

## An overview of multi-modal medical image fusion

Jiao Du <sup>a</sup>, Weisheng Li <sup>a,\*</sup>, Ke Lu <sup>b</sup>, Bin Xiao <sup>a</sup>

<sup>a</sup> Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

<sup>b</sup> College of Engineering and Information Technology, University of Chinese Academy of Sciences, Beijing 100049, China



### ARTICLE INFO

#### Article history:

Received 31 January 2015

Received in revised form

30 June 2015

Accepted 1 July 2015

Available online 10 June 2016

#### Keywords:

Multi-modal

Image fusion

Image decomposition

Image reconstruction

Image fusion rules

Image quality assessment

### ABSTRACT

Multi-modal medical image fusion is the process of merging multiple images from single or multiple imaging modalities to improve the imaging quality with preserving the specific features.

Medical image fusion covers a broad number of hot topic areas, including image processing, computer vision, pattern recognition, machine learning and artificial intelligence. And medical image fusion has been widely used in clinical for physicians to comprehend the lesion by the fusion of different modalities medical images. In this review, methods in the field of medical image fusion are characterized by (1) image decomposition and image reconstruction, (2) image fusion rules, (3) image quality assessments, and (4) experiments on the benchmark dataset. In addition, this review provides a factual listing of scientific challenges faced in the field of multi-modal medical image fusion.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Multi-modal medical image fusion is the combination of multiple images from single or multiple imaging modalities. The purpose of the medical image fusion is to improve imaging quality with preserving the specific features for increasing the clinical applicability of images for diagnosis and assessment of medical problems. Medical image fusion methods cover a broad number of areas, including image processing, computer vision, pattern recognition, machine learning and artificial intelligence with wide applications in clinical for physicians to comprehend the lesion by fusing different modalities medical images [1].

Medical image fusion mainly concentrates on magnetic resonance imaging (MRI), computerized tomography (CT), positron emission tomography (PET) and single-photon emission computed tomography (SPECT) modalities [1]. MRI images provide better soft tissue definition and higher spatial resolution, but they are short of movement information such as body metabolism. CT images gain importance as a three dimensions (3D) imaging technique with the characteristic of short scan times and high imaging resolutions. Nevertheless, tissue characterization is limited and the restrictions of CT scan equipment are to invert slices of images to one image in the short scan times. Furthermore, the PET images have the property of high sensitivity due to the molecular imaging technique, but they are with lower resolution. SPECT images are used to study the blood flow of tissues and organs by the imaging

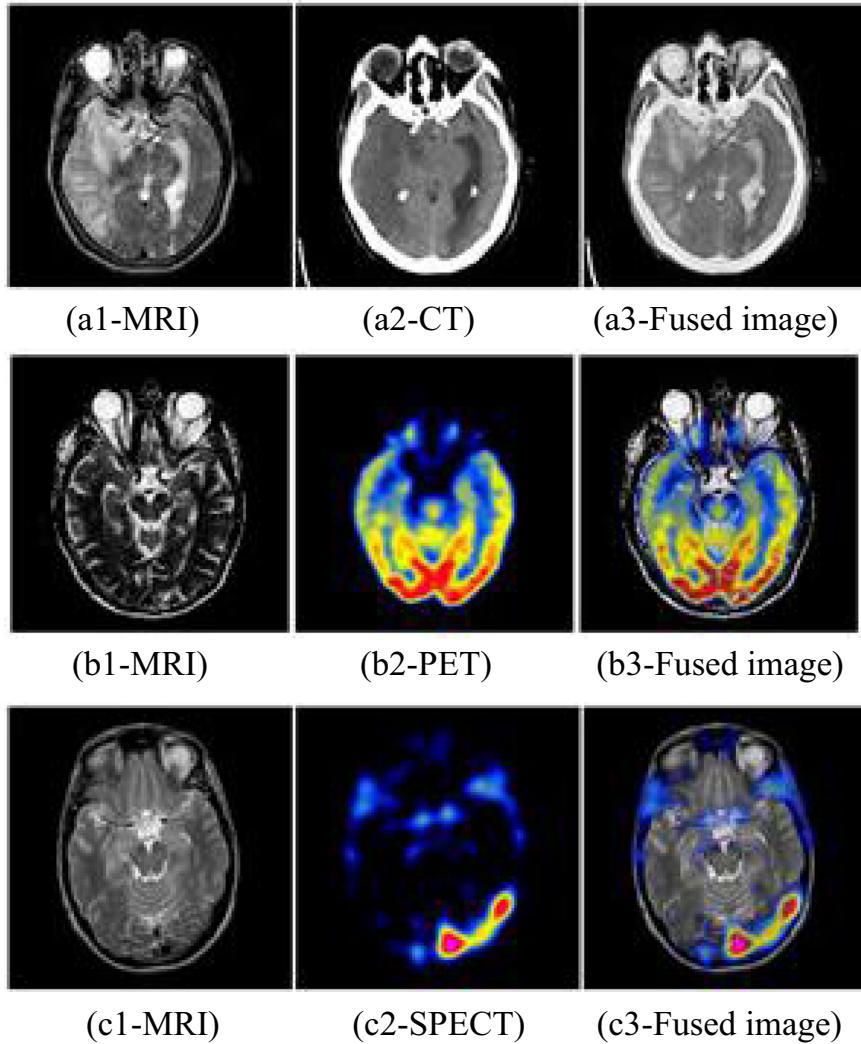
technique of nuclear. In summary, every modality of imaging has its own characteristics and practical limitations. This enforces to explore new imaging technologies or new fusion methods for combining information from multiple imaging modalities. The latter seems to be more meaningful because of lower cost and shorter time, compared with the former. The multi-modal medical image fusion traditionally centers on three categories: MRI-CT, MRI-PET and MRI-SPECT images fusion, as shown in Fig. 1.

In addition to the rapid development of the imaging techniques, there have been a large number of approaches in the field of medical image fusion. Traditionally, image fusion method consists of image decomposition and reconstruction, image fusion rules and image quality assessments, shown in Fig. 2. Firstly, input images are decomposed into a series of sub-images by image decomposition algorithms. Secondly, the image fusion rules are used to combine the multiple features into sub-images at different resolutions. Thirdly, the fused image is reconstructed from the fused sub-images at different resolutions by image reconstruction algorithms. Finally, the image quality assessment seeks to predict the quality of fused images.

Fusion of medical images with two dimensions (2D) has been moved to high dimension, such as 3D and four dimensions (4D) images fusion. Simulation with tools, such as MATLAB has transited to design a visual system for displaying the image fusion. And, researchers are eager to introduce the image fusion approaches such as multi-focus image fusion and multi-sensor image fusion into medical image analysis.

The rest of this paper is structured as follows: Section 2 gives a

\* Corresponding author.



**Fig. 1.** Examples of multi-modal medical image fusion (The fusion of MRI-CT, MRI-PET and MRI-SPECT images.).

brief introduction of the image decomposition and reconstruction of multi-modal medical image fusion algorithms. [Section 3](#) contains the overview of image fusion rules. The image quality assessments of the multi-modal medical image fusion methods are discussed in [Section 4](#). Finally, the conclusion is given in [Section 5](#).

## 2. Image decomposition and reconstruction

In terms of multi-modal medical image fusion, scheme for image decomposition and reconstruction closely relates to the quality extracted from the images. Characteristic of approaches in this frame aims at decomposing the original image into a sequence of images and then reconstructing the decomposition images into a single image.

It is of interest to look at a snapshot of five different key methodologies: (1) color space [[3,4](#)], (2) pyramid [[5,6](#)], (3) wavelet [[7–16](#)], (4) sparse representation [[17–21](#)] and (5) salient feature [[22–24](#)]. From [Table 1](#), the methods of image decomposition and reconstruction are compared with the indexes of spatial domain, frequency domain, multi-scale, scale invariance, dictionary and directive filter. Due to the images shown in pseudo-color, color space based fusion method is employed to multi-modal medical image fusion. Unlike other methods, color space methods are used to process input images in spatial domain. Then multi-scale

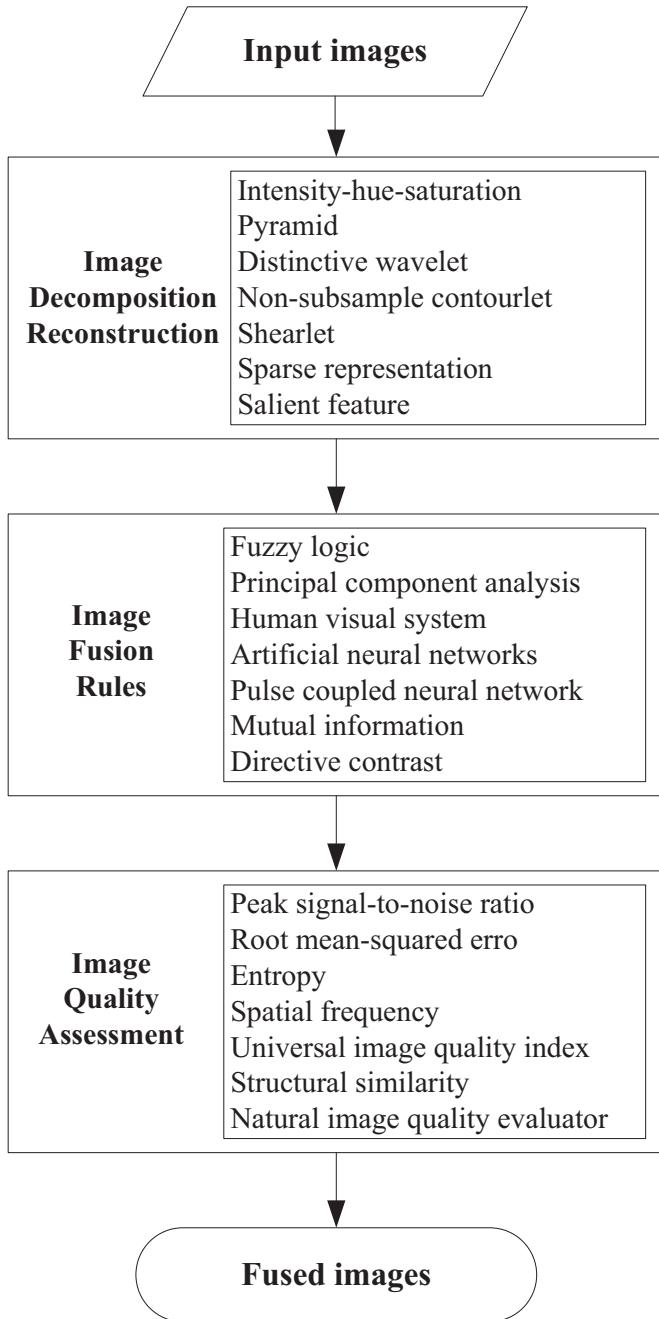
decomposition (MSD) is used to extract and combine salient features of medical images at different scales [[2](#)]. Methods based on MSD are pyramid, wavelet and salient feature. The advantages of wavelet methods are scale invariance and directive filter. Moreover, salient feature methods are used as the MSD tool in spatial domain. In addition, the sparse representation methods, inspired from the compressed sensing algorithms, construct a dictionary of input images.

### 2.1. Color space methods

An intensity-hue-saturation (IHS) fusion method [[3,4](#)] is expected to be a well quality images with a visually beautiful color images (shown in [Fig. 3](#)). In the IHS based medical image fusion methods, the IHS transformation converts the input image from the red, green and blue channels (RGB) to an IHS model by matrix (RGB-IHS). The fused image is reconstructed by the inversed IHS transformation (IHS-RGB).

- 1) Two medical images with difference in modalities are converted to the images in presentation of IHS color space by the RGB-IHS transformation

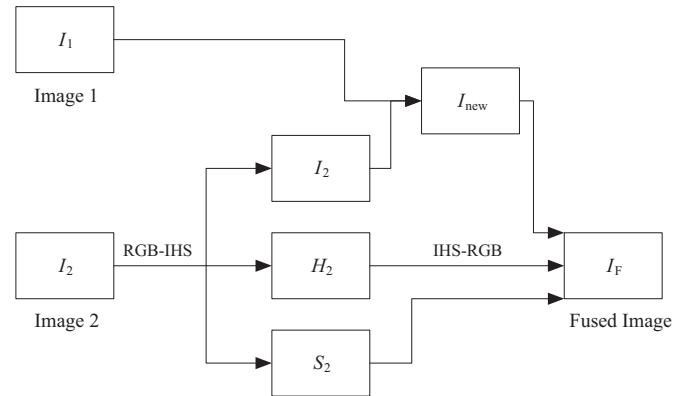
$$I = \frac{R + G + B}{3} \quad (1)$$



**Fig. 2.** The summary of three stages in the multi-modal medical image fusion.

**Table 1**  
Comparison of the five methods of image decomposition and reconstruction.

Method	Color space	Pyramid	Wavelet	Sparse representation	Salient feature
Spatial domain Frequency	Yes No	No Yes	No Yes	Yes Yes	Yes No
Multi-scale Scale invariance	No No	Yes No	Yes Yes	No No	Yes Yes
Dictionary Directive filter	No No	No No	No Yes	Yes No	No No



**Fig. 3.** Diagram of rules in IHS-based fusion of two medical images  $I_1$ ,  $I_2$ .

$$\begin{cases} H = \frac{G - B}{3I - 3B}, S = \frac{I - B}{I}, & \text{if } B < R, G \\ H = \frac{B - R}{3I - 3R}, S = \frac{I - R}{I}, & \text{if } R < B, G \\ H = \frac{R - G}{3I - 3G}, S = \frac{I - G}{I}, & \text{if } G < R, B \end{cases} \quad (2)$$

2) The final image is then obtained by the HIS-RGB transformation

$$\begin{cases} R = I(1 + 2S - 3S \times H), G = I(1 - S + 3S \times H), \\ B = I(1 - S), & \text{if } B < R, G \\ R = I(1 - S), G = I(1 + 5S - 3S \times H), \\ B = I(1 - 4S + 3S \times H), & \text{if } R < B, G \\ R = I(1 - 7S + 3S \times H), G = I(1 - S), \\ B = I(1 + 8S - 3S \times H) & \text{if } G < R, B \end{cases} \quad (3)$$

In this light, researchers propose various approaches for image fusion based on fast IHS fusion with an appropriate tradeoff parameter between the spatial and spectral resolution of the image to be fused.

