ORIGINAL ARTICLE

A sum-modified-Laplacian and sparse representation based multimodal medical image fusion in Laplacian pyramid domain

Xiaoqing Li[1] · Xuming Zhang[1] · Mingyue Ding[1] 🔵

Abstract

Fusion of multimodal medical images provides complementary information for diagnosis, surgical planning, and clinical outcome evaluation. Although the multiscale decomposition–based fusion methods have attracted much attention among researchers, the challenges of determining the decomposition levels and the loss of contrast hindered their applications. Here, we present a multimodal medical images fusion method combining the sum-modified-Laplacian (SML) with sparse representation (SR) in the Laplacian pyramid domain. In this method, we first transformed the original images into the high-pass and low-pass bands by the Laplacian pyramid (LP). Then, we use SML and SR to fuse the high- and low-pass bands, respectively. The proposed method has been compared with different methods including NSST_VGG_MAX, DWT_ARV_BURTS, CVT_MAX_LIS, and NSCT_SR_MAX. We also conducted multiple experiments on four groups of medical images, including CT and MR, T1-weighted MR and T2-weighted MR, PET and MR, as well as SPECT and MR, to demonstrate the advantages of our method. Visual and quantitative results illustrate that our method can produce the fused images with better brightness contrast and retain more image details than other evaluated methods on the basis of MI, $L^{AB/F}$, $Q^{AB/F}$, and Q^w. Furthermore, our method could preserve more fine and useful functional information with better image contrast, which is highly relevant in the assessment of lesion shapes and positions.

Keywords Sparse representation · Sum-modified-Laplacian · Laplacian pyramid · Medical image fusion

1 Introduction

With the development of various specialized medical imaging equipment types, medical imaging has become an important tool in clinical diagnosis and treatment. Different modalities of medical images were used to explore diverse and specific aspects of the human bodies and can be typically divided into functional imaging and anatomical imaging [1]. Functional imaging reflects the metabolic and physiological changes of the human body with usual low spatial resolution [2–4]. Anatomical imaging clarifies the shape, position, and relationship of structures like the bones [5–7]. For example, high-resolution CT images can clearly examine the bones and joints due to their high-density properties, while MR images excel at distinguishing the soft tissue types between the nerves, tendons, ligaments, blood vessels, cartilages, and so on based on their varying magnetic susceptibilities. The combination of CT and MR images can complementarily display the pathophysiological changes and morphological structures of lesions, which could improve the accuracy of diagnosis. For the convenient, precise, and effective diagnosis of diseases by the doctors, a fusion of the various modalities of medical images is needed to generate the comprehensive images which highlight both the metabolic changes and anatomic structures to clearly display the exact location and severity of pathological tissues.

Various multimodal medical image fusion techniques have been developed. These techniques are usually separated into spatial domain techniques and frequency domain techniques. Spatial domain techniques just fuse the source images with the contrast reduction and spectral characteristics distortion [8–11]. A fusion method has been introduced based on backpropagation (BP) neural network and human visual system (HVS) by Yang et al. However, this method is unsuitable

✉ Mingyue Ding
myding@hust.edu.cn

1 Department of Biomedical Engineering, School of Life Science and Technology, Ministry of Education Key Laboratory of Molecular Biophysics, Huazhong University of Science and Technology, Wuhan 430074, China

to fuse medical images such as ultrasonic images due to high sensitivity to image noise [12]. Li brought up a medical image fusion approach by combining the spiking cortical model (SCM) with the rolling guidance filter (RGF) [13]. Due to the self-adaptation of SCM, this fusion method can overcome high noise sensitivity, the distortion of fused images, and image information loss. Compared with the spatial domain techniques, the frequency domain techniques have attracted more interests since they could retain more weak structures with better brightness-contrast, strong robustness, and less sensitivity to noise. Frequency domain techniques transform the source images into the different frequency subbands, fuse these bands with different rules and perform the inverse multi-scale decomposition to reconstruct the fused image.

Multi-scale decomposition transforms the source images into multiple sub-images at different frequency subbands. Multi-scale decomposition techniques include several different groups: the techniques based on pyramid [14–16], wavelet [17–19], geometric analysis [20], or contourlet [1]. In ref. 21, a novel approach based on the gradient map has been proposed to fuse multi-resolution images without information loss, which significantly improves the reliability of the information fusion processes. Venkataraman developed a DTCWT-based fusion method which utilized the strong focused image and the initial fused image to fuse the low-dose CT images [22]. This method is able to generate a fusion image with effective noise suppression at the cost of the loss of some critical details. Petrović depicted a NSCT-based fusion method for medical images with enhanced details and the improved visual effect. The authors proposed a fusion method based on the region mosaicking on Laplacian pyramids (RMLP) with the insensitivity to noise and the reduced color distortion [23]. However, there are two drawbacks in the multi-scale decomposition including the loss of contrast and the difficulty in selecting decomposition level.

Sparse representation [24–27] has attracted much attention in signal processing and has a vigorous development in the past 20 years. In ref. 28, a novel fusion approach based on the sparse representation radiomics (SRR) system has been proposed for brain tumor diagnosis It can quantify the high-throughput texture features of images with improved specificity and accuracy. The authors have shown a fusion method based on nonsubsampled shearlet transform (NSST) and SR for CT and MR images. [29] The subjective and objective evaluations both showed that their method can outperform other compared fusion methods, but the sensitivity to noise is still a problem. SR-based methods can prevent fine details from being smoothed, which would achieve better spatial consistency.

Therefore, the subbands at different levels can capture specified information of edges and textures, the rules for fusing different frequency bands need to be carefully determined. Here, we propose a fusion algorithm based on SML and RP in

LP domain. In the proposed method, LP has been used for transforming the source images into the different frequency subbands, while the low-pass bands are fused by SR, and SML is used to fuse the high-pass bands. Our method is evaluated on four different pairs of medical images, including CT-MR, T1-weighted MR and T2-weighted MR, PET-MR, and SPECT-MR. The results indicate that our method performance is clearly superior to the other methods, according to four performance metrics including MI, $L^{AB/F}$, $Q^{AB/F}$, and Q^w, which retain more functional details with better contrast for the diagnosis of pathological lesions.

The rest sections are arranged as follows. Section 2 depicts our fusion method. Section 3 includes the fusion results and evaluation metrics. Finally, Section 4 provides the conclusion.

