat{\psi}_2 \in C^\infty(\widehat{\mathbb{R}})$ ,  $\text{supp } \widehat{\psi}_1 \subset [-1/2, -1/16] \cup [1/16, 1/2]$ , and  $\text{supp } \widehat{\psi}_2 \subset [-1, 1]$ .

Each element of  $\widehat{\psi}_{j,l,k}$  is supported on a pair of trapezoids, of approximate size  $2^j \times 2^{2j}$ , oriented along lines of slope  $l2^{-j}$ .

Then the ST function is obtained:

$$\psi_{j,l,k}^{(0)}(x) = 2^{3j/2} \psi^{(0)}(S_0^l A_0^j x - k) \quad (5)$$

$$\psi_{j,l,k}^{(1)}(x) = 2^{3j/2} \psi^{(1)}(S_1^l A_1^j x - k),$$

where  $j \geq 0$ ,  $-2^j \leq l \leq 2^j - 1$ , and  $k \in \mathbb{Z}^2$ .

ST has the following properties: well spatial and frequent localization, strong selectivity of anisotropic directionality,

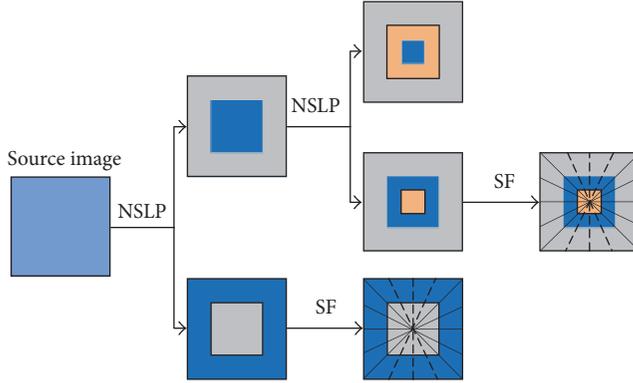


FIGURE 1: Two-level multiscale and multidirectional decompositions of the NSST.

well parabolic scaling, and approximately sparse representation. However, ST engenders the Gibbs phenomena because of the lack of shift invariance. NSST is the shift-invariant version of ST. It utilizes nonsubsampled Laplacian pyramid (NSLP) filters as the substitutes for the LP filters which are used in the ST mechanism. Multiscale decomposition is implemented by NSLP. One low-frequency component and one high-frequency component can be produced at each NSLP decomposition level. And the next NSLP decomposition is performed on the last (previous) low-frequency component to capture the singularities of the source image iteratively. When the decomposition level is set to  $J$ , the source image is decomposed into  $J + 1$  components with the same size of the source image, in which one is the low-frequency component.

Shearing filter (SF) is performed on high-frequency components of each NSLP decomposition level without subsampling to achieve the multidirectional factorization, which ensures the shift invariance property of NSST. Assume that SF performs the directional decomposition with  $l$  stages on a high-frequency component decomposed by NSLP; then  $2^l$  directional subbands with the same size as the source image are produced [17]. Two-level decomposition of the NSST is shown in Figure 1. This schematic diagram illustrates the NSLP decomposition and the corresponding directional decomposition by SF.

**2.2. Sparse Representation (SR).** SR theory is based on the assumption that the natural signals can be represented or approximately represented by a linear combination of a “few” atoms from dictionary [22]. For signals  $\mathbf{\Gamma} \in \mathbb{R}^n$ , SR theory indicates the existence of a dictionary  $\mathbf{D} \in \mathbb{R}^{n \times m}$ , which contains  $m$  prototype signals that are referred to atoms. Then, for any signal  $\mathbf{x} \in \mathbf{\Gamma}$ , there is a linear combination of atoms from  $\mathbf{D}$  which can approximate to  $\mathbf{x}$  very well. The signal  $\mathbf{x}$  can be expressed as  $\mathbf{x} \approx \mathbf{D}\boldsymbol{\alpha}$ , where  $\boldsymbol{\alpha} \in \mathbb{R}^m$  is the unknown sparse coefficient vector. It usually assumes that  $n < m$ , which implies that  $\mathbf{D}$  is the overcompleted and redundant dictionary. So there are numerous feasible solutions of the

underdetermined system  $\mathbf{x} \approx \mathbf{D}\boldsymbol{\alpha}$ . Finding the sparsest  $\boldsymbol{\alpha}$  that contains the fewest nonzero entries involves solving the following optimization problem:

$$\begin{aligned} \arg \min_{\boldsymbol{\alpha}} \quad & \|\boldsymbol{\alpha}\|_0 \\ \text{s.t.} \quad & \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{x}\|_2 < \varepsilon, \end{aligned} \quad (6)$$

where  $\|\boldsymbol{\alpha}\|_0$  denotes the number of nonzero entries in  $\boldsymbol{\alpha}$ ;  $\varepsilon > 0$  is an error tolerance.

The above optimization is an NP-hard problem. The main idea to solve this problem is obtaining an approximate solution instead of the sparsest one. OMP is the representative algorithm to get the approximate solution.

### 3. Proposed Fusion Method

The proposed fusion method of CT and MR images mainly contains four phases: NSST decomposition, high-frequency fusion, low-frequency fusion, and NSST reconstruction. The schematic diagram of the proposed method is shown in Figure 2.

**3.1. NSST Decomposition.** NSST decomposition is performed on CT and MR images  $\{I_C, I_M\}$  to obtain high-frequency components that are denoted as  $\{H_C, H_M\}$  and low-frequency components that are denoted as  $\{L_C, L_M\}$ .

**3.2. High-Frequency Fusion.** The high-frequency components at different levels contain much salient information of regional boundary, edge, and line features. The larger absolute values of the coefficients in each of the high-frequency components indicate more important salient information. Therefore, the popular absolute-maximum rule is appropriate to merge the high-frequency components  $\{H_C, H_M\}$ . The coefficients of high-frequency components  $H_C$  and  $H_M$  and the fused image  $H_F$  are denoted as  $h_C$  and  $h_M$  and  $h_F$ , respectively. The rule is expressed as follows:

$$h_F = \begin{cases} h_C & \text{if } |h_C| \geq |h_M| \\ h_M & \text{otherwise.} \end{cases} \quad (7)$$

**3.3. Low-Frequency Fusion.**  $L_C$  and  $L_M$  are merged with SR-based fusion approach in our research. Image fusion methods always depend on the local information of source images. Besides, the stability and efficiency of sparse coding should be taken into account. Hence, the source images need to be divided into image patches.

The sliding window technique is adopted to divide  $L_C$  and  $L_M$  into image patches of size  $d \times d$  from upper left to lower with a suitable step length, respectively [23]. Then all patches of  $L_C$  are transformed into vectors via lexicographic ordering and normalization operation. And all these vectors constitute the matrix  $\mathbf{V}_C$ , in which each column corresponds to one patch of  $L_C$ . The size of  $\mathbf{V}_C$  is  $(d \cdot d) \times ((M - d + 1) \cdot (N - d + 1))$ .  $\mathbf{V}_M$  can be obtained in the same way.