## 2.2. Pyramid methods

The pyramid representation [5,6] has been used in fusing multi-modal medical images (shown in Fig. 4). It is well known as the diverse resolutions in the level and the iteration of the images.

The steps involved in pyramid methods are:

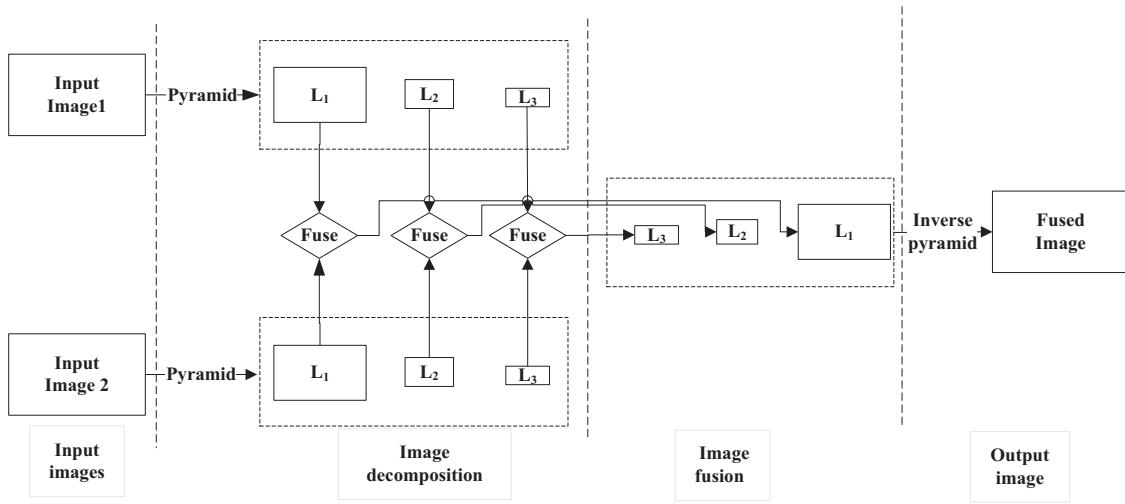
1) A sequential order  $\{p_1, p_2, \dots, p_i\}$  of input image  $I$  is constructed by the pyramid transformation

$$r_i = F_i * p_i, p_{i+1} = p_i - r_i, i = 1, \dots, \max\_level, p_1 = I \quad (4)$$

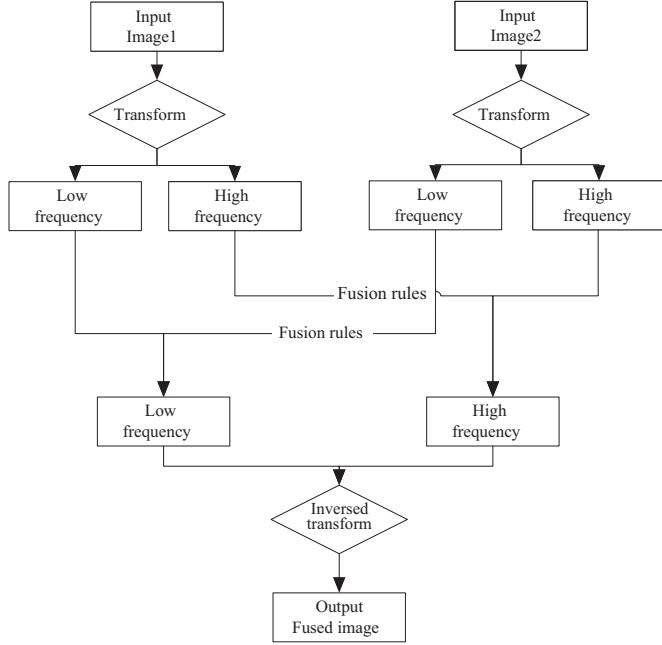
where  $r_i$  is the residual image of the input image at the  $i$ -th level,  $F_i$  is the filter for down-sampling and  $p_i$  is the pyramid representation for the input image at the  $i$ -th level and  $\max\_level$  is the length of levels.

2) The fused image  $I_F$  is reconstructed by the inverse pyramid transformation

$$I_F = p_{\max\_level+1} + \sum_{i=1}^{\max\_level} r_i \quad (5)$$



**Fig. 4.** Framework of three-levels pyramid based image fusion.



**Fig. 5.** Framework of wavelet based methods.

### 2.3. Wavelet transformation based methods

Wavelet transformation (WT) based fusion methods for multi-modal medical image fusion [7–16] is regarded as multi-scale geometric analysis tools (shown in Fig. 5). Firstly, the input image is decomposed into low and high frequency components. Secondly, different image fusion rules are selected to fuse different frequency components. Finally, the fused image is obtained by inverse transformation.

In 2001, discrete wavelet transformation (DWT) is applied for fusing multi-modal medical images. DWT [7–9] preserves different frequency information in a stable form and permits perfect localization both in time and spatial frequency domain. However, the method fails to meet the requirements of shift-invariance. Then redundant DWT (RDWT) [10] aims to overcome the shift-variance problem of DWT through removing down-sampling operation from traditional critically sample DWT. Since multi-wavelet filter banks have no strict division of low-pass and high-pass, multi-wavelet (MWT) [11] has an advantage of preserving more details and texture information with purpose to overcome the

shortcomings of the scalar wavelet. Lifting wavelet transformation (LWT) [12] is also famous as the second generation wavelet transformation with the properties of reducing the computational complexity of wavelet transformation. The construction of LWT is divided into three phases: split, prediction and update.

Methods based on wavelets are widely used as optimal tools for analyzing one-dimensional piecewise smooth signals, whereas, there are serious limitations exist when dealing with high dimensional signals such as images. The limitation on image processing with 2D separate wavelet is that it performs not well in detecting the smoothness along the edges and the only three directional sub-bands of high resolutions. Contourlet (CT) [13] performs better for representing for images than the wavelets with an advantage of owning contour segments to capture the geometrical structures of images. Two stages of the method are multi-scale decomposition by pyramid transformation and directional decomposition by various directional filter banks. Pyramid transformation and directional filter banks without sub-sampled constitute the non-subsampled contourlet transformation (NSCT) method [14].

Method based on Shearlet (ST) [15,16] has been proved to be the best sparse directional image representation. Due to the properties of no restrictions on the number of directions for shearing, ST [15] is used in fusion of medical images. Later, 3D ST theory is available [16] extending from 2D ST with information between image sequence. 3D ST consists of 3D pyramid filter and pseudo-spherical Fourier transformation with its application in the field of multi-modal medical image fusion [16].

In a review, WT, CT and ST based methods meet multi-scale information properties to improve the quality of fused image resulting from the pyramid transformation. Table 2 shows an example of a single scale of medical image processing by WT, CT and ST methods. The image by traditional WT is decomposed into one low frequency and three high frequency bands in horizontal, vertical and diagonal directions. And, the image by CT owns much more high frequency bands compared with the traditional WT. In ST, the decomposed image owns infinite high frequency bands because they are controlled by the parameter.

**Table 2**  
Comparison of wavelet based fusion methods.

Method	Wavelet	Contourlet	Shearlet
Low frequency band	1	1	1
High frequency band	3	finite	infinite

#### 2.4. Sparse representation methods

Different from the traditional MSD methods, sparse representation is proposed as a tool for fusing multi-modal medical images with the assumption that both the high and low frequency images share the same set of sparse coefficients [17,18]. The fusion method based on sparse representation is derived from the compressed sensing and developed to joint sparse representation (JSP) [19,20] and group sparse representation (GSP) [21].

The whole procedure of the JSP methods [19,20] (shown in Fig. 6) is summarized as follows:

- 1) Firstly, the input images  $I_1, I_2$  are transformed into vectors  $V$  via sliding window.
- 2) Secondly, the vectors of input images are sparsely represented by an over-complete dictionary

$$V = V^C + V^U, V^C = DS^C, V^U = DS^U \quad (6)$$

where  $V^C$  denotes the intersection of vectors in images  $I_1$  and  $I_2$ ,  $V^U$  denotes the difference of vectors in images  $I_1$  and  $I_2$ ,  $D$  denotes the over-complete dictionary, and  $S^C$  denotes the sparse coefficient matrix of  $V^C$  and  $S^U$  denotes the sparse coefficient matrix of  $V^U$ .

- 3) Thirdly, the spare coefficients are combined by using the fusion rules.

$$S_F = f(S^C, S^U) \quad (7)$$

where  $S_F$  denotes the fused sparse coefficient and  $f$  denotes the linear function.

- 4) Fourthly, the fused vectors are obtained by the over-complete dictionary.

$$V_F = DS_F \quad (8)$$

where  $V_F$  denotes the fused vector.

- 5) Finally, the fused image  $I_F$  is reconstructed by iterating the vectors  $V^C$  and  $V^U$ .

#### 2.5. Salient feature methods

Salient feature based fusion methods [22–24] are different from the other methods. The advantages of the fusion methods are shift-invariance, retained saliency features and low computational complexity. Edge-preserving filters have been a hot research topic of methods based on salient feature such as medical image fusion with multi-scale directional bilateral filter [22], medical image fusion with guided filter [23] and medical image fusion with local extrema scheme [24].

The common procedure of edge-preserving filters based fusion

methods (shown in Fig. 7) is listed as follows:

- 1) Firstly, the input images are decomposed into multi-scale representation by edge-preserving filters. The base layer of each input image at different scales is computed by Eq. (9)

$$B_i^1 = I_1 * F_i, B_i^2 = I_2 * F_i \quad (9)$$

where  $I_1, I_2$  depict the two input images,  $B_i^1, B_i^2$  depict the base layer of the input images  $I_1, I_2$  at  $i$ -th level and  $F_i$  depict the edge-preserving filters at  $i$ -th level. The detailed layer of each input image at different scales is computed by Eq. (10)

$$D_i^1 = I_1 - B_i^1, D_i^2 = I_2 - B_i^2 \quad (10)$$

where  $D_i^1, D_i^2$  depict the detailed layer of the input images  $I_1, I_2$  at  $i$ -th level.

- 2) Secondly, image fusion rules are applied to fuse the base and detailed layer of each input image at different scales

$$B_F = f_B(B_i^1, B_i^2), D_F = f_D(D_i^1, D_i^2) \quad (11)$$

where  $B_F, D_F$  are the fused base and detailed layer of each input image and  $f_B, f_D$  are the image fusion rules of base and detailed layer.

- 3) Finally, the fused image  $I_F$  is reconstructed by the addition with fused base layer  $B_F$  and fused detailed layer  $D_F$ .

$$I_F = B_F + D_F \quad (12)$$

### 3. Image fusion rules

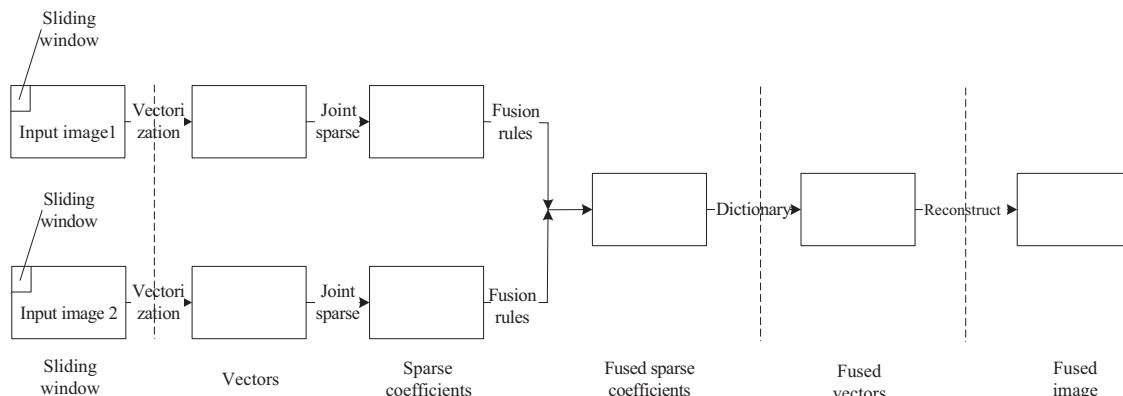
Image fusion rules refer to algorithms that seek to highlight the features of interest in images and restrain the features of insignificance. The main contribution of image fusion rules is the combination of multiple original images into a single image. Traditionally, an image fusion rule includes four components (shown in Fig. 8): activity-level measurement, coefficient grouping, coefficient combination and consistency verification [2].