2 Methods

In this section, a novel method for fusing multimodal medical images is presented based on LP, SML, and SR. The multi-scale decomposition based on LP, combining the predictive characteristics with the frequency transform methods, is adopted to transform the source images into the multiscale subbands. Due to the effective characterization of the image edge feature information, SML [30] is employed for high-pass bands fusion. SR is selected as the low-pass bands' fusion rule because of its improvements in brightness-contrast robustness and good performance on anti-noise [31].

2.1 Multi-scale transformation based on Laplacian pyramid

Multi-scale decomposition based on the Laplacian pyramid could generate a series of image sets that are gradually reduced in the size of LP from the same source image with relatively simple but effective local computations. The multi-resolution images are produced through applying the Gaussian filter, down-sampling of input image, and the interpolation expansion for the Gauss Pyramid. For image analysis, the low-pass information is the slowly/smoothly changing part of the image, which typically indicates the larger frame, outline of shapes, while the high-pass information is the part that changes rapidly and generally contains the details of the image. Moreover, the LP-based approaches, which combine predictive characteristics with the frequency transform methods, are definitely suitable for image analysis because the perfect reconstruction in these methods can lead to subtle loss of image details. The LP is acquired by computing the difference between the neighboring layers of the Gaussian pyramid. The kth layer of the Gaussian pyramid is formulated as:

$$S_k(\mathrm{m},n) = \sum_{i=-2}^{2}\sum_{j=-2}^{2} w(i,j)S_{k-1}(2m+i,2n+j), \qquad (1)$$
$$k\in[1,K), m\in[0,R_k), n\in[0,H_k)$$

where K represents the layer number, $w(i,j)$ is a window function with a Gaussian low-pass filter, while R_k and H_k are the numbers of columns and rows in the kth layer pyramid respectively. The later layers of the Gaussian pyramid can be acquired by the former layer through a window by Eq. (1). The LP is defined as follows:

$$\mathrm{LP}_k = \begin{cases} S_k - S_{k+1}^{*}, & k\in[0,K) \\ S_k, & k = K \end{cases} \qquad (2)$$

where S_{k+1}^{*} is the expanded S_{k+1} with the same size of S_k.

2.2 Sum-modified-Laplacian

Sum-modified-Laplacian (SML) can efficiently reflect the edge feature information and appropriately mirror the focusing characteristics and clarity of the image, to acquire better visual characteristics, as well as more detailed information and better fusion effect. The SML is formulated as:

$$SML(m,n) = \sum_{i=m-a}^{m+a}\sum_{j=n-b}^{n+b} \nabla^2 I(i,j) \qquad (3)$$

where the small window of the size $(2a+1)\times(2b+1)$ is defined around the pixel (i,j) and the $\nabla^2 I(i,j)$ is the gradient which is defined as follows:

$$\nabla^2 I(i,j) = |2I(i,j) - I(i-l,j) - I(i+l,j)|$$
$$+ |2I(i,j) - I(i,j-l) - I(i,j+l)| \qquad (4)$$

where $I(i,j)$ is the pixel value at the position (i,j) and l is the step between different coefficients.

2.3 Sparse representation

Sparse representation (SR) aims to succinctly express the signals through a linear of least atoms from an overcomplete dictionary, which makes it easier to get the meaningful messages out of signals and more convenient for further signal processing. The two major tasks of SR are generating dictionary and sparse coding. Due to the strong self-adaptation, dictionary learning is commonly used for dictionary generation, which includes the method of optimal directions MOD algorithm, generalized PCA algorithm, and K-SVD algorithm. Moreover, sparse coding performs on local patches

and calculates the sparest α with the fewest nonzero atoms. $V\in \mathbf{R}^n$ means a signal vector from the source images and $\mathbf{D}\in \mathbf{R}^{n\times m}$ $(m>n)$ is an overcomplete dictionary. The signal vector is formulated as $V=\mathbf{D}\alpha$, and $\alpha\in \mathbf{R}^k$ means the sparse coefficient vector. The SR coding is presented as follows:

$$\hat{\alpha} = \underset{\alpha}{\mathrm{argmin}}\, \|\alpha\|_0 \quad s.t. \quad \|V-\mathbf{D}\alpha\|_2 < \varepsilon \qquad (5)$$

where $\|\alpha\|_0$ denotes the number of nonzero entries in α and ε is a permit-able error. Calculating the sparse coefficient vectors α utilizes the orthogonal matching pursuit (OMP) algorithm [25], which is a representative method to solve Eq. (5). SR can improve image contrast and be more robust to noise.

2.4 The proposed fusion method framework

The proposed fusion method is depicted in Fig. 1, which is composed of four procedures: LP decomposition, low-pass bands fusion, high-pass bands fusion, and LP reconstruction.
    The proposed method

1. Procedure multi-scale decomposition based on LP
2. Perform LP on the source images $\{I_E, I_N\}$ to acquire the high-pass bands $\{H_E, H_N\}$ and the low-pass bands $\{L_E, L_N\}$ using Eqs. (1) and (2a.). The level of decomposition is set to be 4.
3. Procedure low-pass bands fusion based on SR
4. Set the parameters for dictionary learning, including the dictionary dimensions, number of examples, sparsity of each example, and input the image patches of 10 high-quality medical images, run K-SVD training 1000 iterations, and obtain the overcomplete dictionary ($\mathbf{D}$) with a size of 64×256.
5. By using the sliding window technique, the original images I_E and I_N are divided into the patches. The patch size is set to be 8×8, and the step length is set to be 1.
6. Rearrange the patches generated above into the column vector and normalize the column vectors. The sparse coefficient vectors α_X^i are calculated by
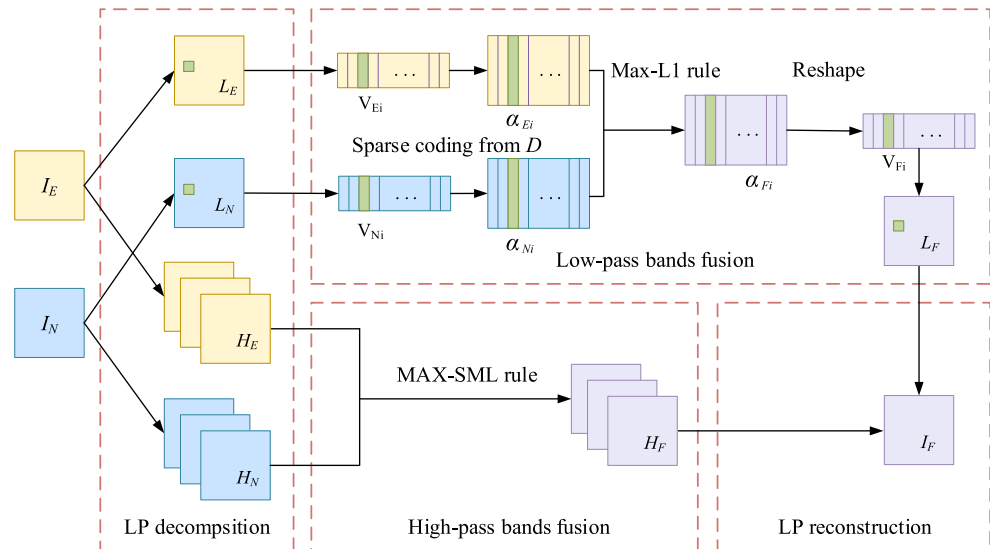
$$\alpha_X^i = \mathrm{argmin}\|\alpha\|_0 \quad s.t. \quad \left\|\hat{v}_X^i - \mathbf{D}\alpha\right\|_2 < \varepsilon \qquad (6)$$