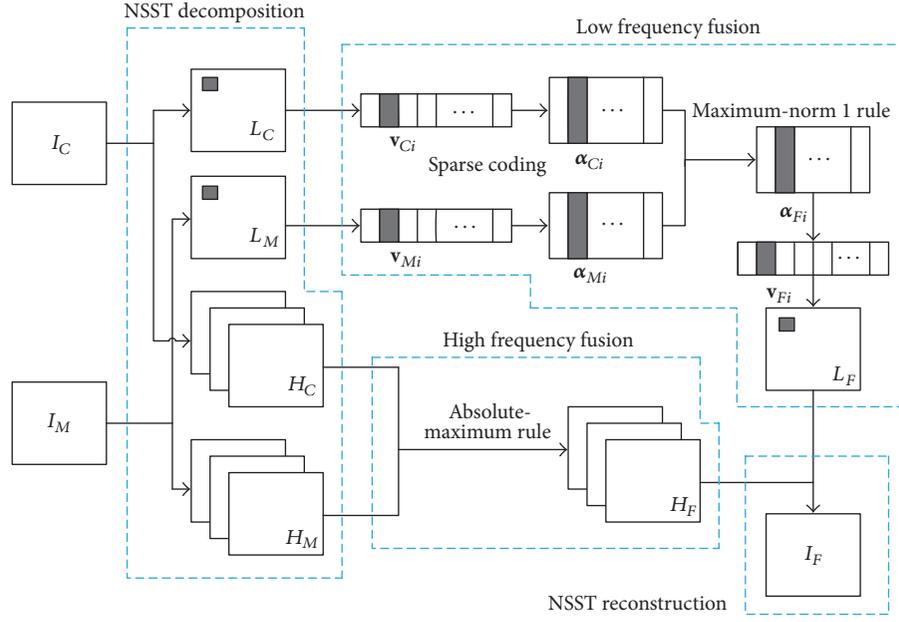


FIGURE 2: The schematic diagram of the proposed method.

For the  $i$ th column vector  $\mathbf{v}_{Ci}$  of  $\mathbf{V}_C$ , its sparse representation is calculated by the DGSR algorithm. According to (6), the sparse coding vector  $\alpha_{Ci}$  for  $\mathbf{v}_{Ci}$  is obtained as follows:

$$\begin{aligned} \alpha_{Ci} &= \arg \min_{\alpha} \|\alpha\|_0 \\ \text{s.t.} \quad & \|\mathbf{D}\alpha - \mathbf{v}_{Ci}\|_2 < \varepsilon. \end{aligned} \quad (8)$$

Similarly, we can get the sparse coding vector  $\alpha_{Mi}$  for  $\mathbf{v}_{Mi}$  which is the  $i$ th column vector of  $\mathbf{V}_M$ .

$$\begin{aligned} \alpha_{Mi} &= \arg \min_{\alpha} \|\alpha\|_0 \\ \text{s.t.} \quad & \|\mathbf{D}\alpha - \mathbf{v}_{Mi}\|_2 < \varepsilon, \end{aligned} \quad (9)$$

where  $\mathbf{D}$  is the offline learned dictionary that is trained by  $K$ -means singular value decomposition ( $K$ -SVD) algorithm [27].

Then  $\alpha_{Ci}$  and  $\alpha_{Mi}$  are merged by maximum-norm 1 rule to produce the fused sparse vectors  $\alpha_{Fi}$ . The fusion rule is expressed as follows:

$$\alpha_{Fi} = \begin{cases} \alpha_{Ci} & \text{if } \|\alpha_{Ci}\|_1 \geq \|\alpha_{Mi}\|_1 \\ \alpha_{Mi} & \text{otherwise.} \end{cases} \quad (10)$$

The  $i$ th column vector  $\mathbf{v}_{Fi}$  of the matrix  $\mathbf{V}_F$  is calculated by

$$\mathbf{v}_{Fi} = \mathbf{D}\alpha_{Fi}. \quad (11)$$

Then  $\mathbf{V}_F$  can be constituted by all these column vectors.

Finally, the low-frequency component  $L_F$  of the fused image  $I_F$  is reconstructed by  $\mathbf{V}_F$ . This reconstruction is the inverse process of normalization and sliding window technique.

In the fusion application of CT and MR images, the nonzero sparse coefficients of source images are not randomly distributed but are group-clustered. They tend to be clustered into groups, although the clustering group structures are dynamic and unpredictable. Therefore, DGSR algorithm that prunes data residues in the iterative process according to both sparsity and the group clustering trend is appropriate for our research. The experimental results in [28] demonstrate that the DGSR algorithm achieves better performance than OMP algorithm. In this paper, the main work is to explore the image fusion method for CT and MR; the details of DGSR algorithm can be found in [28]. The MATLAB codes of DGSR algorithm and relevant experiments can be downloaded at <http://ranger.uta.edu/~huang/Downloads.htm>.

**3.4. NSST Reconstruction.** The inverse NSST is performed on both low frequency  $L_F$  and high frequency  $H_F$  to reconstruct the final fused image  $I_F$ .

## 4. Experiments and Discussions

**4.1. CT and MR Images.** CT imaging techniques provide anatomical images of tissues or lesions. MR-T1W, MR-T2W, and MR-PDW are derived from different settings of longitudinal relaxation time T1, transversal relaxation time T2, echo pulse sequence, and proton density of tissues. T1W means that the contrasts of tissues in the MR image mainly depend on T1 weight. T2W and PDW are of the similar meaning. MR-CE image is acquired by adding gadolinium-containing contrast medium, which is mainly used for the definite diagnosis of cerebral tumor. They are all MR anatomical images.

Because of the limited space of paper, only one representative group of clinical brain CT and MR images are presented