- 1) Activity-level measurement

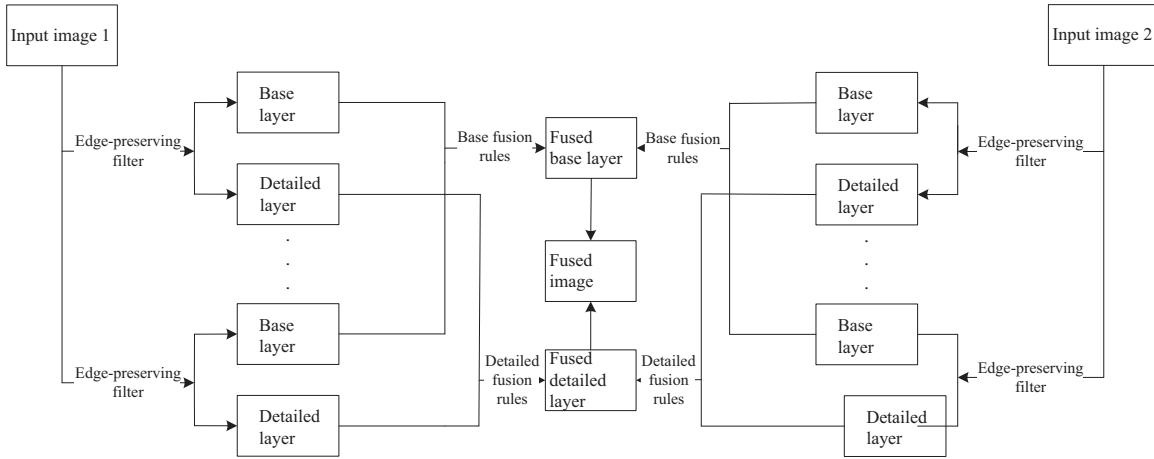
The activity-level scheme reflects the saliency of each coefficient at different scales and it can be categorized into coefficient-based activity (CBA), window-based activity (WBA) and region-based activity (RBA).

- 2) Coefficient grouping

The coefficient grouping scheme roughly contains no-grouping (NG), single-scale grouping (SG) and multi-scale grouping (MG). SG scheme means that coefficients between different sub-images at the same scale are fused by the same strategy.



**Fig. 6.** Framework of joint sparse representation based methods.



**Fig. 7.** Framework of edge-preserving filter based methods.

### 3) Coefficient combination

The categories of coefficient combination scheme include maximum rules (MR), average rules (AR) and weighted average rules (WAR). One common coefficient combination scheme is the MR

$$C_F = \begin{cases} C_i^1, & \text{if } C_i^1 > C_i^2 \\ C_i^2, & \text{if } C_i^1 < C_i^2 \end{cases} \quad (13)$$

where  $C_F$  is the combined coefficient,  $C_i^1, C_i^2$  are the coefficient of input images  $I_1, I_2$  at the  $i$ -th level. AR can be applied to use different original images, that is

$$C_F = \frac{1}{2}(C_i^1 + C_i^2) \quad (14)$$

WAR can be realized by weights of input images  $I_1, I_2$  at  $i$ -th level  $w_i^1, w_i^2$

$$C_F = \frac{1}{2}(w_i^1 \times C_i^1 + w_i^2 \times C_i^2) \quad (15)$$

### 4) Consistency verification

The consistency verification scheme ensures that the coefficients in the neighborhood are fused with the same rules.

An effective fusion rule plays a significant role in the objective quality assessments of the fused image. Conventionally, fusion rules [1] can be implemented at three levels: pixel level, feature level and decision level. In pixel level, the fused image is obtained with the corresponding pixel values of source images. Pixel level fusion rule deals with information associated with each pixel. Feature level

fusion rule deals with information associated with regions and features like saliency feature, visibility or texture, are applied for fusion. In decision level, the fused image is based on rules of statistics, fuzzy logic and machine learning. In the following, five classes of fusion rules for multi-modal medical image fusion have been introduced.

#### 3.1. Fuzzy logic

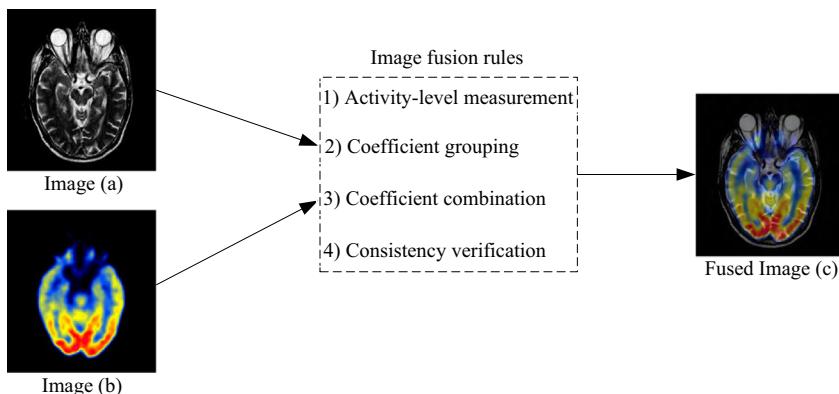
Fuzzy logic methods [25–27] belong to the image fusion rules implement in decision level. Image fusion rules based on fuzzy logic are used for coping with difficulties of blurry fused images. There are two models of fuzzy logic methods: Mamdani and T-S model. Compared with the Mamdani model, T-S model is more accurate avoiding the defuzzing.

Let  $C^1, C^2$  are image features extracted from input images by the combination with fuzzy logic to calculate weight for each pixel value. Firstly, the inputs are handled by fuzzy logic of IF-THEN rules as follows.

- R<sup>1</sup>: IF  $C^1$  is high and  $C^2$  is high THEN  $C_w$  is high.
- R<sup>2</sup>: IF  $C^1$  is low and  $C^2$  is high THEN  $C_w$  is medium.
- R<sup>3</sup>: IF  $C^1$  is high and  $C^2$  is low THEN  $C_w$  is medium.
- R<sup>4</sup>: IF  $C^1$  is low and  $C^2$  is low THEN  $C_w$  is low.

Secondly, the high, medium and low components are computed as follows.

$$C_{w,i}(y) = e^{-(y-\mu_i)/\sigma_i}, i = 1, 2, 3 \quad (16)$$



**Fig. 8.** Schematic diagram of fusion rules ((a), (b) are two input images, (c) is the fused image.).

where  $i$  with the value of 1, 2 and 3 are for high, medium and low, respectively,  $\mu$  is the mean value and  $\sigma$  is the variance value. Finally, the weight of fuzzy logic is obtained by processing fuzzy outputs with center average defuzzifier.

### 3.2. Statistics

Statistics based method is related to data-driven technique as well as higher order statistics to reveal hidden saliency structure. Examples of statics methods for multi-modal medical image fusion are principal component analysis (PCA) [28–30] and Hidden Markov Tree (HMT) [31,32].

PCA methods [25–27] are used as dimensionality reduction tools and refer to the linear combination of vectors forming new irrelevant principal components.

Let  $C^1, C^2$  are the two coefficients of the input images

$$C^1 = [x_1^1, x_2^1, \dots, x_N^1]^T, C^2 = [x_1^2, x_2^2, \dots, x_N^2]^T \quad (17)$$

where  $x_i^1, x_i^2$  are the column vectors of the coefficients  $C^1, C^2$ , the covariance matrix of two coefficients is given by

$$\text{Cov}(C^1, C^2) = E[(C^1 - \mu_1)(C^2 - \mu_2)] \quad (18)$$

where  $E$  is the expectation of vectors,  $\mu_1, \mu_2$  are the average of the coefficients  $C^1, C^2$

$$\mu_1 = \frac{1}{N} \sum_{i=1}^N x_i^1, \mu_2 = \frac{1}{N} \sum_{i=1}^N x_i^2 \quad (19)$$

Then the normalized components  $w_1, w_2$  are computed by the eigen values  $V$  of covariance matrix Cov

$$w_1 = \frac{V(1, 1)}{V(1, 1) + V(2, 1)}, w_2 = \frac{V(2, 1)}{V(1, 1) + V(2, 1)} \quad (20)$$

Finally, the fused coefficients  $C_F$  of two input images are

$$C_F = w_1 \times C^1 + w_2 \times C^2 \quad (21)$$

Unlike PCA methods, the two-state HMT method is adopted to model the coefficients. The HMT method depicts the intra-coefficients by a mixture of two Gaussian random distributions and the inter-coefficients by links between the hidden states of one parent and four child coefficients in a quad-tree model. For each coefficient  $C$ , the neighbor  $NC$  represents the adjacent coefficients of  $C$  in the same band, the parent  $PC$  represents the coefficient in the same spatial location in the immediately scale. Apart from the  $NC$  and  $PC$ , the coefficient is named as  $CC$  which is in the same spatial location and scale but in different directional frequency. The HMT method is associated with a state transition probability

$$P_{mn} = \begin{bmatrix} p_{mn}^{11}, p_{mn}^{12} \\ p_{mn}^{21}, p_{mn}^{22} \end{bmatrix} \quad (22)$$

where  $p_{mn}^{xy}$  is the transfer probability from the coefficient  $m$  in state  $x$  to the parent coefficient  $n$  in state  $y$ . Each coefficient is determined by probability density function

$$C^i = \sum_{m=0}^1 p_i(m) \times f(C^i | S_i = m) \quad (23)$$

where  $f(C^i, S_i = m)$  is the probability density function of coefficient  $C^i$  in the state  $m$  ( $m=0,1$ ). Then the fused coefficients are

$$C_F = \begin{cases} C^1, & \text{if } |C^1| \geq |C^2| \\ C^2, & \text{if } |C^1| < |C^2| \end{cases} \quad (24)$$

### 3.3. Human visual system

Human visual system (HVS) [33–39] based method simulates the process of images recognition and comprehension. Inspired by the HVS, the algorithms for detecting the corners, edges and saliency features are used as fusion rules in multi-modal medical image fusion, such as visibility [30], smallest univalue segment assimilating nucleus (SUSAN) [33], artificial neural networks (ANN) [34–37] and retina-inspired model (RIM) [4,38,39].

The visibility [33] demonstrates the sharpness of the image. The bigger the visibility is, the smaller blurriness is. The fused coefficients are the choice of max visibility scheme.

$$V(x, y) = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n \left( \frac{1}{\mu} \right)^{\alpha} \frac{|I(x, y) - \mu|}{\mu} \quad (25)$$

The visibility in a block is inspired by the HVS of an image. Mathematically, visibility of the image  $I$  is illustrated by Eq. (25). In this formula,  $m \times n$  is the dimension of image  $I$ ,  $\mu$  is the mean value of image  $I$  and  $\alpha$  is a constant ranging from 0.6 to 0.7.

The SUSAN [33] is a function for extracting feature from images inspired by HVS. SUSAN associates each pixel with its local area of similar intensity. The small local area contains the most significant information on the structure of the image in one block around the center pixel. The SUSAN is computed as

$$d(\vec{r}, \vec{r}_o) = \exp \left( \frac{-(\vec{r} - \vec{r}_o)}{2\eta^2} \right) \exp \left( - \left( \frac{F(\vec{r}) - F(\vec{r}_o)}{T} \right)^6 \right) \quad (26)$$

where  $T$  is the threshold of brightness difference,  $\eta$  is a factor of distance scaling.  $F(\vec{r})$  is the input image  $F$  processing by a circular mask with the radius of  $|\vec{r}|$  and  $F(\vec{r}_o)$  is the neighborhood pixel near the central pixel with a radius of  $|\vec{r}_o|$ .

The ANN model [34–37] is able to learn from inputs for processing features. Examples of ANN models are mapping neural network (MNN) [34] and pulse coupled neural network (PCNN) [35–37]. MNN model is inspired from the self-organizing neural network and provides multi-level fusion strategies. In addition, PCNN model has been widely used as the image fusion rules for low and high frequency bands decomposed by WT methods. PCNN model [35–37] is developed from ANN based on the biological experimental observations of synchronous pulse in the visual cortex of the cat. Each neural cell is composed of accept domain, modulation domain and pulse generator. The intensity of each pixel in the image corresponds to the neural cell in PCNN model. Firstly, two channels are the input and the coefficients of the output. Secondly, the inputs from two channels are modulated by weighted summation with the purpose to obtain an intermediate state. Finally, the pulse generates according to the threshold.

In addition, the RIM model [4,38] is used as the fusion rules for components from IHS decomposition methods. The RIM is composed of five basic layers. The first layer is the representation of an array of high resolution cone photoreceptors. The second and the third layers are a spatial feature extractor and horizontal cells. The fourth and the fifth layers are the bipolar and ganglion cells. In conclusion, the RIM based image fusion rule is explained by

$$C_F = f_1 \times C^1 + f_2 \times C^2 \quad (27)$$

where  $C^1, C^2$  are the intensity components of two input images,  $f_1, f_2$  are the filters of feature extractors. The principle of filter  $f_1$  is a high-scale spatial feature extractor which is the difference in spatial positions between high-resolution and low-resolution images. Furthermore, the filter  $f_2$  is the combination of outputs of horizontal cells. Later, the RIM model is developed to multi-scale RIM [39].

### 3.4. Objective evaluation metrics

The measured value of the objective evaluation metrics is used as the feedback to an image fusion rule such that the image quality of the fused image is potentially be improved. In multi-modal medical image fusion, there are several instances where the metrics are derived from the definition of the objective evaluation metrics such as spatial frequency (*SF*) [3], the ratio of spatial frequency error (*rSFe*) [40], wavelet entropy (*WE*) [41], directive contrast (*DC*) [42], signal noise ratio (*SNR*) [43] and mutual information (*MI*) [44].

*SF* metric [3] is used as the fusion rule for multi-modal medical image fusion in IHS integrated with PCA approach. *SF* [3] introduced by Eskicioglu and Fisher is used to measure the activity of images. The expression of *SF* is defined as

$$SF = \sqrt{(RF)^2 + (CF)^2} \quad (28)$$

where *RF* is the row frequency and *CF* is the column frequency.