$$\hat{v}_X^i = \mathbf{v}_X^i - \overline{v}_X^i \cdot \mathbf{1} \qquad (7)$$

where $\mathbf{1}$ denotes an all-one valued $n\times1$ vector, $\overline{v}_X^i$ is the mean value of all elements in $\mathbf{v}_X^i$.

7. Obtain the fused sparse column vector α_F^i by merging α_E^i and α_N^i with the "maximal L1-norm" rule.

Fig. 1 Schematic framework of the proposed fusion method based on RP and SML

The merged α^i_F is calculated by:

$$\alpha^i_F = \begin{cases} \alpha^i_E, if\ \left\| \alpha^i_E \right\|_1 \geq \left\| \alpha^i_N \right\|_1 \\ \\ \alpha^i_N, \quad \text{otherwise} \end{cases} \tag{8}$$

The fused vector is calculated by:

$$\mathbf{v}^i_F = \mathbf{D}\alpha^i_F + \overline{v}^i_F \cdot \mathbf{1} \tag{9}$$

$$\overline{v}^i_F = \begin{cases} \overline{v}^i_E, if\ \alpha^i_F = \alpha^i_E \\ \\ \overline{v}^i_N, \text{otherwise} \end{cases} \tag{10}$$

where $\mathbf{v}^i_F$ is the fused vectors for the patch reconstruction.

8. All the patches are overlapped and the overlapped pixels between two neighbor patches are set to be 6, and the sparse reconstruction error is set to be 0.1 for the images. The fused column vectors are reshaped into the patches and plugged into its original position.
9. Procedure high-pass bands fusion based on SML
10. Merge the high-pass bands with SML rule as follows:

$$C^F(i,j) = \begin{cases} C^E(i,j), if\ SML^N(i,j) \geq SML^E(i,j) \\ \\ C^N(i,j), \text{otherwise} \end{cases} \tag{11}$$

where $C^E(i, j)$ and $C^N(i, j)$ represent the pixel values at the position (i, j) of the high-pass bands H_E and H_N, separately. The step in Eq. (4) is set to be 1.

11. Procedure reconstruction
12. Perform LP inverse multi-scale transform to fuse the merged high-pass and low-pass bands to reconstruct the final images.

3 Experimental results

3.1 Image quality metrics for the fused image

Usually, the image quality assessment is separated into two major types: subjective quality assessment and objective quality assessment. Due to the inconsistent evaluation by human subjects, subjective quality assessment is not as stable as objective quality assessment. Hence, the various objective image quality metrics are proposed to assess the images in a quantitative and automatical way. Four commonly used metrics, including mutual information (MI), the local structural similarity ($L^{AB/F}$), edge-based similarity measure ($Q^{AB/F}$), and the universal image quality index (Q^w), are applied to evaluate the fusion performance of all methods. With regard to all these four metrics, the larger values mean better results. Brief discussions of the metrics are presented below.

3.1.1 Mutual information

As a well-known technique for evaluating the image quality of the similarity, mutual information (MI) clarifies the relationship between two variables and the strength of its relationship. The larger value of MI represents the greater amount of mutual information shared. The MI is defined as follows:

$$\text{MI} = \text{MI}_{AF} + \text{MI}_{BF} \tag{12}$$

$$\text{MI}_{AF} = \sum_{f=0}^{L-1} \sum_{a=0}^{L-1} p_{AF}(a,f) \log_2 \left(\frac{p_{AF}(a,f)}{p_A(a)p_F(f)} \right) \tag{13}$$

$$\text{MI}_{BF} = \sum_{f=0}^{L-1} \sum_{b=0}^{L-1} p_{BF}(b,f) \log_2 \left(\frac{p_{BF}(b,f)}{p_B(b)p_F(f)} \right) \tag{14}$$

where L is the number of the bins, $p_{AF}(a,f)$ and $p_{BF}(b,f)$ represent the joint gray-level histograms of the fused image and

the original images, while $p_A(a)$, $p_B(b)$, and $p_F(f)$ represent the normalized gray-level histograms of the two source images and fused image, respectively. MI indicates how much source image information is conveyed by the fused image.

3.1.2 Edge-based similarity measure ($L^{AB/F}$)

The $L^{AB/F}$ computes the similarity of the transferred edges from the original images to the final image. The higher value of $L^{AB/F}$ metric represents better edge similarity. The $L^{AB/F}$ is formulated as follows:

$$L^{AB/F} = \frac{\sum_{i=1}^{H}\sum_{j=1}^{W}\left[Q^{AF}(i,j)w^A(i,j) + Q^{BF}(i,j)w^B(i,j)\right]}{\sum_{i=1}^{H}\sum_{j=1}^{W}\left[w^A(i,j) + w^B(i,j)\right]}$$

$$\tag{15}$$

$$Q^{AF}(i,j) = Q_g^{AF}(i,j)Q_\alpha^{AF}(i,j) \tag{16}$$

$$Q^{BF}(i,j) = Q_g^{BF}(i,j)Q_\alpha^{BF}(i,j) \tag{17}$$

where $w^A(i,j)$ and $w^B(i,j)$ represent the corresponding gradient strengths for the original images. $Q_g^{AF}(i,j)$, $Q_g^{BF}(i,j)$, $Q_\alpha^{AF}(i,j)$, and $Q_\alpha^{BF}(i,j)$ represent the orientation preservation values and edge intensity at the position (i,j) for the two input images, respectively.