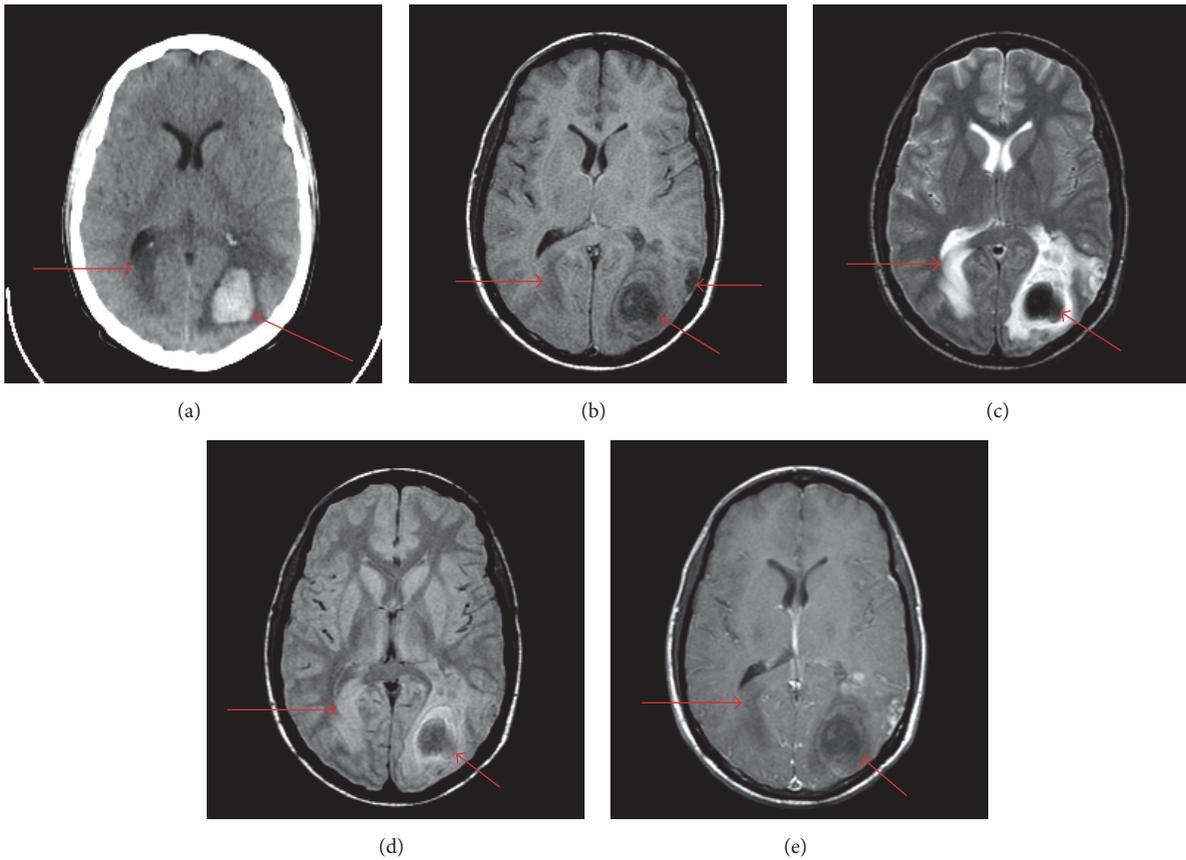


FIGURE 3: The clinical brain CT and MR images. (a) CT image, (b) MR-T1W image, (c) MR-T2W image, (d) MR-PDW image, and (e) MR-CE image.

as the source images which are shown in Figure 3. The images of Figure 3 are the 12th slices of a Ewing's sarcoma patient at the age of 22. He had a right homonymous hemianopia, a left inferior quadrantanopia, right lower extremity hyperreflexia, and right extensor plantar response. The images are of the same slice of the same patient so that we can find the difference of different imaging modalities obviously. They can be downloaded from <http://www.med.harvard.edu/aanlib/>.

**4.2. Experimental Setting.** The clinical brain CT and MR images above have the same size,  $256 \times 256$ , and they are preregistered in pairs. All the experiments are implemented in MATLAB 7.14 and on a PC with 2.66-GB CPU, 4 GB RAM, and 32-bit OS.

To assess the performance of the proposed method, the GP [10], DTCWT [12], CVT [18], NSCT [19], NSST [20], SR-OMP [23], and NSCT-SR-OMP [25, 26] based fusion methods are implemented for comparison.

The fusion rules for low-frequency components of these MST-based methods are all chosen as "averaging" rule. The fusion rules for high-frequency components of these MST-based methods are all chosen as "max-absolute" rule [2]. For SR-OMP based method, the fusion rule is chosen as "max-norm 1" rule [23]. For NSCT-SR-OMP based method, the fusion rules for low-frequency and high-frequency components are the same as the proposed method's fusion rules [25,

26]. For these MST-based methods, NSCT-SR-OMP based method, and the proposed method, the decomposition levels are all set to 4, which are most appropriate according to our experiments and the analysis in [2, 26]. For DTCWT-based method, the filters of first level and other levels are selected as "5-3" and "q-a," respectively. For NSCT-based and NSCT-SR-OMP based methods, the pyramidal filters and directional filters are selected as "pyrexc" and "7-9," respectively. And the directional numbers from coarser to finer scales are chosen as  $2^3$ ,  $2^3$ ,  $2^4$ , and  $2^4$ , respectively. For NSST-based method and the proposed method, NSLP filters are selected as "pyrexc." The directional numbers are the same as aforementioned. And the Meyer wavelet window is used for SF. The sizes of SFs which are ordered from coarser to finer scales are chosen as 32, 32, 16, and 16, respectively [17]. For SR-OMP and NSCT-SR-OMP based methods and the proposed method, the size of image patches is set to  $8 \times 8$ , the step length is set to 2, the error tolerance is set to 0.1, and the dictionary size is set to  $64 \times 256$  [2, 23].

**4.3. Objective Evaluation Metrics.** Generally, image fusion results can be evaluated in subjective and objective ways. Subjective ways are difficult to execute because they are based on psychovisual testing and cannot be applied in an automatic system. Besides, there is little visual difference among the fusion results in most cases, so subjective ways