Similar to the definition of *SF*, *rSFe* [40] along with the fused image (*SF<sub>F</sub>*) compared with *SF* of the input images (*SF<sub>1</sub>*, *SF<sub>2</sub>*). Firstly, the *SF* is extended to the combination of *RF*, *CF*, main diagonal *SF* (*MDF*) and secondary diagonal *SF* (*SDF*)

$$SF = \sqrt{(RF)^2 + (CF)^2 + (MDF)^2 + (SDF)^2} \quad (29)$$

Then the *SF<sub>F</sub>* and *SF<sub>1</sub>*, *SF<sub>2</sub>* are computed by Eq. (28). Finally, *rSFe* is defined as follows:

$$rSFe = \frac{1}{2} \left[ (SF_F - SF_1)/SF_1 + (SF_F - SF_2)/SF_2 \right] \quad (30)$$

*WE* [41] is calculated by multi-scale entropy and based on the Shannon entropy to measure the information of signal distribution. The *WE* is calculated by

$$WE = - \sum_{j<0} p_j \cdot \ln p_j \quad (31)$$

where *j* denotes the resolution level, *p<sub>j</sub>* is the density distribution computed by energy of the detail signal *E<sub>j</sub>* and total energy *E<sub>tot</sub>*

$$p_j = E_j/E_{tot} \quad (32)$$

*DC* is a measurement of the difference between one pixel and its neighbor pixels. The visual system is highly sensitive to the intensity contrast *DC* [42] which is the ration of high frequency intensity *I<sub>H</sub>* and low frequency intensity *I<sub>L</sub>*.

$$DC = I_H(x, y)/I_L(x, y) \quad (33)$$

To solve the problems of distortion in the images, the image fusion rules based on *SNR* is proposed [43]. The fusion method is defined as

$$A(k) = \frac{1}{M_k} \sum_{1 \leq l \leq M_k} p_l \quad (34)$$

where *A(k)* is the activity level of each region, *M<sub>k</sub>* is the number of pixels in the region *k* of the input image and *p<sub>l</sub>* is the probability of pixel activity computed by Eq. (35)

$$p_l = \frac{w}{n} \sum_{1 \leq i \leq n} \frac{1}{3 \cdot 2^{2(n-i)}} \sum_{j_1} \sum_{j_2} |d_i(j_1, j_2)| \quad (35)$$

with *w* is the weight from *SNR* of the image, *n* is the number of decomposition levels and *d<sub>i</sub>(j<sub>1</sub>, j<sub>2</sub>)* is the detailed wavelet coefficients.

*MI* [44] is used as the fusion rule for multi-modal medical image fusion based on WT. The fused coefficients *C<sub>F</sub>* have been calculated by the maximizing the *MI* between the input images

*I<sub>1</sub>*, *I<sub>2</sub>* and then the fused image *I<sub>F</sub>* is obtained by Eq. (36)

$$C_F = \frac{1}{2} [MI(I_1, I_F) + MI(I_2, I_F)] \quad (36)$$

### 4. Image quality assessment

Image quality assessment plays a significant role in multi-modal medical image fusion. The image quality assessment research aims at designing algorithms with achievement of automatically assessing the quality of images in a perceptually consistent manner [45]. Additionally, the image quality assessment is not specially designed to assess the performance of medical image fusion methods. For example, structural similarity (*SSIM*), root-mean-square error (*RMSE*), *MI* and *SF* are used as image quality assessment tools in both medical image fusion and the other image processing algorithms. *SSIM* metric is used as tools in evaluating the quality of images in video [46–49] and image compression [50]. *MI* and *SF* metrics are used as tools in assessing the quality of images in multi-focus image fusion [51,52] and multi-sensor image fusion [53]. Moreover, *RMSE* metric is used as tools in predicting the quality of images in image denoising [54].

So far, subjective trial methodology and objective metric are provided for multi-modal medical image fusion assessments [55]. The subjective image fusion evaluation trials have been the most reliable way for image fusion methods in which the potential users are employed to evaluate the quality of the fused image. However, subjective tests are impractical in many applications due to the difficulty of tremendous organizations and strict tests. Objective fusion metrics are a far easier computed without the complex organization of observers. The objective metrics produce a single numerical score that indicates the distortion signals of the fusion algorithms with the consideration of the input images and the fused image.

The assessment of the image requires that a subjective method is consistent with an objective method [45,46]. Recently, there exist several quality assessment algorithms that make comparisons between these two methods. Any new objective quality measure for image fusion is remarkable with requirements of correlating well with subjective criteria as well as with other existing performance measures.

#### 4.1. Subjective quality assessment

Subjective quality metric is easy to obtain within a nonlinear questionnaire since it is convenient for people to compensate survey. In addition, the survey is close with dependent of the subjective validation application scope and methodology [45].

The traditional way of subjective quality assessment [45] is to calculate scores for each subject with normalized by the mean and variance of scores for that subject from a number of people ranging from 1 to 100. Firstly, five indexes marked with adjectives of bad, poor, fair, good and excellent are with the value assignments of 1, 2, 3, 4 and 5. Secondly, human subjects are asked to assign each image with a score for their perception of quality indicating their assessment of that image, defined as the extent to which the artifacts were visible and annoying. The raw scores are altered to scores between the test and the reference images and then converted to Z-scores with 1–100 ranges. Finally a difference mean opinion score for each test image is computed [45]. Finally, a double stimulus methodology for more accurate measurement of quality for realignment purposes is proposed [56]. The double stimulus method is a study of combination the first view with the second view with representations of reference and corresponding test images.

In addition, the collection of the subjective scores consists of two sets: symmetric and asymmetric comparisons with the purpose of

revealing the psychological responses of observers [47]. The total subjective scores of two sets are processing by eliminating the data noise. Then, we exclude the unqualified observers from the processed data to improve the confidence of the collected data.

#### 4.2. Objective quality assessment

Quality metric for image fusion is a pertinent quality assessment tool to evaluate the visual quality degradation of images suffering from various distortions during fusion procedure. Image signals are generally in instability while image quality is usually desired to evaluate an image with a single mathematical value. The purpose of quality metric is that multiple comparisons are obtained with an optimal setting of parameters for a specific fusion algorithm.

Objective image quality measure is foremost in various image processing applications. There are basically three classes known as full-reference, reduced-reference and no-reference of objective quality or distortion assessment approaches according to the availability of an original image with which the test image is to be compared [57]. In full-reference image metric [48], the reference image is assumed to be known. And in reduced-reference quality assessment, the reference image is only partially available to help evaluate the quality of test image. No-reference image metric means that the reference image is not available [57].

The most widely used objective quality assessments for multi-modal medical image fusion are full-reference quality metrics. The simplest full-reference metrics are based on signal distortion and HVM. The metrics based on signal distortion defined with strict mathematical theory [15,58–61] is incorporated with entropy ( $EN$ ), difference of entropy ( $DEN$ ), overall cross entropy ( $OCE$ ), standard deviation ( $STD$ ), sharpness ( $SP$ ),  $RMSE$ , peak signal to noise ratio ( $PSNR$ ), etc. The second class [39,57,61–67] based on HVS is to measure the information of salient feature transferred from the input images to the fused image, such as  $SSIM$ , the phase congruency based index ( $Q_G$ ), the gradient based index ( $Q_{AB/F}$ ), etc. It is meaningful to compute the error between test and reference images with loss in salient information of gradient, contrast and edge between different image components.

##### 4.2.1. Metrics based on signal distortion

The metrics based on arithmetic theory [15,45] for the evaluation of multi-modal medical image fusion algorithms are  $RMSE$ ,  $STD$  and  $SP$ .

$$RMSE = \left( \sum_{x=1}^M \sum_{y=1}^N [I_I(x, y) - I_F(x, y)]^2 \right)^{1/2} \quad (37)$$

$RMSE$  is the variance of the arithmetic square root.  $RMSE$  is defined by the input images and the fused image with size of  $M \times N$  by Eq. (37). It is the difference of each pixel value between the input image  $I_I$  and the fusion image  $I_F$ .  $STD$  is the square of  $RMSE$ .

$$STD = (RMSE)^2 \quad (38)$$

$SP$  reflects the details of the fused images

$$SP = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N \sqrt{\frac{1}{2} \{ [I_F(x, y) - I_F(x, y-1)]^2 + [I_F(x, y) - I_F(x-1, y)]^2 \}} \quad (39)$$

Moreover, some other metrics based on signal distortion are inspired from the information theory, such as  $PSNR$ ,  $EN$ ,  $DEN$ ,  $OCE$  [15], visual information fidelity (VIF) [45],  $MI$  [58] and  $SF$  [24].

$PSNR$  is one of the most widely used as an objective quality measurement and computed by Eq. (40)

$$PSNR = 10 \cdot \lg [(M \times N)^2 / RMSE] \quad (40)$$

$EN$  [15] indicates the information in the fused image

$$EN = - \sum_{x=0}^{255} p_x \ln p_x \quad (41)$$

where  $p_i$  is the probability distribution of the pixels with the value equals to  $i$  over the total number of the pixels.  $DEN$  denotes the difference of entropy between the input images  $EN_I$  and the fused image  $EN_F$

$$DEN = |EN_F - EN_I| \quad (42)$$

$OCE$  [15] reflects the entropy of two input images  $I_1$ ,  $I_2$  and the fused image  $I_F$

$$OCE = \frac{1}{2} [ (CE(I_1, I_F) + CE(I_2, I_F)) ] \quad (43)$$

where  $CE$  is the cross entropy of the images

$$CE = \sum_{x=0}^{255} p_i^{I_1} \ln \left| p_i^{I_1} / p_i^{I_2} \right|, I_1 = I_1, I_2 \quad (44)$$

$VIF$  [45] defined as the image quality assessment is raised with consideration of the relationship between image information and visual quality during the distortion process, since distortions with changes of lighting and brightness are nonstructural and that these should be treated differently except for the structural ones. The framework of VIF (shown in Fig. 9) is the connection of natural image source, distortion channel and human visual system models.

A Gaussian model as a probability density function is applied to construct natural source models based on models of natural scene statistics. Then a formula of signal attenuation and additive noise models are used to signify the signal distortion.

$$VIF = I(C; F) / I(C; E) \quad (45)$$

The reference images  $E$  and test images  $F$  are outputs of input image  $N$  together with natural source model  $C$  and HVS  $D$ , and  $I$  is the function for computing mutual information entropy of an image.

$MI$  [58] mainly concentrates on estimating the amount of information transferred from the input images into the fused image  $MI(I_I, I_F) = H(I_I) + H(I_F) + H(I_I, I_F)$ ,  $I_I = I_1, I_2$  (46)

where  $H(I_I, I_F)$  is the joint entropy between the input image  $I_I$  and the fused image  $I_F$  and  $H(I_I)$ ,  $H(I_F)$  are the marginal entropy of  $I_I$ ,  $I_F$ . Then, Hossny [59] modified the  $MI$  as the normalized mutual information ( $NMI$ )

$$NMI = 2 \left[ \frac{MI(I_I, I_F)}{H(I_I) + H(I_F)} + \frac{MI(I_2, I_F)}{H(I_2) + H(I_F)} \right] \quad (47)$$

$SF$  [24] is used to measure the overall clarity level of the fused images. It is a regular tool for assessing the active level of the image and can be obtained by Eq. (28), described in Section 3.3. Moreover, the  $RF$  and  $CF$  are defined as:

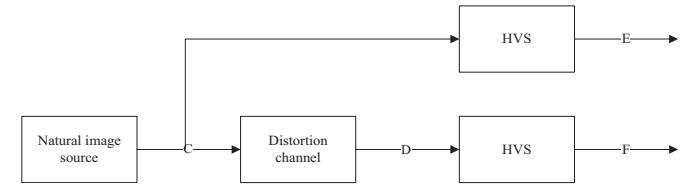


Fig. 9. Framework of visual information fidelity.

$$RF = \sqrt{1/M \times N \sum_{x=1}^M \sum_{y=2}^N [I_F(x, y) - I_F(x, y-1)]^2} \quad (48)$$

$$CF = \sqrt{1/M \times N \sum_{y=2}^N \sum_{x=1}^M [f(x, y) - f(x-1, y)]^2} \quad (49)$$

Later, the universal image quality index (*UIQI*) [60–62] inspired by HVS is designed based on the structural information of fused images by a combination of loss of correlation, luminance distortion and contrast distortion. The proposed index is suitable for achieving better quality predictions between the test and reference images. The new index is mathematically defined as

$$UIQI = \frac{\sigma_{I_1I_F}}{\sigma_{I_1} \times \sigma_{I_F}} \cdot \frac{\mu_{I_1I_F}}{(\mu_{I_1})^2 + (\mu_{I_F})^2} \cdot \frac{2\sigma_{I_1} \times \sigma_{I_F}}{(\sigma_{I_1})^2 + (\sigma_{I_F})^2} \quad (50)$$

where  $\sigma$  is the variance, and  $\mu$  is the average. Cvejic [61] modified the *UIQI* to a no-reference image quality assessment  $Q_b$

$$Q_b = \mu(I_1, I_2, I_F) UIQI(I_1, I_F) + [1 - \mu(I_1, I_2, I_F)] UIQI(I_2, I_F) \quad (51)$$

with the measurement of the important information from the input images  $\mu(I_1, I_2, I_F)$  is calculated to measure how well the preserved important information from the input images

$$\mu(I_1, I_2, I_F) = \begin{cases} 0, & \text{if } \sigma_{I_1I_F}/(\sigma_{I_1I_F} + \sigma_{I_2I_F}) < 0 \\ \sigma_{I_1I_F}/(\sigma_{I_1I_F} + \sigma_{I_2I_F}), & \text{if } 0 \leq \sigma_{I_1I_F}/(\sigma_{I_1I_F} + \sigma_{I_2I_F}) \leq 1 \\ 1, & \text{if } \sigma_{I_1I_F}/(\sigma_{I_1I_F} + \sigma_{I_2I_F}) > 1 \end{cases} \quad (52)$$

#### 4.2.2. Metrics based on salient feature

Another type of objective image quality assessments is implemented by measuring how well the salient features are transferred from the input images to the fused image. The ability to predict the fused image enables several changes for full-reference, reduced-reference and no-reference images.

*SSIM* [57] is presented for full-reference image quality assessment with the assumption that human visual perception is highly adapted for extracting structural information. *SSIM* measures how well the structural information of input images is preserved

$$SSIM(I_1, I_F) = [l(I_1, I_F)]^a \cdot [c(I_1, I_F)]^b \cdot [s(I_1, I_F)]^c \quad (53)$$

The task of *SSIM* for the input image  $I_1$  and the fused image  $I_F$  is separated into three components: luminance  $l$ , contrast  $c$  and structure  $s$ . Moreover,  $a, b$  and  $c$  are the same value of 1/3.