3.1.3 Local structural similarity-based quality metrics ($Q^{AB/F}$)

$Q^{AB/F}$ can assess the local structural similarity and is presented by:

$$LSSIM(I,J) = [s(i,j)]^\alpha [l_n(i,j)]^\beta [c_n(i,j)]^\gamma \tag{18}$$

$$Q^{AB/F} = \frac{1}{|W|}\sum_{w\in W}[\lambda(w)LSSIM(a,f|w) + (1-\lambda(w))LSSIM(b,f|w)]$$

$$\tag{19}$$

$$\lambda(w) = \frac{s(a|w)}{s(a|w) + s(b|w)} \tag{20}$$

where the parameters α, β, and γ are used to adjust the relative weight of the structure, luminance, and contrast components, and are all set to be 1. $s(i,j)$, $l_n(i,j)$, and $c_n(i,j)$ represent the three components mentioned above and can be derived using the method mentioned in [32]. $s(a|w)$ and $s(b|w)$ are the variances of the source images in the sliding window to calculate $\lambda(w)$. The sliding window is set to be 5×5. The higher value of $Q^{AB/F}$ metric means the greater amount of edge information contained in the fused image.

3.1.4 Universal image quality index–based fusion metric (Q^w)

Q^w represents the distortions of illumination, coefficient correlation, and contrast between the original images and the final

images. The higher value of the Q^w metric means less distortion of image quality. The Q^w is defined as:

$$Q^w = \sum_{w\in W} c(w)(\lambda(w)Q_0(A,F|w))$$

$$+ ((1-\lambda(w))Q_0(B,F|w)) \tag{21}$$

where $Q_0(A,F|w)$ and $Q_0(B,F|w)$ can be obtained by in a local sliding window w. [33] The saliency weight $\lambda(w)$ is defined as follows:

$$\lambda(w) = \frac{s(a|w)}{s(a|w) + s(b|w)} \tag{22}$$

where the salience measures $s(A|w)$ and $s(B|w)$ are calculated with the variances of A and B in window w, respectively. $c(w)$ is the normalized salience of w among all the local windows, which is calculated by:

$$c(w) = \frac{\max(s(A|w), s(B|w))}{\sum_{w'\in W}\max(s(A|w'), s(B|w'))} \tag{23}$$

3.2 Results and discussion

We used the medical images from the Harvard Whole Brain Atlas [34] to test the adaption capacity and robustness of the fusion methods. These images are classified into four pairs: (1) CT and MR [31], (2) T1-weighted MR and T2-weighted MR, (3) PET and MR, and (4) SPECT and MR. All images have the same size of 256×256 pixels. Figure 2 a, h, and o are the MR images of a human brain, providing good contrasts between gray matter, white matter, and cerebrospinal fluid of the brain, while Fig. 2 b, i, and p are the corresponding CT images, providing enhanced contrasts for bone-related tissues. Figure 3 a, h, and o show the T1-weighted images of the brain hemisphere, while Fig. 3 b, i, and p show the corresponding T2-weighted MR images. Figure 4 shows the MR (first left most column) and PET (second left most column) images of the brain hemisphere. Figure 5 shows the MR (first left most column) and SPECT images (second left most column) at the brain hemisphere with glioma. The presented method and the other evaluated methods are applied to fuse the above-mentioned medical images.

For all these four groups of images, the results of LP_SR_SML method are compared with the following methods: (1) fusion method based on the nonsubsampled shearlet transform using the VGG and the absolute-max value fusion rules (NSST_VGG_MAX); (2) fusion method based on the discrete wavelet transform using the average value and Burt's fusion rules (DWT_ARV_BURTS); (3) fusion method based on curvelet transform-based using the absolute-max value and Lis rules (CVT_MAX_LIS); (4)
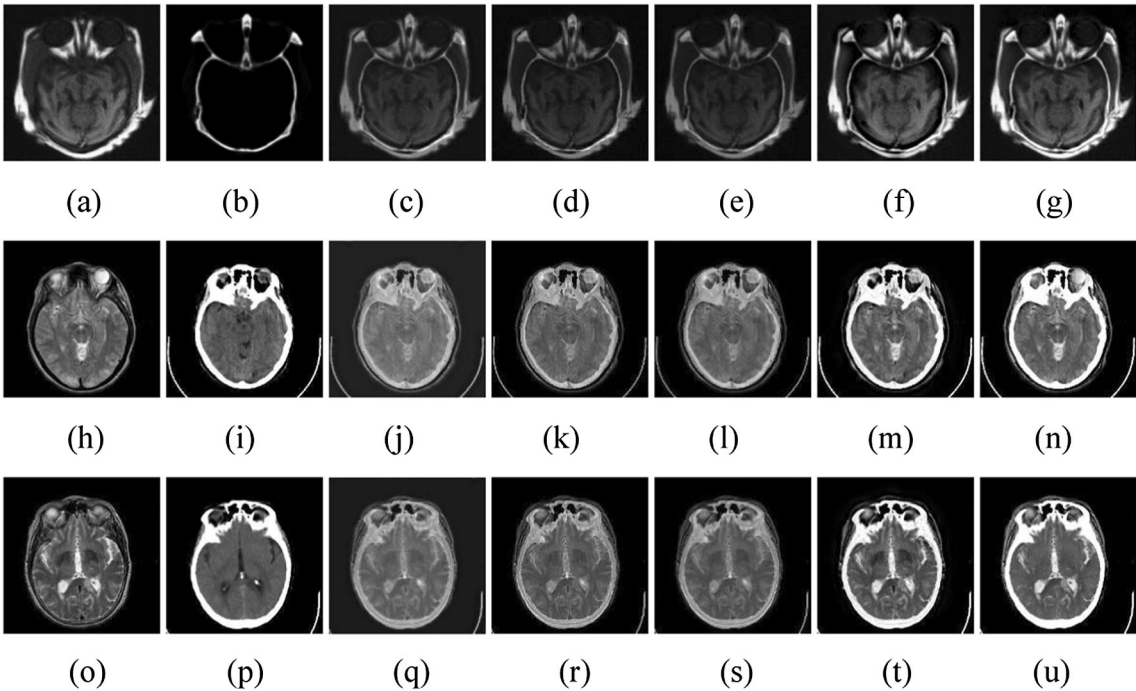
Fig. 2 The fusion results produced by different methods for three groups of CT-MR images. a–b, h–i, and o–p are the source images needed to be fused. c–g, j–n, and q–u are fusion results produced by NSST_VGG_ MAX, DWT_ARV_BURTS, CVT_MAX_LIS, NSCT_SR_MAX, and LP_SR_SML, respectively

NSCT-based fusion method using the SR and the absolute-max value (NSCT_SR_MAX). For a fair comparison, we applied the same parameters for each compared method implemented on all images.