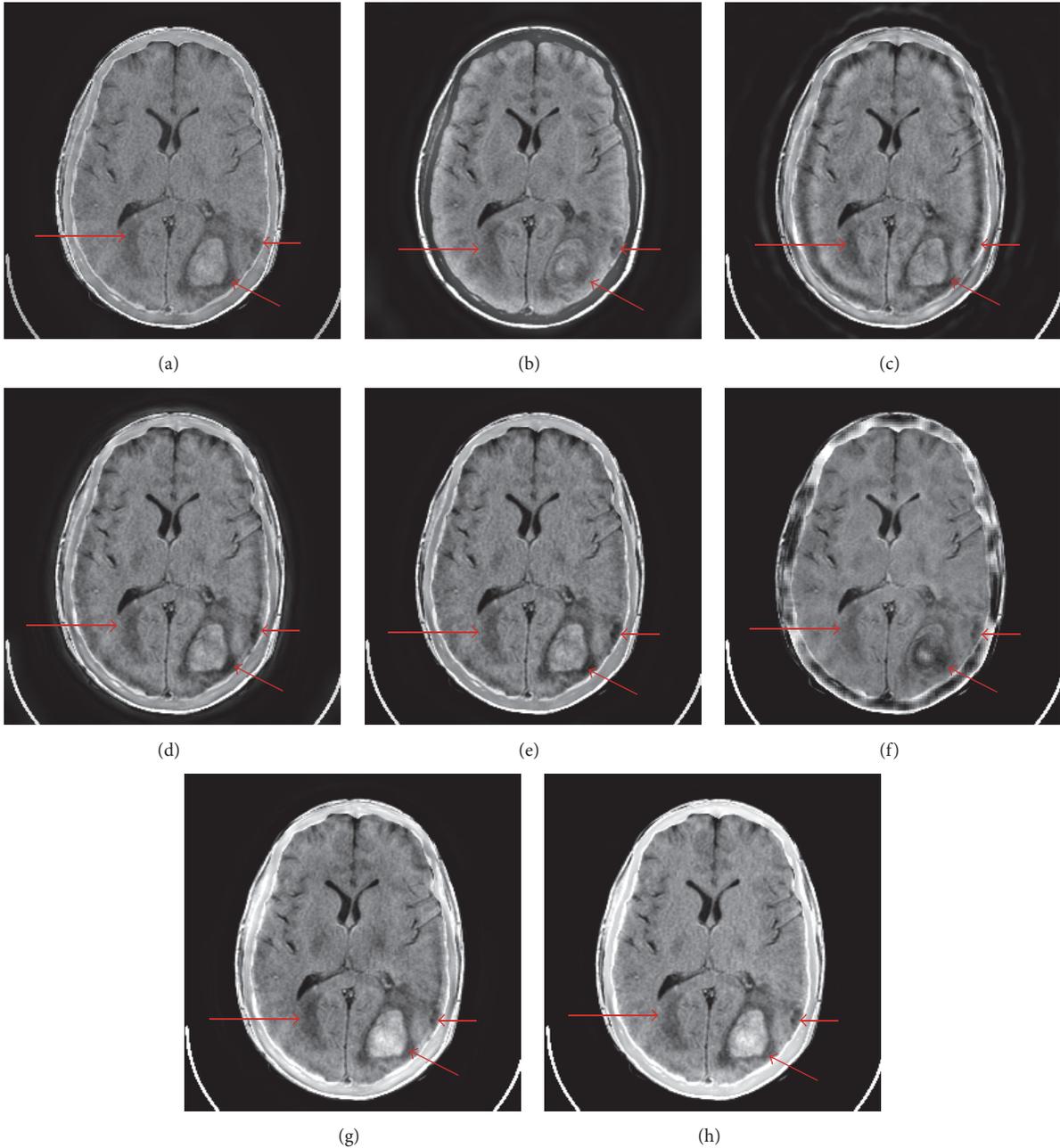


FIGURE 4: The fusion results of Figures 3(a) and 3(b). ((a)–(h)) GP, DTCWT, CVT, NSCT, NSST, SR-OMP, NSCT-SR-OMP, and proposed method.

are difficult to correctly evaluate the fusion results. Therefore, many objective evaluation metrics have been developed [2, 3]. In this paper, standard deviation (STD) [29], mean structural similarity (M-SSIM) [30], mutual information (MI) [31], sum of the correlations of differences (SCD) [32],  $Q_0$  [33],  $Q_w$  [34],  $Q_E$  [34],  $Q^{AB/F}$  [35], and visual information fidelity for fusion (VIFF) [36] are used to evaluate the fusion results objectively. The larger values of these metrics imply the better fusion results. It is worthwhile to note that if the fused images are used for preoperative planning, intraoperative guidance, and interventional treatment, which depend on further image

processing, the superiority of objective evaluation is important and valuable. We also give simply subjective assessments on the fusion results.

**4.4. Results and Discussions.** Figure 4 shows the fusion results of Figures 3(a) and 3(b) with different fusion methods. In the right halves of Figure 3(b) and Figures 4(a)–4(h), there are two lesions labeled by the arrows. Actually they are the one lesion. And the small one is the diffusion of the large one. The small lesion in right half and the lesion in left half of Figures 4(a), 4(b), 4(c), and 4(f) is blurred to some extent. Although

TABLE 1: Objective evaluation metrics of the fusion results of Figures 3(a) and 3(b).

Method	STD	M-SSIM	MI	SCD	$Q_0$	$Q_W$	$Q_E$	$Q^{AB/F}$	VIFF
Proposed	<b>83.1426</b>	<b>0.7511</b>	<i>2.7804</i>	<b>1.4081</b>	<b>0.3614</b>	<b>0.6502</b>	0.4119	0.4911	<b>0.4494</b>
NSCT-SR-OMP	78.2445	0.6689	2.6225	1.1685	0.3478	0.6320	0.4133	0.4765	0.4124
SR-OMP	67.3711	0.7392	<b>3.1058</b>	0.8483	0.3164	0.5886	0.4131	0.5091	0.2517
NSST	69.2338	0.6853	2.5469	1.0222	0.3535	0.6504	<b>0.4364</b>	0.5205	0.3862
NSCT	66.1168	0.5354	2.3739	0.8571	0.3266	0.5942	0.3935	0.4852	0.3302
CVT	67.1317	0.5582	2.2612	0.8859	0.3010	0.6186	0.4017	0.4629	0.3319
DTCWT	64.1289	0.5115	2.6224	0.8886	0.2430	0.3381	0.2401	<b>0.5746</b>	0.1849
GP	66.3231	0.6104	2.5530	0.9997	0.3505	0.6201	0.4042	0.4868	0.3440

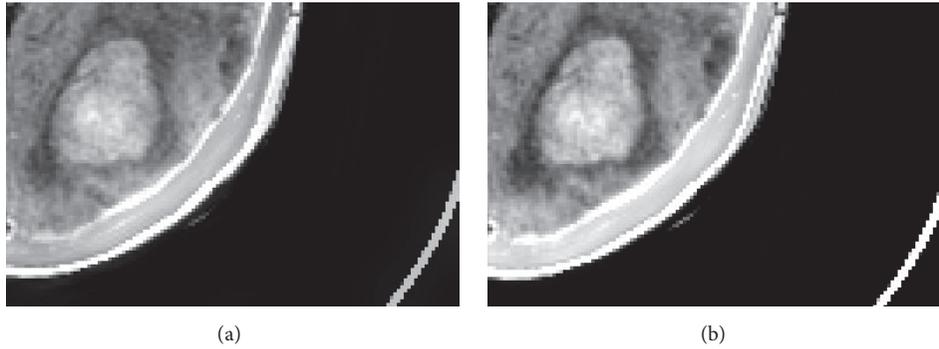


FIGURE 5: The zoomed-in version of the bottom-right part of Figures 4(g) and 4(h): (a) the zoomed-in version of Figure 4(g); (b) the zoomed-in version of Figure 4(h).

Figures 4(d) and 4(e) exhibit the small lesion in right half well, they cannot exhibit the lesion in left half with good visual quality. The cerebral bony structure is the basic sclerous tissue in brain anatomical image. Only Figures 4(g) and 4(h) show the cerebral bony structure as well as the expected effect. On the whole, the visual information of tissues and lesions of CT and MR-T1W images is most perfectly preserved in Figures 4(g) and 4(h). Besides, the CVT- and NSCT-based methods produce artifacts along the cerebral bony structure. Although SR-OMP based method can capture the salient information, it cannot extract the blurry features (such as the lesion in left half) effectively.

Table 1 gives the objective evaluation metrics of the fusion results of Figures 3(a) and 3(b) with different fusion methods. The best result for each metric is labeled in bold. The second best result for each metric is labeled with italic font. The proposed method provides the best results for six metrics and the second best result for one metric. Although the fusion results of NSCT-SR-OMP based method and the proposed method are very similar in subjective quality, the proposed method is superior to NSCT-SR-OMP based method in all of the objective evaluation metrics. In addition, the computational efficiency of the proposed method is much higher than that of NSCT-SR-OMP based method because the decomposition and reconstruction of NSST are much faster than those of NSCT [17].