$Q_G$  [62] is defined based on the salient feature of images and provides an absolute measurement of image features.  $P$  is used to compute the preserved salient features from the input images and defined by the three correlation coefficients as:

$$Q_G = (P_p)^\alpha (P_M)^\beta (P_m)^\gamma \quad (54)$$

where  $P_p, P_M, P_m$  denote the phase congruency value, maximum moment and minimum moment, respectively and  $\alpha, \beta, \gamma$  are the exponential parameters.

$Q_{AB/F}$  [63–66] measures the success of edge information transferred from the input images to the fused image. It is computed as follows:

$$Q_{AB/F} = \frac{\sum_{x=1}^M \sum_{y=1}^N (Q_{I_1I_F}(x, y) w_{I_1}(x, y) + Q_{I_2I_F}(X, Y) w_{I_2}(x, y))}{\sum_{x=1}^M \sum_{y=1}^N (w_{I_1}(x, y) + w_{I_2}(x, y))} \quad (55)$$

where  $Q_{I_1I_F}(x, y), Q_{I_2I_F}(X, Y)$  are the preserved values of edge and orientation information at location  $(x, y)$  and  $w_{I_1}(x, y), w_{I_2}(x, y)$  reflect the importance of  $Q_{I_1I_F}(x, y), Q_{I_2I_F}(X, Y)$ .

The metric  $Q_{CF}$  [39] that embraces both contrast enhancement and image fusion to measure the performance of multi-modal medical image fusion algorithms, is defined as follows:

$$Q_{CF} = \frac{1}{M} \sum_{(x,y) \in \bar{\Omega}} S(x, y) \cdot O(x, y) \quad (56)$$

where  $\bar{\Omega}$  is the non-flat regions of the images,  $M$  is the total number of pixels in  $\bar{\Omega}$ ,  $S(x, y)$  is the edge strength increment at position  $(x, y)$  and  $O(x, y)$  is the edge coincidence at position  $(x, y)$ .  $Q_{CF}$  owns a larger value only if both the contrast enhancement and preserved salient features are larger over all non-flat regions.

Recently, natural image quality evaluator (*NIQE*) [67] is used as the no-reference objective quality assessment tool in medical image fusion [24]. *NIQE*, one of the blind image quality analysis models, measures the image quality without knowledge of distortions or human opinions in advance.

$$NIQE = \sqrt{(\mu_1 - \mu_2)^T \cdot \left( \frac{\sigma_1 + \sigma_2}{2} \right)^{-1} \cdot (\mu_1 - \mu_2)} \quad (57)$$

where  $\mu_1, \mu_2$  are the averaged vectors of the natural multivariate Gaussian model (MVG) and MVG of the distorted image and  $\sigma_1, \sigma_2$  are the covariance matrices of the natural MVG and MVG of the distorted image.

## 5. Experiments on database

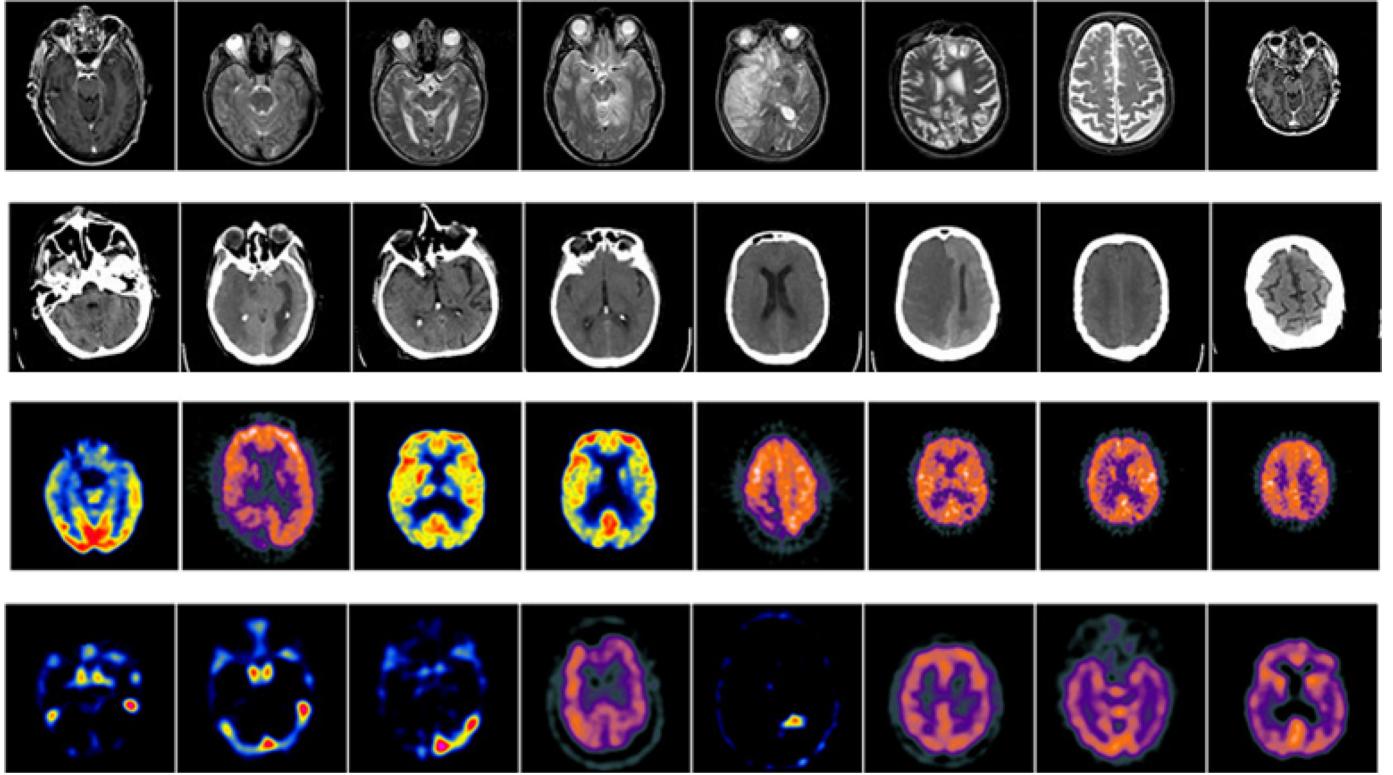
The Whole Brain Atlas [68] is a benchmark database for evaluating the performance of multi-modal medical image fusion methods established by Keith A. Johnson and J. Alex Becker in Harvard Medical School. The Whole Brain Atlas database consists of four imaging types: CT, MRI, PET and SPECT with the description of normal and abnormal brain structure (Fig. 10). And more importantly, all the images of the database are co-aligned.

In this section, three groups of brain images in the Whole Brain Atlas database [67] are used as the empirical data. Each group owns a pair of medical images from different imaging devices. The corresponding pixels of the two input images have been already registered. In this section, the methods for comparison are six image fusion algorithms based on IHS+PCA method [3], Morphology Pyramid (MoR)+MR method [5], DWT+Entropy [7], NSCT+PCNN [37], Joint sparse representation (JSP)+AR [20] and LES+DC [24] fusion methods, respectively. To evaluate the performance of different multi-modal medical image fusion methods described in Section 2 and Section 3, eight metrics are adopted as the objective quality assessments, such as *SSIM* in Eq. (53) [56], *RMSE* in Eq. (37) [15], *MI* in Eq. (46) [52], *PSNR* in Eq. (40) [15], *SF* in Eq. (28) [24],  $Q_{AB/F}$  in Eq. (55) [58–60],  $Q_G$  in Eq. (54) [57] and  $Q_{CF}$  in Eq. (51) [39].

### 5.1. Evaluations on the group of MRI-CT fusion

In this subsection, the empirical images are three examples of MRI-CT fusion are for the disease of Fatal stroke, Multiple embolic infarctions and Acute stroke. Each example of disease includes 23 pairs of MRI and CT imaging modalities.

Figs. 11–13 illustrate the input images and the fused image by the methods, described above. Three pairs of MRI and CT images are the input images (images marked with slice 11 in the disease of Fatal stroke, marked with slice 11 in the disease of Multiple embolic infarctions and marked with slice 12 in the disease of Acute stroke). The structure of brain is very clear in the fused image (shown in Fig. 11(d)). In Fig. 11(c)–(d), the fusion methods can well preserve the complementary information of the CT image.

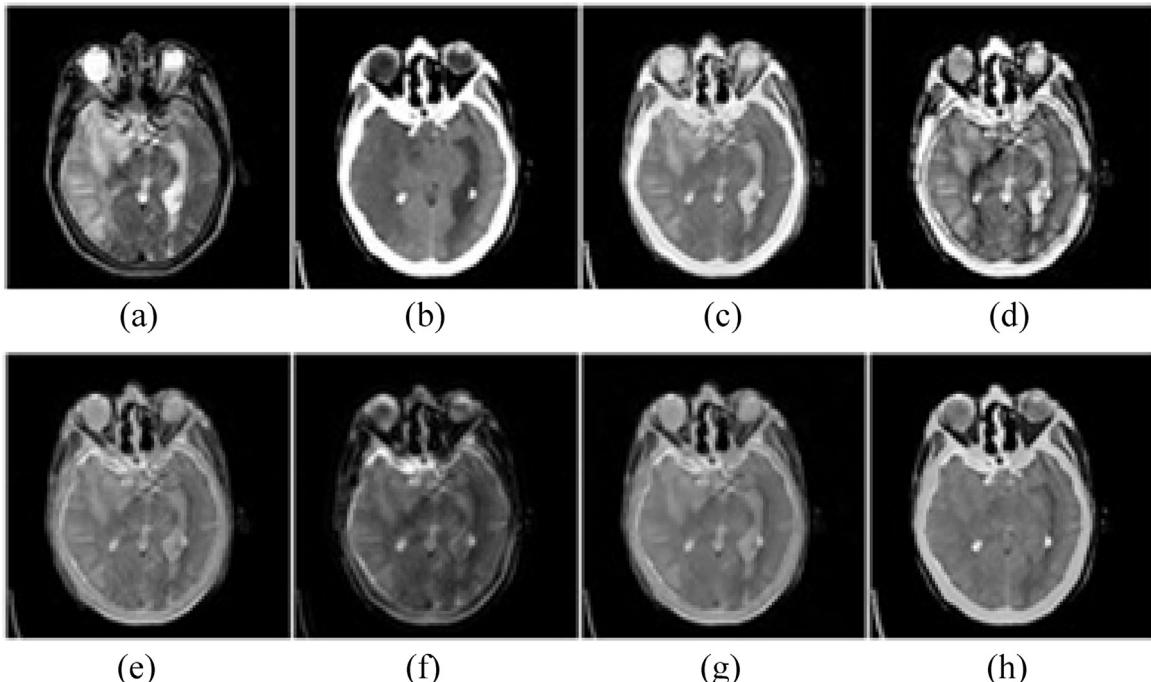


**Fig. 10.** Multi-modal image fusion database composed by four imaging types: MRI, CT, PET and SPECT (The first row shows the slices of brain in MRI modality. The second row shows the slices of brain in CT modality. The third row shows the slices of brain in PET modality. The fourth row shows the slices of brain in SPECT modality).

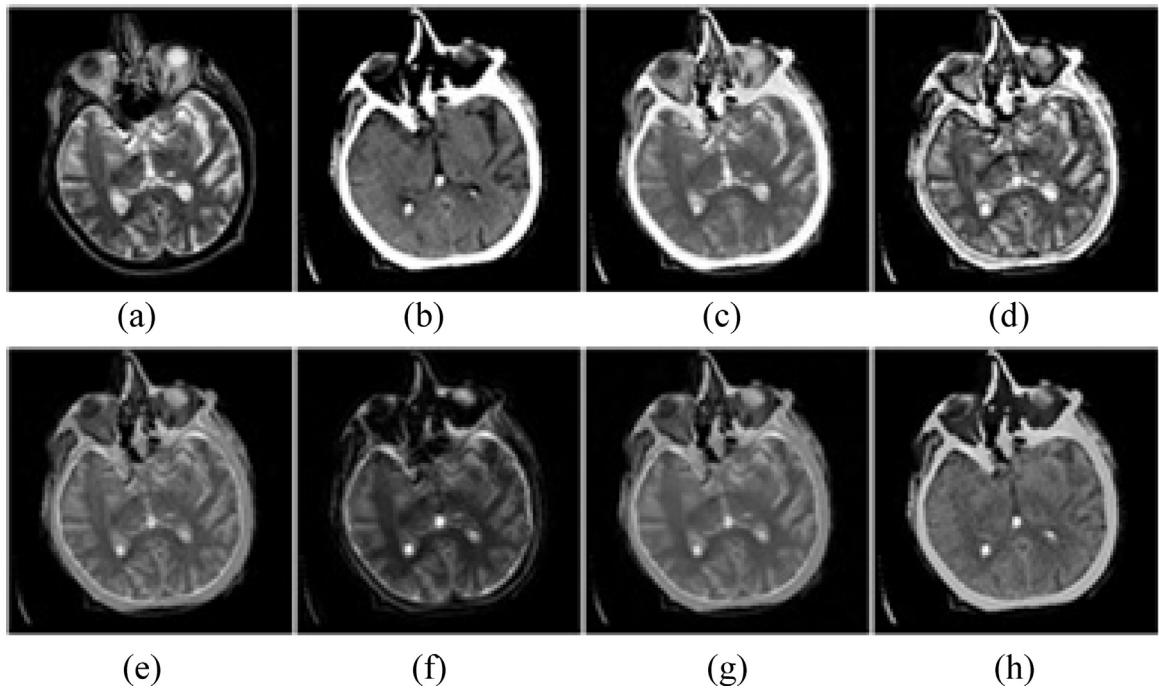
[Fig. 12\(e\)](#) shows that the fused image can reserve the details without producing artifacts and distortions. In [Fig. 13\(c\)](#) and [\(d\)](#), the structure is completely reserved from the input images. However, the texture information is disappeared in the fused images (shown in Figure. [\(e\)–\(h\)](#)).