Figure 2 shows the experimental results of three groups of MR and CT images. From Fig. 2, we can see that our method can successfully preserve both the features such as the bony structures highlighted in the CT and the high metabolic areas
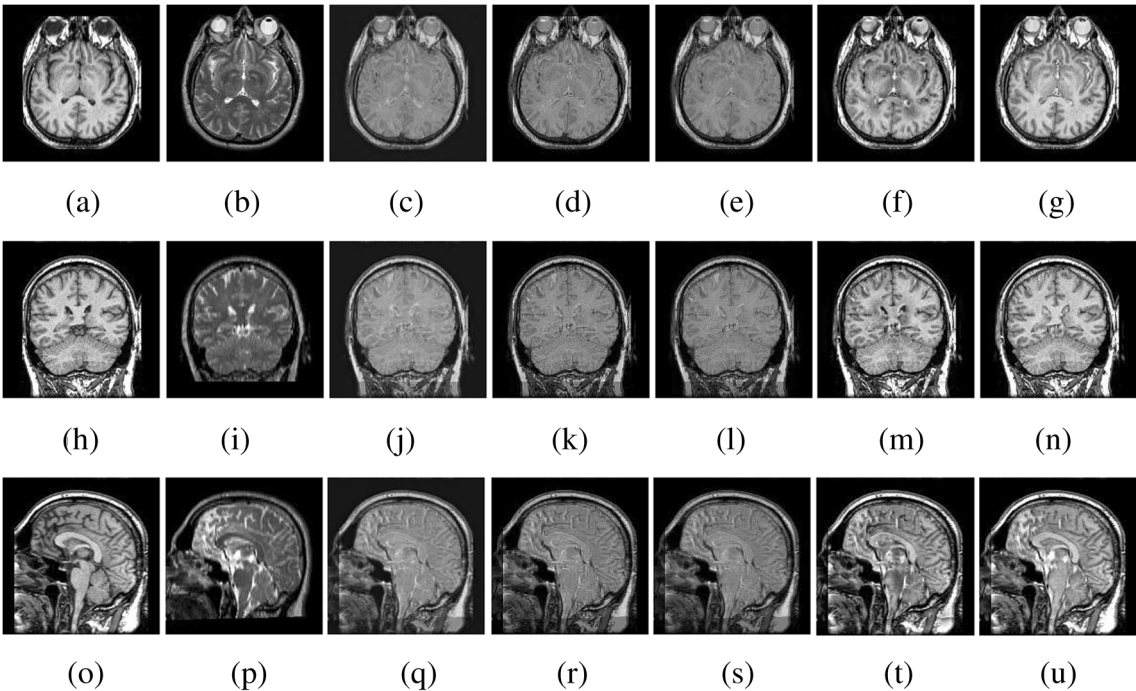

Fig. 3 The fusion results obtained by different methods for MR T1-MR T2 images. a–b, h–i, and o–p are the three groups of the source images needed to be fused. c–g, j–n, and q–u are fusion results produced by NSST_VGG_MAX, DWT_ARV_BURTS, CVT_MAX_LIS, NSCT_ SR_MAX, and LP_SR_SML, respectively

Fig. 4 The fusion results obtained by different methods for MR T1-MR T2 images. a–b, h–i, and o–p are the three groups of the source images needed to be fused. c–g, j–n, and q–u are fusion results produced by NSST_VGG_MAX, DWT_ARV_BURTS, CVT_MAX_LIS, NSCT_SR_MAX, and LP_SR_SML, respectively
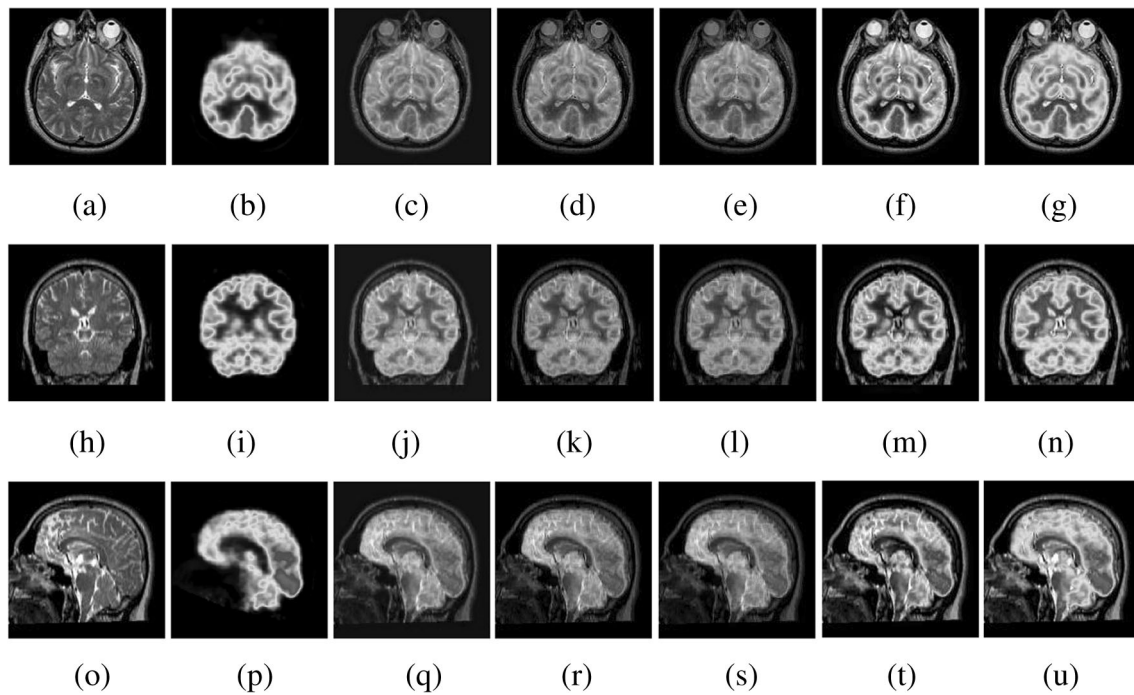
from the MR images. Table 1 shows the objective evaluation metric values, and it is clear that the LP_SR_SML method demonstrates better objective performance compared to the other methods on MI, $Q^{AB/F}$, $L^{AB/F}$, and Q^w, which indicates that most high-frequency details are conserved by the proposed method. The above analysis indicates that our method performs superior to the other four methods for the fusion of MR and CT images.
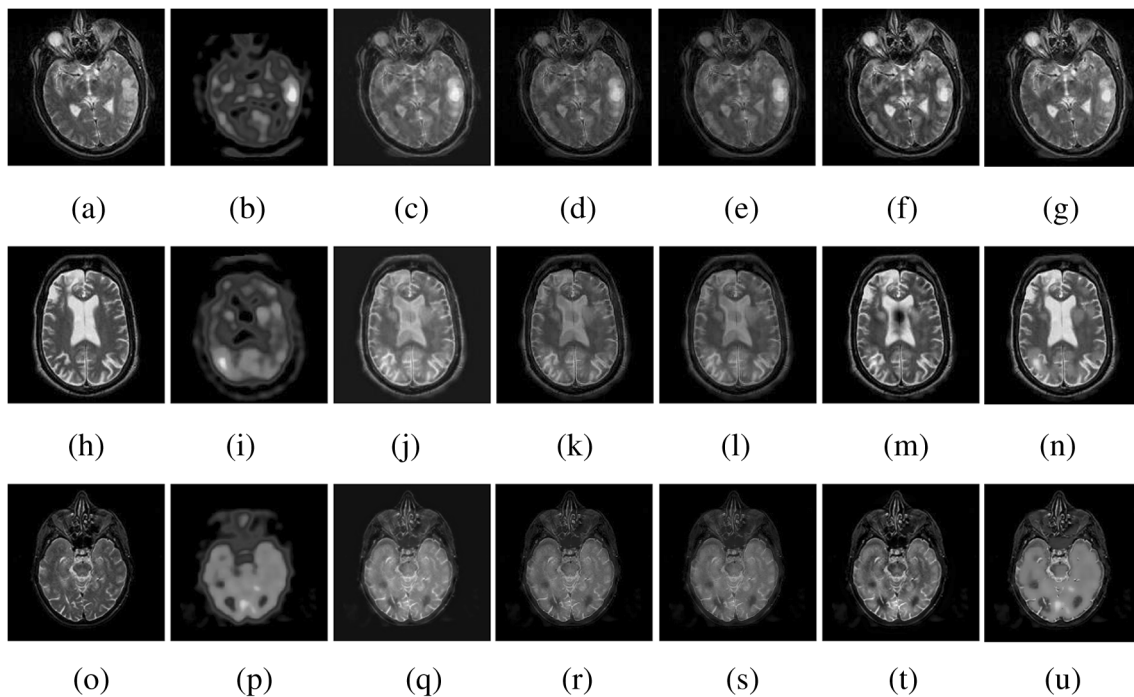