The zoomed-in version of the bottom-right part of Figures 4(g) and 4(h) is shown in Figure 5. We can find that the contrast of the larger lesion and its limbic region in Figure 5(b) is higher than the corresponding part in

Figure 5(a). This verdict coincides with the STD metric evaluation. Besides, the visual quality of Figure 5(b) is higher than that of Figure 5(a), which is consistent with the VIFF metric. (Note that medical display system provides better visual effect than common display system; it distinguishes the difference between Figures 5(a) and 5(b) more clearly.)

Figure 6 shows the fusion results of Figures 3(a) and 3(c) with different fusion methods. The subjective analysis of Figure 6 is similar to the subjective assessment of Figure 4. The proposed method provides the best visual quality of the lesion labeled by the arrow, sulcus, and bony structure. NSCT-SR-OMP based method provides similar but little worse subjective visual quality with careful inspection.

Table 2 gives the objective evaluation metrics of the fusion results of Figures 3(a) and 3(c) with different fusion methods. The proposed method provides the best results for six metrics and the second best result for one metric.

Figures 7 and 8 show the fusion results of Figures 3(a) and 3(d) and Figures 3(a) and 3(e) with different fusion methods, respectively. The proposed method provides the best visual fusion performances in the two cases.

Tables 3 and 4 give the objective evaluation metrics of the fusion results of Figures 3(a) and 3(d) and Figures 3(a) and 3(e) with different fusion methods, respectively. The proposed method provides the best results for six metrics and the second best result for one metric in Table 3. It also provides the best results for five metrics and the second best result for one metric in Table 4.

Table 5 gives the objective evaluation metrics of the fusion results of forty pairs of clinical brain CT and different MR

TABLE 2: Objective evaluation metrics of the fusion results of Figures 3(a) and 3(c).

Method	STD	M-SSIM	MI	SCD	$Q_0$	$Q_W$	$Q_E$	$Q^{AB/F}$	VIFF
Proposed	<b>85.4593</b>	<b>0.7386</b>	2.9224	<b>1.4621</b>	<b>0.3824</b>	<b>0.7690</b>	0.4827	0.5113	<b>0.4131</b>
NSCT-SR-OMP	76.8654	0.6591	2.8395	1.1346	0.3718	0.7587	<b>0.5001</b>	0.5217	0.3669
SR-OMP	73.0928	0.7235	<b>3.1352</b>	1.0630	0.3353	0.7170	0.4729	0.5192	0.2911
NSST	69.3967	0.6643	2.6868	1.0580	0.3670	0.7372	0.4851	<b>0.5332</b>	0.3338
NSCT	66.7627	0.5297	2.6151	0.9171	0.3433	0.6916	0.4680	0.4921	0.2897
CVT	67.2029	0.5603	2.4766	0.9319	0.3152	0.7006	0.4542	0.4715	0.2837
DTCWT	61.8421	0.4939	2.5247	0.8863	0.2558	0.2433	0.1876	0.4946	0.1420
GP	66.1162	0.6455	2.7886	1.0829	0.3662	0.6868	0.4535	0.4983	0.3151

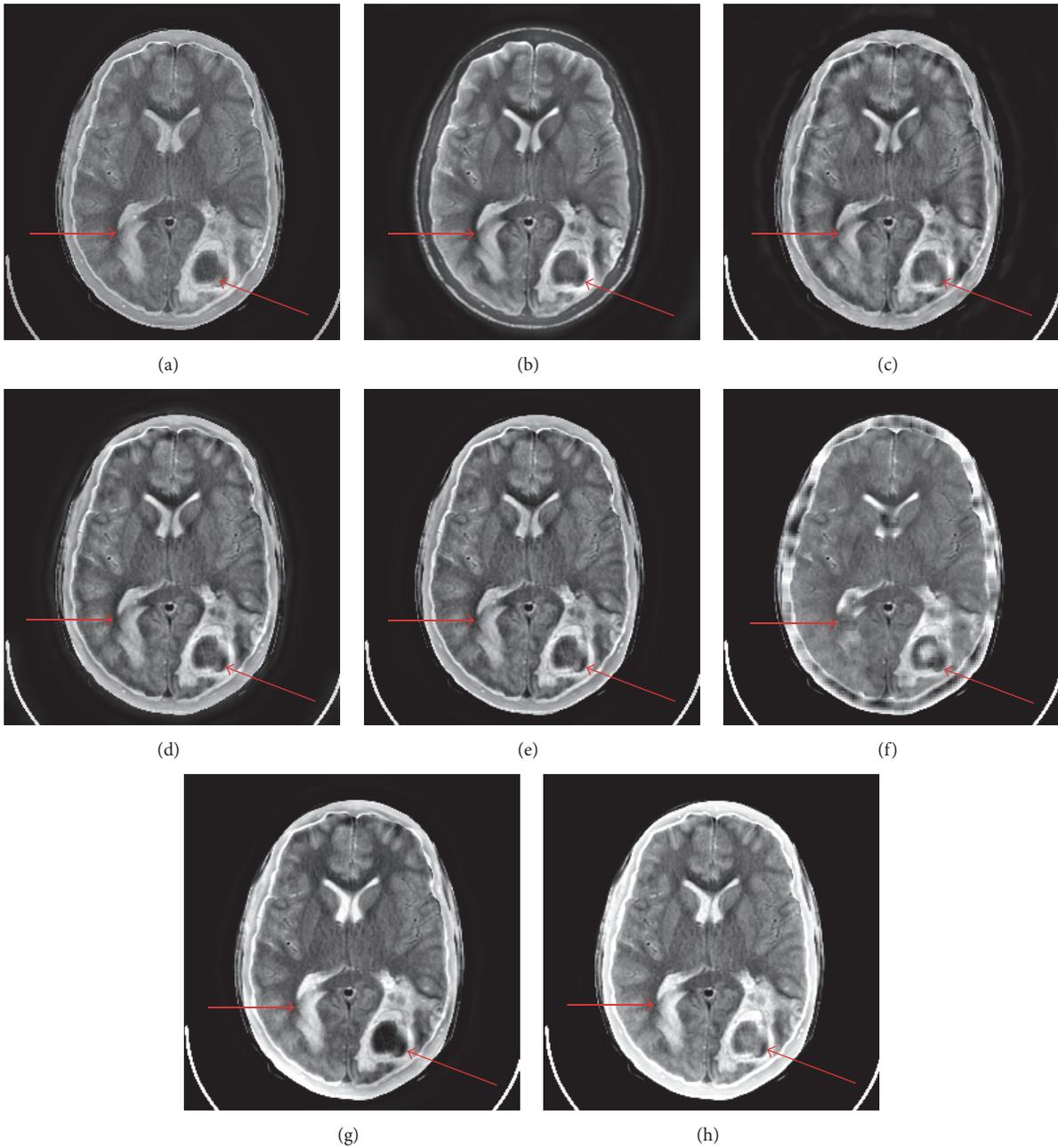


FIGURE 6: The fusion results of Figures 3(a) and 3(c). ((a)–(h)) GP, DTCWT, CVT, NSCT, NSST, SR-OMP, NSCT-SR-OMP, and proposed method.