Then, the objective performances of different methods are shown in [Table 3](#). It can be seen that, for the images in the group of

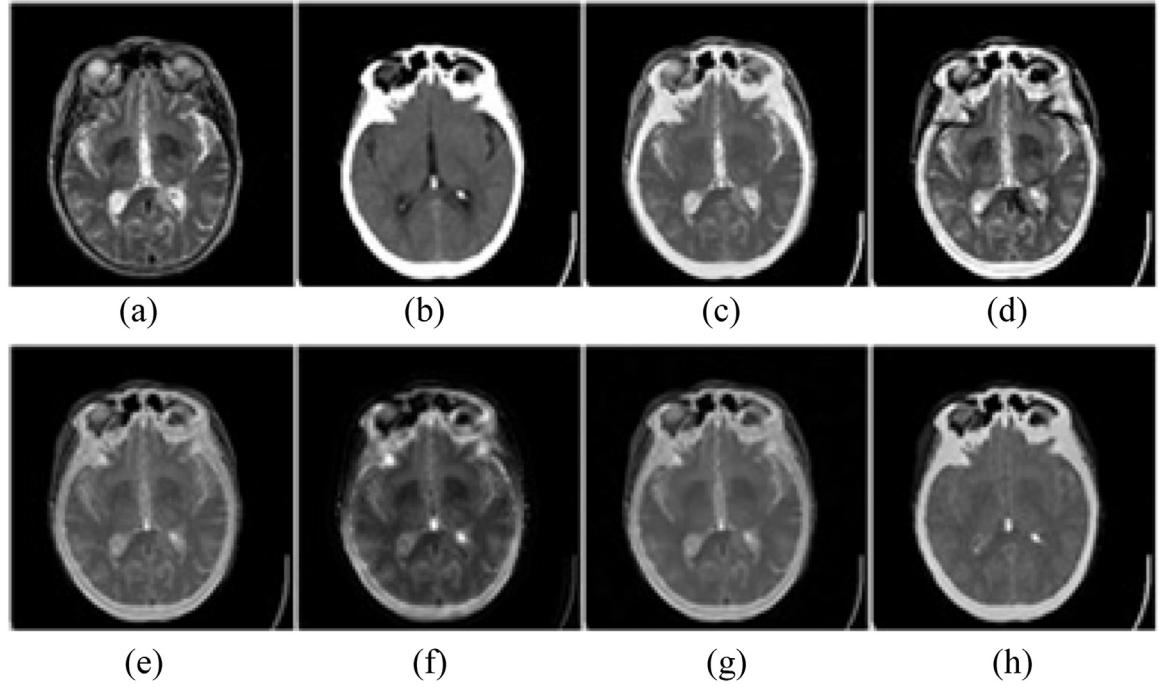
MRI-CT fusion, the IHS+PCA based method gives the largest quality indexes for  $MI$ ,  $Q_G$  and  $Q_{CF}$ . Moreover, the  $SF$  and  $Q_{AB/F}$  values of the MoR+MR based method are the largest for the fused images. The  $RMSE$  value of the NSCT+PCNN based method is the largest. The LES+DC based method performs best in terms of  $SSIM$  and  $PSNR$  metrics.



**Fig. 11.** Performance comparison of different methods of group1. ((a) MRI, (b) CT, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).



**Fig. 12.** Performance comparison of different methods of group1. ((a) MRI, (b) CT, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).



**Fig. 13.** Performance comparison of different methods of group1. ((a) MRI, (b) CT, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).

### 5.2. Evaluations on the group of MRI-PET fusion

In this subsection, three examples of MRI-PET fusion are for the disease of Glioma, Mild Alzheimer and Glioma-GD. Each example of disease includes 81 pairs, 23 pairs and 17 pairs of MRI and PET imaging modalities.

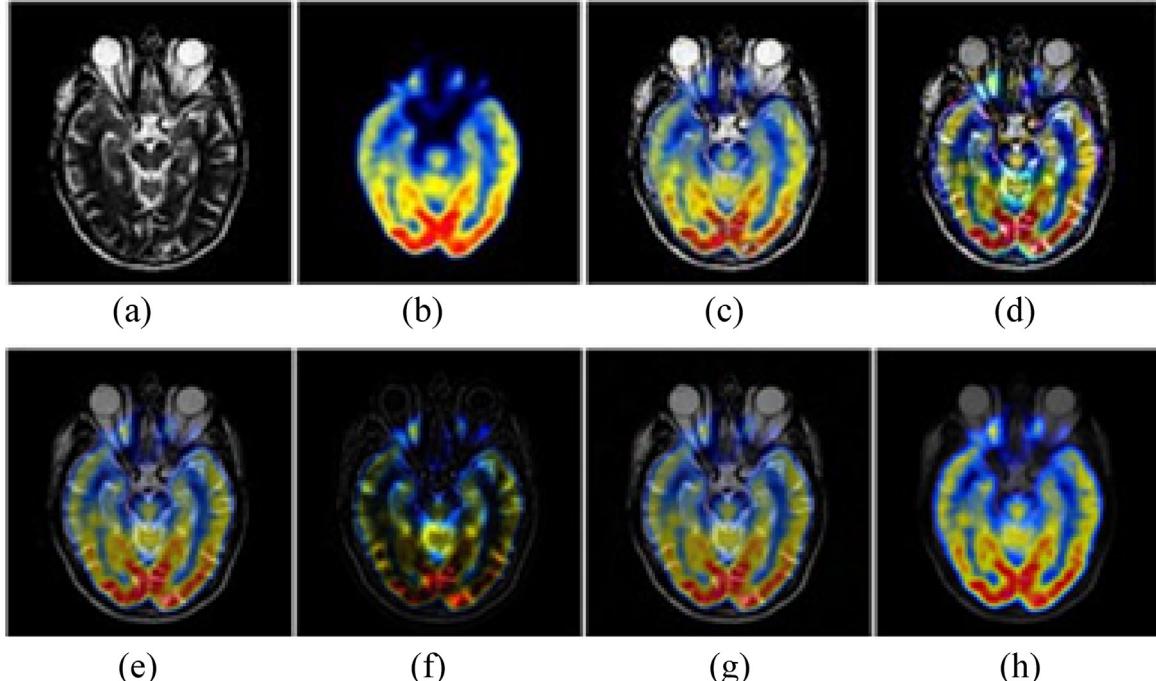
Figs. 14–16 illustrate the input images and the fused image by

the methods, described above. Three pairs of MRI and PET images are the input images (images marked with slice 9 in the disease of Glioma, marked with slice 15 in the disease of Mild Alzheimer and marked with slice 55 in the disease of Glioma-GD). In Fig. 14(c), (e), (g)–(h), the color is perfectly preserved in the fused images. However, the NSCT+PCNN based method introduces the color distortion to the fused images (shown in Fig. 14(f), Fig. 15(f) and

**Table 3**

Objective quality assessment of different fusion methods in group of MRI-CT fusion.

Method	IHS + PCA	MoR + MR	DWT + Entropy	NSCT + PCNN	JSP + AR	LES + DC
SSIM	0.6956	0.62743	0.6453	0.4365	0.1736	<b>0.7941</b>
RMSE	0.2592	0.2481	0.1618	<b>0.3153</b>	0.1604	0.2222
MI	<b>1.7048</b>	1.0635	1.2759	0.9203	1.0128	0.8703
PSNR	18.8465	15.9394	16.2031	19.7288	15.9977	<b>20.2037</b>
SF	27.0171	<b>39.6167</b>	17.0408	39.5724	14.8905	22.8919
$Q_{AB/F}$	0.5663	<b>0.5849</b>	0.3562	0.2043	0.2241	0.4486
$Q_G$	<b>0.4214</b>	0.3542	0.3460	0.1347	0.2219	0.2554
$Q_{CF}$	<b>0.7460</b>	0.6028	0.6055	0.1317	0.5809	0.6887

**Fig. 14.** Performance comparison of different methods of group 1. ((a) MRI, (b) PET, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).

**Fig. 16(f)).** The edge information is completely retained in the fused images by the IHS+PCA based method (shown in Fig. 14(c), Fig. 15(c) and Fig. 16(c)).

Then, the comparison of objective quality assessment is shown in Table 4. The IHS+PCA based method gives the largest value for the quality metrics of MI. It means that the fused image by the IHS+PCA method has well preserved the original information from the input images. Moreover, the PSNR value of the LES+DC based method is the largest for the images from the group of MRI-PET fusion.

### 5.3. Evaluations on the group of MRI-SPECT fusion

In this subsection, eight examples of MRI-SPECT fusion are for the disease of Subacute Stroke, Cavernous angioma, Vascular dementia, Hypertensive Encephalopathy, Glioma, Metastatic bronchogenic cancer, Alzheimer and AIDS. Each example of disease includes 21 pairs, 25 pairs, 38 pairs, 19 pairs, 20 pairs, 21 pairs, 43 pairs and 15 pairs of MRI and SPECT imaging modalities.

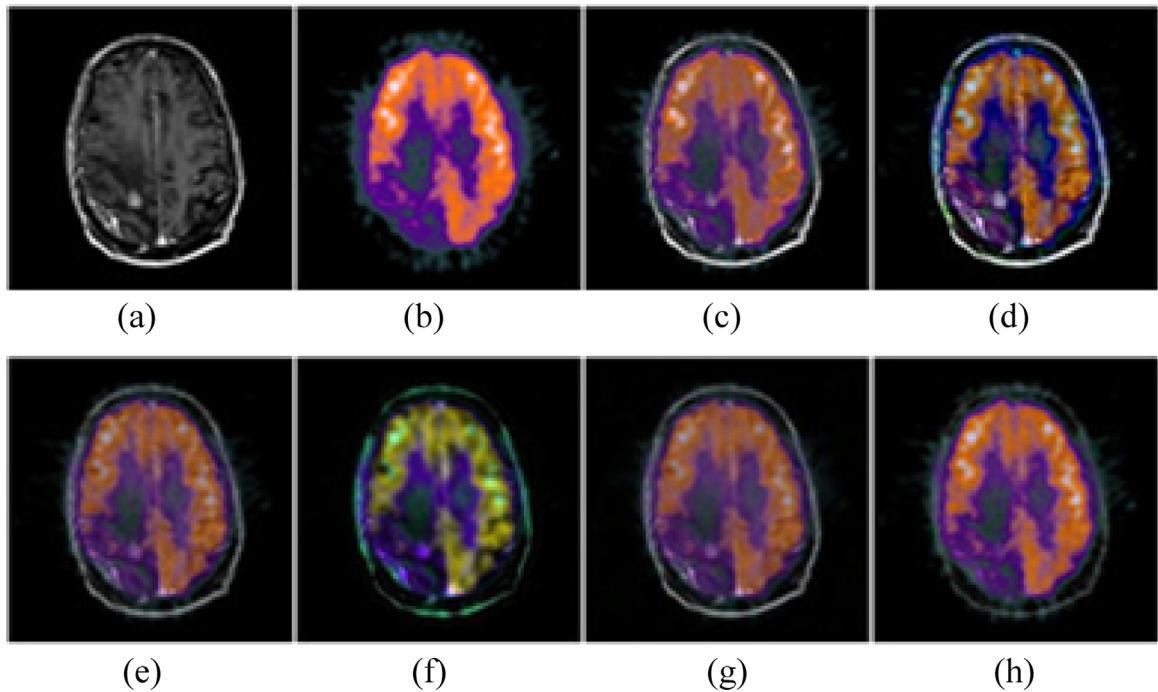
Figs. 17–19 illustrate the input images and the fused image by the methods, described above. Three pairs of MRI and SPECT images are the input images (MRI and SPECT images from the images marked with slice 10 in the disease of Cavernous angioma, marked with slice 12 in the disease of Hypertensive

Encephalopathy and marked with slice 11 in the disease of Metastatic bronchogenic cancer). In Fig. 17(d), the fused image looks very close to the input image of SPECT image (shown in Fig. 17(b)). Furthermore, the enhanced-contrast of the brightness in lesion is clear in the fused image by the NSCT+PCNN method (shown in Fig. 17(f)). However, color distortion exists in the fused images by the NSCT+PCNN (shown in Fig. 18(f) and Fig. 19(f)). In other words, the NSCT+PCNN based method gives a very unstable performance in the images from the group of MRI-SPECT. The white area, located in the input image of MRI, is completely retained in the fused image by IHS+PCA method (shown in Fig. 18(c)).

In addition, Table 5 shows the objective performances of different fusion methods. IHS+PCA based method is always the largest quality indexes for SSIM, RMSE, MI, PSNR,  $Q_G$  and  $Q_{CF}$ . Moreover, the  $Q_{AB/F}$  value of the MoR+MR based method is the largest. It means that most of edge information are transferred from the input images to the fused image.