Fig. 5 The fusion results obtained by different methods for MR T1-MR T2 images. a–b, h–i, and o–p are the three groups of the source images needed to be fused. c–g, j–n, and q–u are fusion results produced by NSST_VGG_MAX, DWT_ARV_BURTS, CVT_MAX_LIS, NSCT_SR_MAX, and LP_SR_SML, respectively

Table 1 Objective metrics for each fused method implemented in Fig. 2

Group	Metrics	Fusion methods				
		NSST_VGG_MAX	DWT_ARV_BURTS	CVT_MAX_LIS	NSCT_SR_MAX	LP_SR_SML
Group 1	MI	3.2062	3.2610	4.0868	3.0122	4.8064
	$Q^{AB/F}$	0.4882	0.4793	0.4233	0.7367	0.7858
	$L^{AB/F}$	0.7128	0.6651	0.6540	0.8478	0.9540
	Q^w	0.8136	0.7171	0.7138	0.8604	0.9264
Group 2	MI	2.3258	3.1997	3.1646	3.1353	3.8590
	$Q^{AB/F}$	0.3228	0.4555	0.3985	0.6194	0.6438
	$L^{AB/F}$	0.3533	0.8237	0.7382	0.7249	0.9233
	Q^w	0.3646	0.8304	0.8264	0.8151	0.8395
Group 3	MI	3.2786	3.3599	3.3720	3.4173	4.0349
	$Q^{AB/F}$	0.2889	0.4180	0.3733	0.5741	0.5898
	$L^{AB/F}$	0.3505	0.8293	0.7520	0.7579	0.9200
	Q^w	0.3636	0.8275	0.8249	0.8174	0.8228

Figures 3 shows the fused images of three groups of T1-weighted and the T2-weighted MR images of the brain hemisphere by the five fusion methods. From Fig. 3, the fusion results indicate that our method not only provides a better contrast but also preserves more precise texture information, which is more consistent with original images. It can be seen that our method provides the best objective metrics on MI, $Q^{AB/F}$, $L^{AB/F}$, and Q^w for T1-weighted MR and T2-weighted MR images, as shown in Table 2. The comparison for the evaluated methods indicates that the LP_SR_SML method maintains the texture information well, especially the weak information in the original images.

Figures 4 and 5 show the MR/PET and MR/SPECT images of the brain hemisphere and the images fused by the five fusion methods. The fusion results demonstrate that our method can preserve more texture information, especially the detailed information of the pathological shapes and positions. The LP_SR_SML method provides the highest objective metrics such as MI, $Q^{AB/F}$, $L^{AB/F}$, and Q^w for fused MR and PET images according to Table 3. In Table 4, it is clear that our method is superior in terms of MI and $L^{AB/F}$ for MR- and SPECT-fused images. Although the objective metrics $Q^{AB/F}$ and Q^w are not the best, the proposed method produces the fused images that are easier for doctors to ascertain the shape and position of the lesion due to the better demonstration of functional details in Fig. 5.

For statistical validation of these extensive comparison results, multiple t tests are performed among all the methods (Table 5). Between the LP_SR_SML method and other compared methods, most of them are significant differences as

Table 2 Objective metrics for each fused method implemented in Fig. 3

Group	Metrics	Fusion methods				
		NSST_VGG_MAX	DWT_ARV_BURTS	CVT_MAX_LIS	NSCT_SR_MAX	LP_SR_SML
Group 1	MI	2.9565	2.9883	2.9874	3.1458	4.8754
	$Q^{AB/F}$	0.3240	0.4058	0.3789	0.5536	0.6255
	$L^{AB/F}$	0.3884	0.7811	0.6894	0.8145	0.9491
	Q^w	0.3817	0.7510	0.7541	0.7791	0.7737
Group 2	MI	2.8594	2.8692	2.8042	3.2816	5.0132
	$Q^{AB/F}$	0.3965	0.4727	0.4417	0.6559	0.7080
	$L^{AB/F}$	0.4645	0.8268	0.7561	0.8649	0.9694
	Q^w	0.4572	0.7924	0.7894	0.8478	0.8557
Group 3	MI	2.5700	2.5524	2.5614	2.7569	4.2249
	$Q^{AB/F}$	0.3577	0.4125	0.3785	0.5795	0.6485
	$L^{AB/F}$	0.5352	0.7651	0.6727	0.8460	0.9372
	Q^w	0.5206	0.7343	0.7319	0.7921	0.7777