TABLE 3: Objective evaluation metrics of the fusion results of Figures 3(a) and 3(d).

Method	STD	M-SSIM	MI	SCD	$Q_0$	$Q_W$	$Q_E$	$Q^{AB/F}$	VIFF
Proposed	<b>83.5111</b>	<b>0.7583</b>	2.7913	<b>1.3901</b>	<b>0.3749</b>	<b>0.6762</b>	0.4247	0.4896	<b>0.4487</b>
NSCT-SR-OMP	74.3309	0.6554	2.5727	1.0432	0.3611	0.6548	0.4312	0.4820	0.3996
SR-OMP	68.0421	0.7522	<b>2.8751</b>	0.8643	0.3365	0.6083	0.4174	0.5042	0.2662
NSST	69.1365	0.6885	2.5367	1.0001	0.3679	0.6619	<b>0.4470</b>	0.5147	0.3804
NSCT	66.1375	0.5378	2.3410	0.8138	0.3371	0.5891	0.3997	0.4491	0.3288
CVT	67.2095	0.5739	2.2814	0.8713	0.3167	0.6222	0.4038	0.4503	0.3214
DTCWT	63.6317	0.5042	2.6157	0.8570	0.2564	0.2516	0.1870	<b>0.5476</b>	0.1790
GP	66.7812	0.6558	2.5425	0.9912	0.3638	0.6216	0.4109	0.4737	0.3361

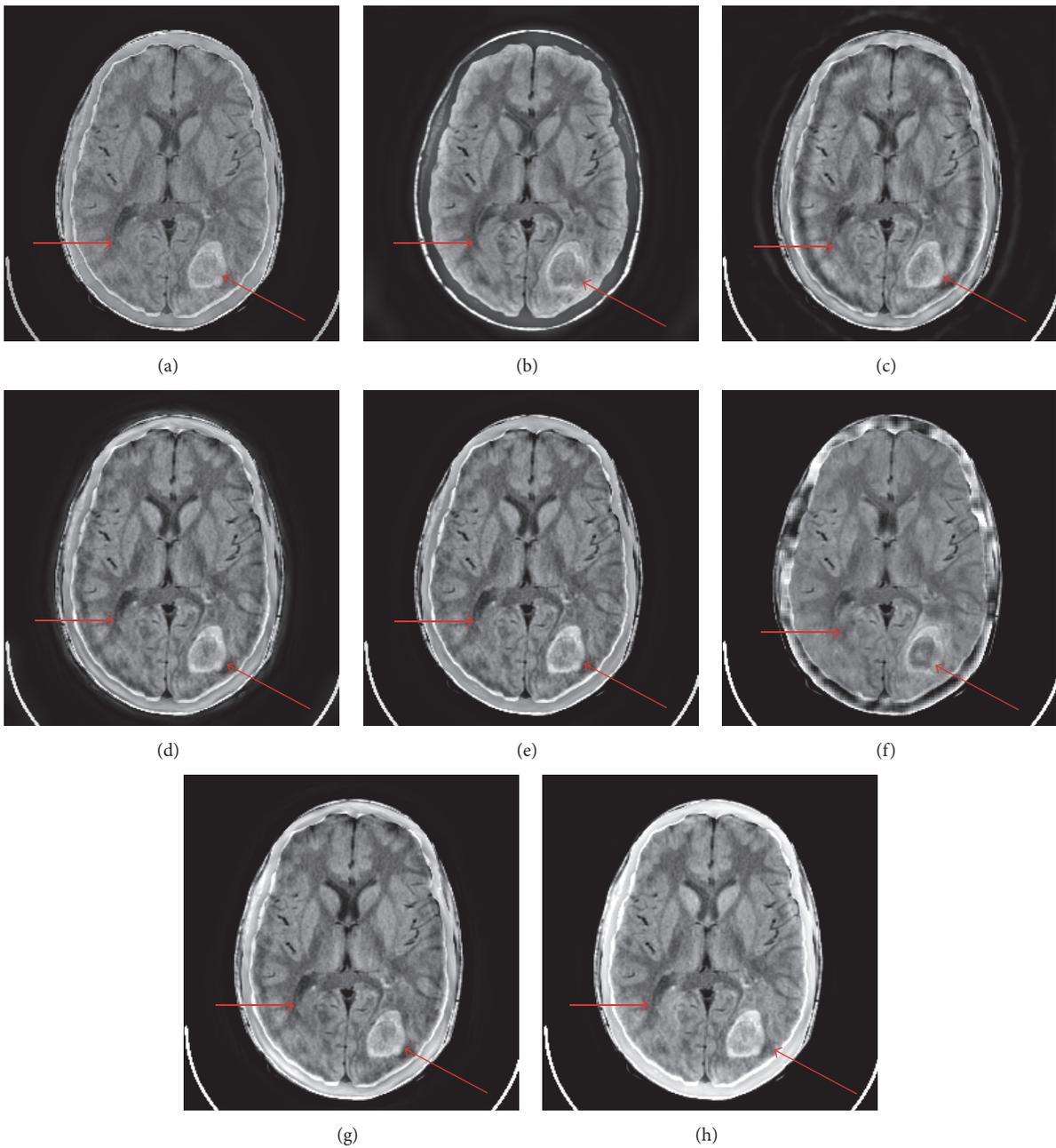


FIGURE 7: The fusion results of Figures 3(a) and 3(d). ((a)–(h)) GP, DTCWT, CVT, NSCT, NSST, SR-OMP, NSCT-SR-OMP, and proposed method.

TABLE 4: Objective evaluation metrics of the fusion results of Figures 3(a) and 3(e).

Method	STD	M-SSIM	MI	SCD	$Q_0$	$Q_W$	$Q_E$	$Q^{AB/F}$	VIFF
Proposed	<b>82.3036</b>	0.7453	2.9034	<b>1.3673</b>	<b>0.3829</b>	<b>0.6598</b>	0.3877	0.5277	<b>0.4598</b>
NSCT-SR-OMP	78.7649	0.6805	2.9842	1.1264	0.3798	0.6478	0.3888	0.5109	0.4472
SR-OMP	72.5496	<b>0.7533</b>	<b>3.4771</b>	0.7848	0.3568	0.6178	0.4028	0.5348	0.3257
NSST	70.0334	0.6944	2.6821	0.9589	0.3725	0.6573	<b>0.4181</b>	<b>0.5514</b>	0.4028
NSCT	66.9852	0.5503	2.5579	0.7918	0.3492	0.6115	0.3820	0.5033	0.3480
CVT	68.0393	0.5602	2.3974	0.8071	0.3157	0.6202	0.3600	0.4694	0.3529
DTCWT	64.8234	0.5444	2.7009	0.8209	0.2000	0.3193	0.2279	0.5390	0.1816
GP	66.4735	0.6117	2.6955	0.9055	0.3758	0.6300	0.4058	0.5206	0.3567