## 6. Conclusions

Multi-modal medical image fusion plays a major role in biomedical research. In general, the purpose of multi-modal medical

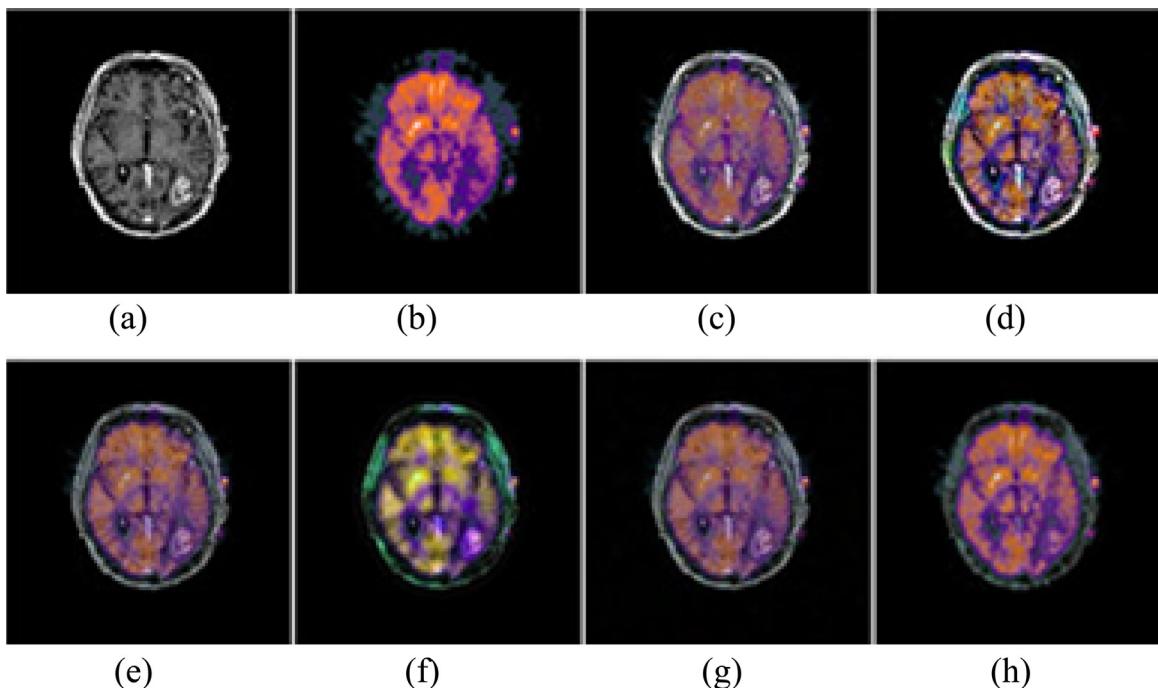


**Fig. 15.** Performance comparison of different methods of group1. ((a) MRI, (b) PET, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).

image is to improve the imaging quality for removing the physical limitations of the imaging technology. The availability of a large number of multi-modal fusion approaches, including image decomposition, image reconstruction, fusion rules and image quality assessments, makes the field of image fusion appealing to be employed by medical imaging community. The main challenge is the gap between the image fusion algorithms and the patient

case in hospital. Image fusion is primarily concerned with image decomposition and reconstruction methods, image fusion rules and image quality assessment resulting from the framework of image fusion. The major problems of the three parts are caused by the additional noise, color distortion, missing features, artificial effects.

A good image decomposition and reconstruction approach is



**Fig. 16.** Performance comparison of different methods of group1. ((a) MRI, (b) PET, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).

**Table 4**

Objective quality assessment of different fusion methods in group of MRI-PET fusion.

Method	IHS + PCA	MoR + MR	DWT + Entropy	NSCT + PCNN	JSP + AR	LES + DC
SSIM	0.6125	0.7611	0.7375	0.5348	0.1720	<b>0.7537</b>
RMSE	0.1221	0.1690	0.1066	<b>0.1891</b>	0.0833	0.1533
MI	<b>1.1666</b>	1.0768	1.1599	0.8447	0.8093	1.1558
PSNR	16.2782	21.8320	20.1295	20.6655	15.1558	<b>25.2012</b>
SF	16.4145	<b>29.2607</b>	14.3663	25.9504	7.8226	7.6548
$Q_{AB/F}$	0.4652	<b>0.6259</b>	0.4003	0.2269	0.1899	0.2557
$Q_G$	<b>0.3246</b>	0.3181	0.2855	0.1377	0.1879	0.2223
$Q_{CF}$	0.6127	<b>0.7676</b>	0.6684	0.4136	0.4673	0.3819

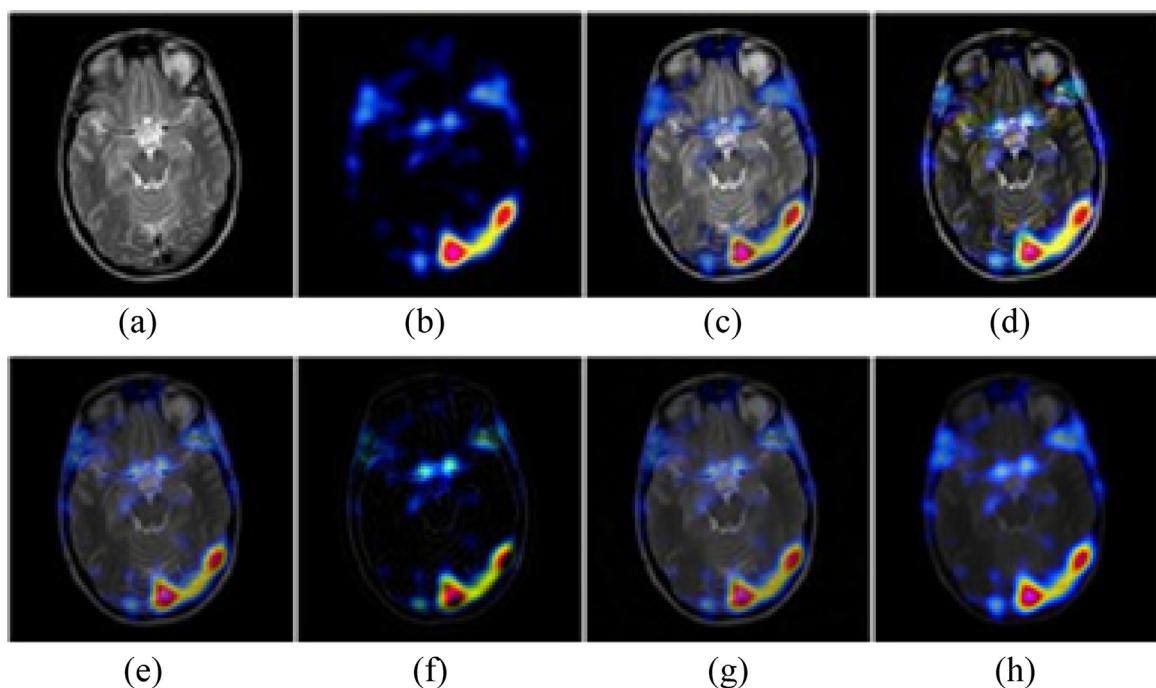
tightly linked to the scale and rotation invariance on medical image fusion. Point of interest is that when addressing the medical image decomposition strategy, the emphasis has been in the direction of developing algorithms that try to extract features at different scales and orientations within images. The tool of multi-scale and multi-orientation is considered as a significant issue in medical image fusion. The methods are introduced from other fields of fusion, such as multi-focus image fusion [69,70], remote sensing image fusion [71], etc. Many multi-scale and multi-orientation tools are proposed, such as PT, DWT, CT, and ST. Another area of interest is to extract the feature in different scales and orientations which is developed from edge-preserving filter [72–74], HVS [75,76], etc.

Conventionally, the image fusion rules such as average, maximum or minimum are simple to implement in a short time. However, it is easy to bring the phenomenon of pseudo-Gibbs, obscure edge. To overcome this problem, an alternate approach is to simulate procedure of image perception and obtainment in the human visual system. The networks of visual neural cells are effective fusion rules reported, such as PCNN, self-generating neural network and visual cortex model. Another approach is the combinations of features rooted in prior knowledge. The prior

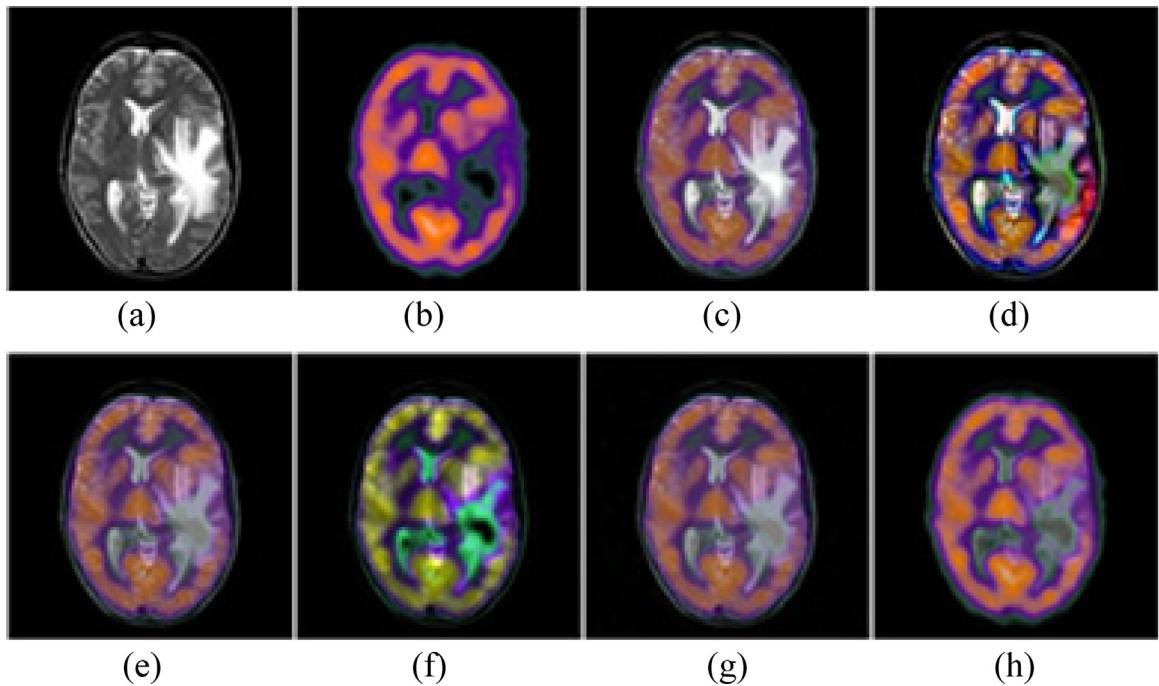
knowledge demonstrates that human is sensitive to high-contrast, intensity, edge information. The prior knowledge based fusion rules are directive contrast, dependencies between different scales or orientations, edge and texture detectors.

Image quality assessment metrics are introduced as the standard to judge the availability of fusion image. The main branches of the metrics are subjective and objective approaches. With respect to the subjective approaches, questionnaire is popular. The key is the design of the item and its score. The most authoritative evaluation may be from the doctor. However, it is difficult to judge the image quality with doctors. Another aspect is objective approach. Objective methods for assessing the perceptual image quality attempt to quantify the error between a fused image and two original images. The kinds of errors are the loss of noise and the increase of information. To quantify the information presented in the fused image, the measurement is that how much information of this fused image can be extracted from the original image. Moreover, the difference of structure between fused and original images is qualified as a measurement.

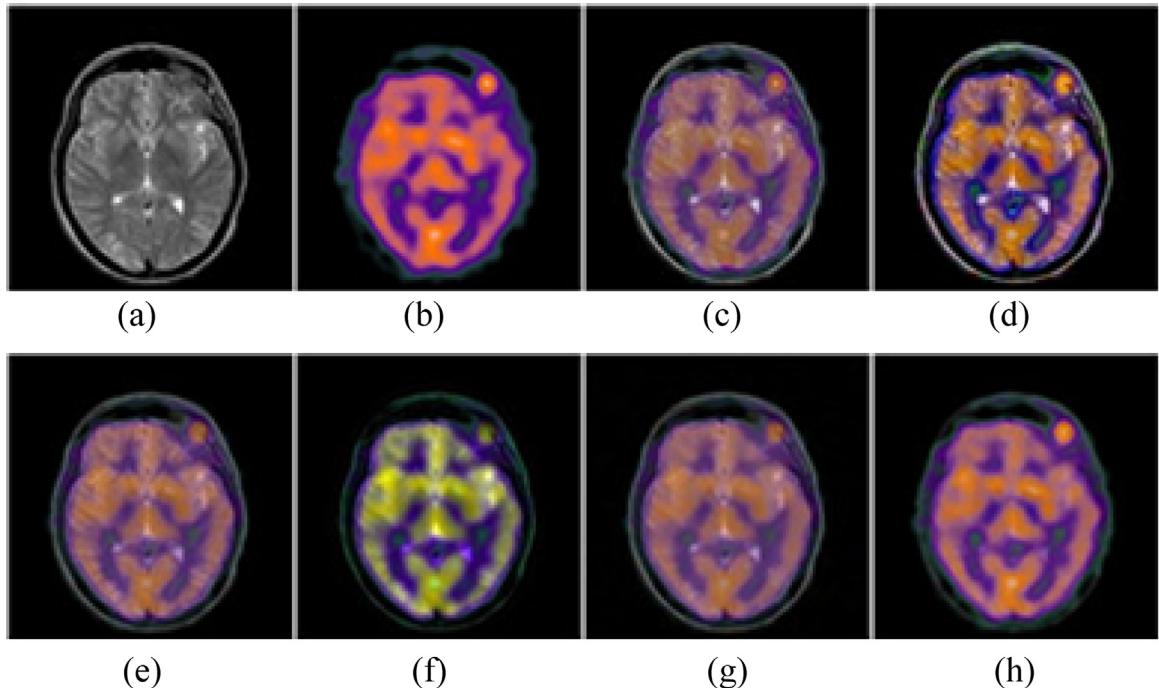
In summary, the advancement in multi-modal medical image fusion largely depends on the trust that the doctors and patients take the remote medical care resulting from the image fusion



**Fig. 17.** Performance comparison of different methods of group1. ((a) MRI, (b) SPECT, (c) IHS + PCA, (d) MoR + MR, (e) DWT + Entropy, (f) NSCT + PCNN, (g) JSP + AR, and (d) LES + DC).



**Fig. 18.** Performance comparison of different methods of group1. ((a) MRI, (b) SPECT, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).