Table 3 Objective metrics for each fused method implemented in Fig. 4

Group	Metrics	Fusion methods				
		NSST_VGG_MAX	DWT_ARV_BURTS	CVT_MAX_LIS	NSCT_SR_MAX	LP_SR_SML
Group 1	MI	2.9791	2.8882	2.9655	2.7589	4.0235
	$Q^{AB/F}$	0.3723	0.4223	0.3993	0.6376	0.7067
	$L^{AB/F}$	0.4436	0.8045	0.7551	0.8192	0.9657
	Q^w	0.4629	0.8061	0.8083	0.8590	0.8757
Group 2	MI	3.0850	3.0235	3.1279	2.8281	4.0882
	$Q^{AB/F}$	0.4207	0.4047	0.3956	0.6181	0.6906
	$L^{AB/F}$	0.4421	0.8118	0.7612	0.7921	0.9633
	Q^w	0.4528	0.8181	0.8219	0.8505	0.8862
Group 3	MI	3.2987	3.2545	3.4449	3.0878	4.2087
	$Q^{AB/F}$	0.3815	0.4231	0.3922	0.6314	0.6914
	$L^{AB/F}$	0.5026	0.7731	0.7141	0.8233	0.9530
	Q^w	0.5341	0.7813	0.7800	0.8524	0.8599

shown in Table 5 ($p < 0.05$). Because $Q^{AB/F}$ and Q^w are influenced by the local spatial details preserved, the LP_SR_SML method preserves more useful information in the functional images and may cover some meaningless local anatomic structure, leading to the fact that there are no significant differences ($p > 0.05$) between the two metrics of the LP_SR_SML method and those of the NSCT_SR_MAX method. The two fusion rules facilitate our fusion method with less loss of contrast and better brightness and preserve more useful information. The low-pass bands include most energy of the image, which is more important to human vision. Hence, SR is suitable for low-pass subbands fusion with better brightness-contrast and more retained spatial details. The high-pass subbands correspond to the image edges and contour with fast and obvious brightness changes with sensitivity to the noise. SML is proper

for high-pass bands fusion due to its good characteristics in preserving the edges, lines, and boundaries for better information interpretation. Due to the combination of SR and SML, our proposed method outperforms the other methods in both the subjective and objective way. Our method preserves more information from fused image sources that help facilitate better disease diagnosis.

4 Conclusion

In this research, a novel fusion method based on SML and SR in the Laplacian pyramid domain is proposed for the multimodal medical image fusion. Our method utilizes LP to acquire the high-pass and low-pass subbands. For better

Table 4 Objective metrics for each fused method implemented in Fig. 5

Group	Metrics	Fusion methods				
		NSST_VGG_MAX	DWT_ARV_BURTS	CVT_MAX_LIS	NSCT_SR_MAX	LP_SR_SML
Group 1	MI	3.2750	3.2260	3.3403	3.7265	5.8928
	$Q^{AB/F}$	0.4964	0.4859	0.4361	0.7211	0.7558
	$L^{AB/F}$	0.5329	0.7674	0.7355	0.9246	0.9643
	Q^w	0.5536	0.8188	0.8106	0.9039	0.9027
Group 2	MI	3.2986	3.2539	3.2732	3.3913	4.7803
	$Q^{AB/F}$	0.5479	0.4694	0.4201	0.6932	0.7083
	$L^{AB/F}$	0.4359	0.8321	0.8050	0.8836	0.9385
	Q^w	0.4423	0.8469	0.8436	0.8887	0.8820
Group 3	MI	3.1487	3.0682	3.0980	2.9139	3.8383
	$Q^{AB/F}$	0.5184	0.4675	0.4132	0.6668	0.567
	$L^{AB/F}$	0.3658	0.8591	0.8197	0.8757	0.9330
	Q^w	0.4004	0.8796	0.8729	0.9085	0.8437

Table 5 p value of overall objective metrics for each fused method

Fusion methods	MI p value	$Q^{AB/F}$ p value	$L^{AB/F}$ p value	Q^w p value
NSST_VGG_MAX	2.61×10^{-7}	1.46×10^{-8}	9.62×10^{-14}	1.54×10^{-9}
DWT_ARV_BURTS	2.49×10^{-7}	1.21×10^{-10}	1.39×10^{-9}	0.01
CVT_MAX_LIS	3.21×10^{-6}	2.45×10^{-12}	2.15×10^{-12}	7.57×10^{-3}
NSCT_SR_MAX	6.67×10^{-7}	0.16	5.53×10^{-7}	0.74
LP_SR_SML	–	–	–	–