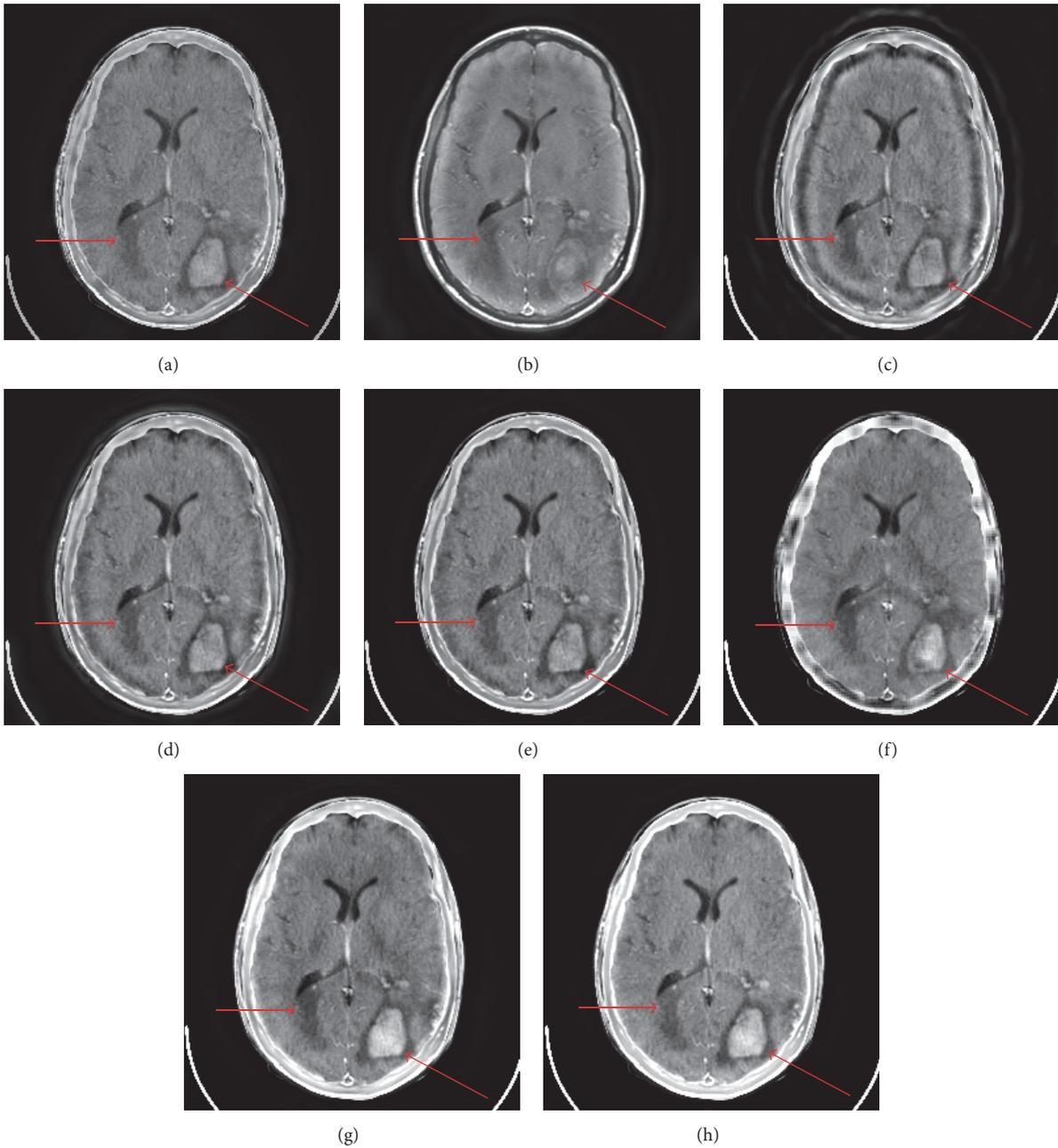


FIGURE 8: The fusion results of Figures 3(a) and 3(e). ((a)–(h)) GP, DTCWT, CVT, NSCT, NSST, SR-OMP, NSCT-SR-OMP, and proposed method.

TABLE 5: Objective evaluation metrics of the fusion results of the forty pairs of clinical brain CT and MR images (mean/STD).

Method	STD	M-SSIM	MI	SCD	$Q_0$	$Q_W$	$Q_E$	$Q^{AB/F}$	VIFF
Proposed	<b>83.5748/1.5512</b>	<b>0.7523/0.0245</b>	2.9104/0.1727	<b>1.4430/0.1321</b>	<b>0.3798/0.0284</b>	<b>0.7166/0.0403</b>	0.4495/0.0464	0.5037/0.0384	<b>0.4874/0.0778</b>
NSCT-SR-OMP	77.1774/2.3899	0.6775/0.0275	2.7336/0.1759	1.1635/0.1341	0.3605/0.0195	0.7117/0.0419	0.4547/0.0385	0.5074/0.0298	0.4432/0.0802
SR-OMP	71.4563/3.1108	0.7381/0.0228	<b>3.1898/0.2426</b>	0.9630/0.1418	0.3414/0.0248	0.6827/0.0446	0.4481/0.0372	0.5284/0.0302	0.3224/0.0584
NSST	69.3828/1.9412	0.6874/0.0245	2.6168/0.1203	1.0590/0.1133	0.3601/0.0179	0.7043/0.0304	<b>0.4612/0.0305</b>	<b>0.5378/0.0293</b>	0.4068/0.0704
NSCT	66.7117/2.0838	0.5470/0.0273	2.4958/0.1426	0.9087/0.1121	0.3363/0.0166	0.6593/0.0362	0.4283/0.0276	0.4850/0.0260	0.3555/0.0681
CVT	67.2008/1.9224	0.5680/0.0230	2.3954/0.1572	0.9241/0.1112	0.3119/0.0162	0.6668/0.0316	0.4176/0.0306	0.4649/0.0239	0.3494/0.0687
DTCWT	62.7511/2.5960	0.5185/0.0331	2.5262/0.1377	0.8533/0.1033	0.2360/0.0280	0.3276/0.0654	0.2409/0.0489	0.5097/0.0383	0.1888/0.0441
GP	65.6848/1.7288	0.6384/0.0271	2.6640/0.1081	1.0159/0.1336	0.3576/0.0174	0.6686/0.0261	0.4413/0.0302	0.5050/0.0325	0.3653/0.0497

modalities with different fusion methods. The means and the standard deviations of different metrics of different methods are shown in Table 5. We can find that the proposed method provides the best results for six metrics.

NSST is the most representative MST tool, which can extract the geometric structure information of source images more effectively than NSCT, CVT, and classical MST tools. The salient information such as edges, curves, and textures can be captured by NSST very well and can be preserved in its high-frequency components. Therefore, NSST-based method obtains the best results of  $Q_E$  or  $Q^{AB/F}$  metrics in some cases. However, NSST cannot express the low-frequency components sparsely. If ordinary fusion rule is utilized to merge the low-frequency components, the fusion performance may be degraded as the low-frequency components contain a lot of energy of source images.