information interpretation, SML is used to fuse the high-pass bands and improve the characteristics of edges, lines, and boundaries, while SR is applied to fuse the low-pass bands and provide better brightness and contrast. The combination of these two fusion rules can facilitate better preservation of more details, and it can further effectively improve the interpretation of the fused results to help doctors identify diseases and conditions more accurately. The experiments on four groups of medical images indicate the superior performance of LP_SR_SML method. Evaluations based on the objective metrics, including MI, $Q^{AB/F}$, $L^{AB/F}$, and Q^w, illustrate that the proposed method performs best among all the evaluated methods. Therefore, it can be concluded as one of the most robust medical image fusion methods.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Yang Y, Tong S, Huang S, Lin P (2014) Log-Gabor energy based multimodal medical image fusion in NSCT domain. Comput Math Methods Med 2014:835481. https://doi.org/10.1155/2014/835481
2. Filippi M, Rocca MA (2013) Present and future of fMRI in multiple sclerosis. Expert Rev Neurother 13(12 Suppl):27–31. https://doi.org/10.1586/14737175.2013.865871
3. Sarikaya I (2015) PET imaging in neurology: Alzheimer's and Parkinson's diseases. Nucl Med Commun 36(8):775–781. https://doi.org/10.1097/MNM.0000000000000320
4. Hutton BF (2014) The origins of SPECT and SPECT/CT. Eur J Nucl Med Mol Imaging 41(Suppl 1):S3–S16. https://doi.org/10.1007/s00259-013-2606-5
5. Mühlenweg M, Schaefers G, Trattnig S (2015) Physical interactions in MRI: some rules of thumb for their reduction. Radiologe 55(8):638–648. https://doi.org/10.1007/s00117-015-2812-1
6. Diaconis JN, Rao KC (1980) CT in head trauma: a review. J Comput Tomogr 4(4):261–270
7. Schellpfeffer MA (2013) Ultrasound imaging in research and clinical medicine. Birth Defects Res C Embryo Today 99(2):83–92. https://doi.org/10.1002/bdrc.21032
8. Kim T, Rivara FP, Mozingo DW, Lottenberg L, Harris ZB, Casella G, Liu H, Moldawer LL, Efron PA, Ang DN (2012) A regionalised strategy for improving motor vehicle-related highway driver deaths using a weighted averages method. Inj Prev 18(1):16–21. https://doi.org/10.1136/ip.2010.030759
9. Sainani KL (2014) Introduction to principal components analysis. PM R 6(3):275–278. https://doi.org/10.1016/j.pmrj.2014.02.001
10. Gloi AM, Buchanan R (2013) Dosimetric assessment of prostate cancer patients through principal component analysis (PCA). J Appl Clin Med Phys 14(1):3882–3849. https://doi.org/10.1120/jacmp.v14i1.3882
11. Vollnhals F, Audinot JN, Wirtz T, Mercier-Bonin M, Fourquaux I, Schroeppel B, Kraushaar U, Lev-Ram V, Ellisman MH, Eswara S (2017) Correlative microscopy combining secondary ion mass spectrometry and electron microscopy: comparison of intensity-hue-saturation and Laplacian pyramid methods for image fusion. Anal Chem 89(20):10702–10710. https://doi.org/10.1021/acs.analchem.7b01256
12. Yang Y, Zheng W, Huang S (2014) Effective multifocus image fusion based on HVS and BP neural network. ScientificWorldJournal 2014:281073. https://doi.org/10.1155/2014/281073
13. Shuaiqi L, Jie Z, Mingzhu S (2015) Medical image fusion based on rolling guidance filter and spiking cortical model. Comput Math Methods Med 2015:156043. https://doi.org/10.1155/2015/156043
14. Toet A (1989) Image fusion by a ratio of low pass pyramid. Pattern Recogn Lett 9(4):245–253
15. Burt P, Adelson E (1987) The Laplacian pyramid as a compact image code. Read Comput Vision 31(4):671–679. https://doi.org/10.1109/TCOM.1983.1095851
16. Petrovic V, Xydeas C (2004) Gradient-based multiresolution image fusion. IEEE Trans Image Process 13(2):228–237
17. Jmail N, Zaghdoud M, Hadriche A, Frikha T, Ben Amar C, Bénar C (2018) Integration of stationary wavelet transform on a dynamic partial reconfiguration for recognition of pre-ictal gamma oscillations. Heliyon 4(2):e00530. https://doi.org/10.1016/j.heliyon.2018.e00530
18. Li H, Manjunath B, Mitra S (1995) Multisensor image fusion using the wavelet transform. Graphical Models Image Proc 57(3):235–245
19. Lewis J, OCallaghan R, Nikolov S, Bull D, Canagarajah N (2007) Pixel- and region-based image fusion with complex wavelets. Inform Fusion 8(2):119–130
20. Nencini F, Garzelli A, Baronti S, Alparone L (2007) Remote sensing image fusion using the curvelet transform. Inform Fusion 8(2):143–156
21. Petrović VS, Xydeas CS (2004) Gradient-based multiresolution image fusion. IEEE Trans Image Process 13(2):228–237
22. Venkataraman A, Alirezaie J, Babyn P, Ahmadian A (2014) Multi dose computed tomography image fusion based on hybrid sparse

methodology. Conf Proc IEEE Eng Med Biol Soc 2014:3901–3904. https://doi.org/10.1109/EMBC.2014.6944476

23. Sun J, Han Q, Kou L, Zhang L, Zhang K, Jin Z (2018) Multi-focus image fusion algorithm based on Laplacian pyramids. J Opt Soc Am A Opt Image Sci Vis 35(3):480–490. https://doi.org/10.1364/JOSAA.35.000480

24. Zhang J, Zhao D, Gao W (2014) Group-based sparse representation for image restoration. IEEE Trans Image Process 23(8):3336–3351. https://doi.org/10.1109/TIP.2014.2323127

25. Ptucha R, Savakis AE (2014) LGE-KSVD: robust sparse representation classification. IEEE Trans Image Process 23(4):1737–1750. https://doi.org/10.1109/TIP.2014.2303648.

26. Chen L, Li J, Chen CL (2013) Regional multifocus image fusion using sparse representation. Opt Express 21(4):5182–5197. https://doi.org/10.1364/OE.21.005182

27. Lan X, Ma AJ, Yuen PC, Chellappa R (2015) Joint sparse representation and robust feature-level fusion for multi-cue visual tracking. IEEE Trans Image Process 24(12):5826–5841. https://doi.org/10.1109/TIP.2015.2481325

28. Wu G, Chen Y, Wang Y, Yu J, Lv X, Ju X, Shi Z, Chen L, Chen Z (2018) Sparse representation-based radiomics for the diagnosis of brain tumors. IEEE Trans Med Imaging 37(4):893–905. https://doi.org/10.1109/TMI.2017.2776967

29. Qiu C, Wang Y, Zhang H, Xia S (2017) Image fusion of CT and MR with sparse representation in NSST domain. Comput Math Methods Med 2017:9308745. https://doi.org/10.1155/2017/9308745

30. Yang Y, Tong S, Huang S, Lin P (2014) Dual-tree complex wavelet transform and image block residual-based multi-focus image fusion in visual sensor networks. Sensors (Basel) 14(12):22408–22430. https://doi.org/10.3390/s141222408.

31. Liu Y, Liu S, Wang Z (2015) A general framework for image fusion based on multi-scale transform and sparse representation. Inform Fusion 24:147–164

32. Yang C, Zhang JQ, Wang XR, Liu X (2008) A novel similarity based quality metric for image fusion. Inform Fusion 9:156–160

33. Wang Z, Bovik A, Sheikh H, Simoncelli E (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612

34. Atlas, the Whole Brain. http://www.med.harvard.edu/aanlib/home.html. Accessed 12 Aug 2019

Xuming Zhang is an associate professor at the Department of Biomedical Engineering, Huazhong University of Science and Technology. His research interests include image processing and image-guided interventional surgery navigation technology.

Mingyue Ding is a professor at the Department of Biomedical Engineering, Huazhong University of Science and Technology. His research interests include medical image processing, three-dimensional ultrasound imaging and image-guided surgery.

Xiaoqing Li is currently pursuing her PhD degree at Huazhong University of Science and Technology, Wuhan, China. Her research focuses on image processing and ultrasound imaging.