SR theory can represent an image in the approximately sparsest sense. It is the best approach of two-dimensional signal in terms of information theory, so SR-based method obtains the best results of MI metric from Tables 1–5. While SR-based fusion approach can represent the underlying salient information of source images very efficiently, it is difficult to reconstruct the small-scale details of source images and the blurred features.

The proposed method combines NSST and SR in order to preserve their advantages and avoid their defects. DGSR algorithm is first explored in the proposed method to ensure the satisfactory performance when SR-based fusion approach is implemented on low-frequency components. It is worthwhile to note that NSST plays a more important role than SR in the proposed fusion method, since the fusion framework is mainly based on NSST. The geometric structure information including edges, curves, and textures of images is preserved in high-frequency components of NSST, while a minority of salient features and a majority of energy are preserved in the low-frequency components of NSST. The SR-based approach is used to fuse the low-frequency components and DGSR algorithm is explored to calculate sparse representation coefficients of the low-frequency components. We can find that the proposed method has the superiority in STD, M-SSIM, SCD,  $Q_0$ ,  $Q_W$ , and VIFF metrics from the tables.

## 5. Conclusions and Future Work

In this paper, we have proposed a novel image fusion method for fusing CT and MR images, which combines the advantages of NSST and SR. Moreover, DGSR algorithm is utilized when the SR-based fusion approach is performed on the low-frequency components of NSST decompositions. To assess the performance of the proposed method, the GP, DTCWT, CVT, NSCT, NSST, SR-OMP, and NSCT-SR-OMP based fusion methods are implemented for comparison. The experimental results show that the proposed method provides superior fusion results in terms of subjective quality and objective evaluation metrics.

However, fusing MR and low-dose CT images is not considered in our paper because the low-dose CT image contains a lot of noise, which is complicated and cannot be

modeled as Gaussian distribution or Rician distribution. We may concentrate on this fusion topic in the future.

Besides, the automatic or guided target delineation in radiotherapy planning which is based on the fused CT and MR images will be studied in the future.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors would like to sincerely thank Dr. Longyan Ni and Dr. Hongxing Liu from Nanjing Brain Hospital affiliated to Nanjing Medical University (China), Professor Shutao Li from Hunan University (China), and Dr. Haitao Yin from Nanjing University of Posts and Telecommunications (China). This work is supported by National Key Research and Development Program of China (no. 2016YFC1306600).

## References

- [1] A. A. Goshtasby and S. Nikolov, "Image fusion: advances in the state of the art," *Information Fusion*, vol. 8, no. 2, pp. 114–118, 2007.
- [2] 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.
- [3] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: a survey of the state of the art," *Information Fusion*, vol. 33, pp. 100–112, 2017.
- [4] A. P. James and B. V. Dasarathy, "Medical image fusion: a survey of the state of the art," *Information Fusion*, vol. 19, no. 1, pp. 4–19, 2014.
- [5] J. Du, W. Li, K. Lu, and B. Xiao, "An overview of multi-modal medical image fusion," *Neurocomputing*, vol. 215, pp. 3–20, 2016.
- [6] G. J. Webster, J. E. Kilgallon, K. F. Ho, C. G. Rowbottom, N. J. Slevin, and R. I. Mackay, "A novel imaging technique for fusion of high-quality immobilised MR images of the head and neck with CT scans for radiotherapy target delineation," *British Journal of Radiology*, vol. 82, no. 978, pp. 497–503, 2009.
- [7] J. Dai, X. Wang, Y. Dong, H. Yu, D. Yang, and G. Shen, "Two- and three-dimensional models for the visualization of jaw tumors based on CT-MRI image fusion," *The Journal of Craniofacial Surgery*, vol. 23, no. 2, pp. 502–508, 2012.
- [8] P. Ganasala and V. Kumar, "CT and MR image fusion scheme in nonsubsampling contourlet transform domain," *Journal of Digital Imaging*, vol. 27, no. 3, pp. 407–418, 2014.
- [9] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, 1983.
- [10] 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.
- [11] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *CVGIP: Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235–245, 1995.
- [12] 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," *Information Fusion*, vol. 8, no. 2, pp. 119–130, 2007.

- [13] E. J. Candès and D. L. Donoho, "Curvelets and curvilinear integrals," *Journal of Approximation Theory*, vol. 113, no. 1, pp. 59–90, 2001.
- [14] 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.
- [15] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampling contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.
- [16] D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss, "Sparse multidimensional representation using shearlets," in *Proceedings of the Wavelets XI*, pp. 1–9, USA, August 2005.
- [17] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 25–46, 2008.
- [18] S. Mehta and B. Marakarkandy, "CT and MRI image fusion using curvelet transform," *Journal of Information, Knowledge and Research in Electronics and Communication Engineering*, vol. 2, no. 2, pp. 848–852, 2013.
- [19] 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.
- [20] S. Singh, D. Gupta, R. S. Anand, and V. Kumar, "Nonsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network," *Biomedical Signal Processing and Control*, vol. 18, pp. 91–101, 2015.
- [21] X. Liu, W. Mei, H. Du, and J. Bei, "A novel image fusion algorithm based on nonsubsampling shearlet transform and morphological component analysis," *Signal, Image and Video Processing*, vol. 10, no. 5, pp. 959–966, 2016.
- [22] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [23] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 884–892, 2010.
- [24] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Information Fusion*, vol. 13, no. 1, pp. 10–19, 2012.
- [25] J. Wang, J. Peng, X. Feng, G. He, J. Wu, and K. Yan, "Image fusion with nonsubsampling contourlet transform and sparse representation," *Journal of Electronic Imaging*, vol. 22, article 043019, no. 4, 2013.
- [26] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion*, vol. 24, pp. 147–164, 2015.
- [27] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [28] J. Huang, X. Huang, and D. Metaxas, "Learning with dynamic group sparsity," in *Proceedings of the 12th International Conference on Computer Vision, ICCV 2009*, pp. 64–71, jpn, October 2009.
- [29] M. Deshmukh and U. Bhosale, "Image fusion and image quality assessment of fused images," *International Journal of Image Processing*, vol. 4, no. 5, pp. 484–508, 2010.
- [30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [31] G. H. Qu, D. L. Zhang, and P. F. Yan, "Information measure for performance of image fusion," *IEEE Electronics Letters*, vol. 38, no. 7, pp. 313–315, 2002.
- [32] V. Aslantas and E. Bendes, "A new image quality metric for image fusion: the sum of the correlations of differences," *AEÜ - International Journal of Electronics and Communications*, vol. 69, no. 12, pp. 1890–1896, 2015.
- [33] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.
- [34] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proceedings of the International Conference on Image Processing (ICIP '03)*, vol. 3, pp. 173–176, September 2003.
- [35] C. S. Xydeas and V. Petrović, "Objective image fusion performance measure," *IEEE Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.
- [36] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Information Fusion*, vol. 14, no. 2, pp. 127–135, 2013.



**Hindawi**  
Submit your manuscripts at  
<https://www.hindawi.com>