**Fig. 19.** Performance comparison of different methods of group1. ((a) MRI, (b) SPECT, (c) IHS+PCA, (d) MoR+MR, (e) DWT+Entropy, (f) NSCT+PCNN, (g) JSP+AR, and (d) LES+DC).

approaches. The overview in terms of image decomposition and reconstruction, image fusion rules and image quality assessment indicates the importance of this research in improving the image quality. The growth of multiple imaging modalities has enabled progress in the image fusion research. In addition, challenges of image fusion research area are image noise, lack of

sufficient of features in each medical image modality, high cost of imaging equipment, increased computational complexity and dimension difference between images in varied modalities. It is another way that the fused images have provided the observers effectively representation of medical images. The algorithms used for image fusion have resulted in the image with high

**Table 5**

Objective quality assessment of different fusion methods in group of MRI-SPECT fusion.

Method	IHS + PCA	MoR + MR	DWT + Entropy	NSCT + PCNN	JSP + AR	LES + DC
SSIM	<b>0.8558</b>	0.7435	0.7523	0.5608	0.2412	0.5612
RMSE	<b>0.1802</b>	0.1331	0.0934	0.1753	0.0939	0.1213
MI	<b>1.8885</b>	1.2135	1.3472	1.0332	1.1545	1.1967
PSNR	<b>25.6892</b>	21.2706	21.6137	23.0925	21.2881	25.0118
SF	18.0463	20.8946	10.9183	<b>24.4434</b>	7.4713	6.3520
$Q_{AB/F}$	0.6451	<b>0.6468</b>	0.4606	0.2647	0.2052	0.1976
$Q_G$	<b>0.4884</b>	0.3620	0.3658	0.2219	0.2452	0.2102
$Q_{CF}$	<b>0.8105</b>	0.6853	0.6899	0.4791	0.5912	0.4346

resolution and low noise. And these algorithms have been used in clinical institutions. It is expected that the practical advancements would be continued in the future.

## Acknowledgement

This work was supported in part by Natural Science Foundation of China (No. 61272195, 61472055, U1401252, U1301251), Program for New Century Excellent Talents in University of China (NCET-11-1085), Chongqing Outstanding Youth Fund (cstc2014jcyjjq40001) and Chongqing Research Program of Application Foundation and Advanced Technology (cstc2012jja40036).

## References

- [1] A.P. James, B.V. Dasarathy, Medical image fusion: a survey of the state of the art, *Inf. Fusion* 19 (2014) 4–19.
- [2] R. Shen, I. Cheng, A. Basu, Cross-scale coefficient selection for volumetric medical image fusion, *IEEE Trans. Biomed. Eng.* 60 (4) (2013) 1069–1079.
- [3] C. He, Q. Liu, H. Li, et al., Multimodal medical image fusion based on IHS and PCA, *Proc. Eng.* 7 (2010) 280–285.
- [4] S. Daneshvar, H. Ghassemian, MRI and PET image fusion by combining IHS and retina-inspired models, *Inf. Fusion* 11 (2) (2010) 114–123.
- [5] Matsopoulos, G.K., Marshall, S., Brunt, J.N.H. 1994. Multiresolution morphological fusion of MR and CT images of the human brain, *IEE Proceedings of Visual Image Signal Processing*, vol. 141 (3), pp. 137–142.
- [6] Patil, U., Mudengudi, U. 2011. Image fusion using hierarchical PCA, *Proceedings of Internal Conference Image Information Processing*, pp. 1–6.
- [7] Q. Guihong, Z. Dali, Y. Pingfan, Medical image fusion by wavelet transform modulus maxima, *Opt. Express* 9 (4) (2001) 184–190.
- [8] R. Singh, A. Khare, Fusion of multimodal medical images using Daubechies complex wavelet transform—A multiresolution approach, *Inf. Fusion* (2012).
- [9] Y. Zheng, A.E. Essock, B.C. Hansen, et al., A new metric based on extended spatial frequency and its application to DWT based fusion algorithms, *Inf. Fusion* 8 (2) (2007) 177–192.
- [10] Singh, R., Vatsa, M., Noore, A. 2009. Multimodal medical image fusion using redundant discrete wavelet transform, In: *IEEE Seventh International Conference on Advances in Pattern Recognition*, 2009, pp. 232–235.
- [11] Liu, Y., Yang, J., Sun, J. 2010. PET/CT medical image fusion algorithm based on multiwavelet transform, In: *IEEE 2nd International Conference on Advanced Computer Control*, 2010, pp. 264–268.
- [12] Xue-jun, W., Ying, M. 2010. A medical image fusion algorithm based on lifting wavelet transform, In: *IEEE International Conference on Artificial Intelligence and Computational Intelligence*, 2010 (3), pp. 474–476.
- [13] L. Yang, B.L. Guo, W. Ni, Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform, *Neurocomputing* 72 (1) (2008) 203–211.
- [14] G. Bhatnagar, Q.M. Wu, Z. Liu, Directive contrast based multimodal medical image fusion in NSCT domain, *IEEE Trans. Multimed.* 15 (5) (2013) 1014–1024.
- [15] Q. Miao, C. Shi, P. Xu, et al., A novel algorithm of image fusion using shearlets, *Opt. Commun.* 284 (6) (2011) 1540–1547.
- [16] L. Wang, B. Li, L. Tian, Multi-modal medical volumetric data fusion using 3D discrete shearlet transform and global-to-local rule, *IEEE Trans. Biomed. Eng.* (2014).
- [17] H. Yin, S. Li, L. Fang, Simultaneous image fusion and super-resolution using sparse representation, *Inf. Fusion* 14 (3) (2013) 229–240.
- [18] B. Yang, S. Li, Pixel-level image fusion with simultaneous orthogonal matching pursuit, *Inf. Fusion* 13 (1) (2012) 10–19.
- [19] H. Yin, S. Li, Multimodal image fusion with joint sparsity model, *Opt. Eng.* 50 (6) (2011) 7–10.
- [20] N. Yu, T. Qiu, F. Bi, et al., Image features extraction and fusion based on joint sparse representation, *IEEE J. Sel. Top. Signal Process.* 5 (5) (2011) 1074–1082.
- [21] S. Li, H. Yin, L. Fang, Group-sparse representation with dictionary learning for medical image denoising and fusion, *IEEE Trans. Biomed. Eng.* 59 (12) (2012) 3450–3459.
- [22] J. Hu J, S. Li, The multiscale directional bilateral filter and its application to multisensor image fusion, *Inf. Fusion* 13 (3) (2012) 196–206.
- [23] S. Li, X. Kang X, J. Hu, Image fusion with guided filtering, *IEEE Trans. Image Process.* 22 (7) (2013) 2864–2875.
- [24] Z. Xu, Medical image fusion using multi-level local extrema, *Inf. Fusion* 19 (2014) 38–48.
- [25] V. Barra, J.Y. Boire, A general framework for the fusion of anatomical and functional medical images, *NeuroImage* 13 (3) (2001) 410–424.
- [26] Wang, P.Y., Dang, J.W., Li, Q., et al. 2007. Multimodal medical image fusion using fuzzy radial basis function neural networks, In: *IEEE International Conference on Wavelet Analysis and Pattern Recognition 2007* (2), pp. 778–782.
- [27] U. Javed, M.M. Riaz, A. Ghafoor, et al., Mri and pet image fusion using fuzzy logic and image local features, *Sci. World J.* (2014).
- [28] W. Hao-quan, X. Hao, Multi-mode medical image fusion algorithm based on principal component analysis, In: *IEEE International Symposium on Computer Network and Multimedia Technology*, 2009, pp. 1–4.
- [29] H. Chen, A multiresolution image fusion based on principle component analysis, In: *IEEE Fourth International Conference on Image and Graphics*, 2007, pp. 737–741.
- [30] R. Vijayarajan, S. Muttan, Iterative block level principal component averaging medical image fusion, *Opt. Int. J. Light. Electron Opt.* 125 (17) (2014) 4751–4757.
- [31] L. Wang, B. Li, L.F. Tian, EGGDD: an explicit dependency model for multi-modal medical image fusion in shift-invariant shearlet transform domain, *Inf. Fusion* 19 (2014) 29–37.
- [32] L. Wang, B. Li, L. Tian, Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients, *Inf. Fusion* 19 (2014) 20–28.
- [33] G. Bhatnagar, Q.M.J. Wu, Z. Liu, Human visual system inspired multi-modal medical image fusion framework, *Expert. Syst. Appl.* 40 (5) (2013) 1708–1720.
- [34] Q.P. Zhang, W.J. Tang, L.L. Lai, et al., Medical diagnostic image data fusion based on wavelet transformation and self-organizing features mapping neural networks, *IEEE Proc. Int. Conf. Mach. Learn. Cybern.* 2004 5 (2004) 2708–2712.
- [35] Z. Liu, H. Yin, Y. Chai, et al., A novel approach for multimodal medical image fusion, *Expert. Syst. Appl.* 41 (16) (2014) 7425–7435.
- [36] W. Li, X.F. Zhu, A new image fusion algorithm based on wavelet packet analysis and PCNN, *IEEE Proc. Int. Conf. Mach. Learn. Cybern.* 2005 9 (2005) 5297–5301.
- [37] N. Wang, Y. Ma, K. Zhan, et al., Multimodal medical image fusion framework based on simplified PCNN in nonsubsampled contourlet transform domain, *J. Multimed.* 8 (3) (2013) 270–276.
- [38] S. Daneshvar, H. Ghassemian, MRI and PET images fusion based on human retina model, *J. Zhejiang Univ. Sci. A* 8 (10) (2007) 1624–1632.
- [39] J.H. Jang, Y. Bae, J.B. Ra, Contrast-enhanced fusion of multisensor images using subband-decomposed multiscale retinex, *IEEE Trans. Image Process.* 21 (8) (2012) 3479–3490.
- [40] Y. Zheng, E.A. Essock, B.C. Hansen, et al., A new metric based on extended spatial frequency and its application to DWT based fusion algorithms, *Inf. Fusion* 8 (2) (2007) 177–192.
- [41] Z. Wencang, C. Lin, Medical image fusion method based on wavelet multi-resolution and entropy, *IEEE International Conference on Automation and Logistics*, 2008, pp. 2329–2333.
- [42] G. Bhatnagar, Q.M.J. Wu, Z. Liu, Directive contrast based multimodal medical image fusion in NSCT domain, *IEEE Trans. Multimed.* 15 (5) (2013) 1014–1024.
- [43] Garg, S., Ushah Kiran, K., Mohan, R., et al. 2006. Multilevel medical image fusion using segmented image by level set evolution with region competition, In: *IEEE Engineering 27th Annual International Conference of Medicine and Biology Society*, 2005, pp. 7680–7683.
- [44] X. Li, X. Tian, Y. Sun, et al., Medical image fusion by multi-resolution analysis of wavelets transform, *Wavel. Anal. Appl.* (2007) 389–396.
- [45] H.R. Sheikh, A.C. Bovik, Image information and visual quality, *IEEE Trans. Image Process.* 15 (2) (2006) 430–444.
- [46] Y. Yang, X. Wang, T. Guan, et al., A multi-dimensional image quality prediction model for user-generated images in social networks, *Inf. Sci.* 281 (2014) 601–610.

- [47] Y. Yang, X. Wang, Q. Liu, et al., User models of subjective image quality assessment on virtual viewpoint in free-viewpoint video system, *Multimed. Tools Appl.* (2014) 1–21.
- [48] Y. Yang, Q. Dai, Contourlet-based image quality assessment for synthesised virtual image, *Electron. Lett.* 46 (7) (2010) 492–494.
- [49] Z. Wang, L. Lu, A.C. Bovik, Video quality assessment based on structural distortion measurement, *Signal Process. Image Commun.* 19 (2) (2004) 121–132.
- [50] A. Hore, D. Ziou, Image quality metrics: PSNR vs. SSIM, *ICPR* 34 (2010) 2366–2369.
- [51] S. Zheng, W.Z. Shi, J. Liu, et al., Multisource image fusion method using support value transform, *IEEE Trans. Image Process.* 16 (7) (2007) 1831–1839.
- [52] H. Zhao, Z. Shang, Y.Y. Tang, et al., Multi-focus image fusion based on the neighbor distance, *Pattern Recognit.* 46 (3) (2013) 1002–1011.
- [53] Q. Xiao-Bo, Y. Jing-Wen, X. Hong-Zhi, et al., Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in non-subsampled contourlet transform domain, *Acta Autom. Sin.* 34 (12) (2008) 1508–1514.
- [54] Buades, A., Coll, B., Morel, J.M. 2005. A non-local algorithm for image denoising, In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005 (2), pp. 60–65.
- [55] V. Petrović, Subjective tests for image fusion evaluation and objective metric validation, *Inf. Fusion* 8 (2) (2007) 208–216.
- [56] H.R. Sheikh, M.F. Sabir, A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, *IEEE Trans. Image Process.* 15 (11) (2006) 3440–3451.
- [57] Z. Wang, A.C. Bovik, H.R. Sheikh, et al., Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (2004) 600–612.
- [58] M. Hossny, S. Nahavandi, D. Creighton, Comments on 'Information measure for performance of image fusion', *Electron. Lett.* 44 (18) (2008) 1066–1067.
- [59] Y. Horibe, Entropy and correlation, *IEEE Trans. System, Man, Cybern.* SMC-15 (1985), pp. 641–642.
- [60] Z. Wang, A.C. Bovik, A universal image quality index, *IEEE Trans. Signal Process. Lett.* 9 (3) (2002) 81–84.
- [61] N. Cvejic, A. Loza, D. Bull, et al., A similarity metric for assessment of image fusion algorithms, *Int. J. Signal Process.* 2 (3) (2005) 178–182.
- [62] J. Zhao, R. Laganiere, Z. Liu, Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement, *Int. J. Innov. Comput. Inf. Control.* 3 (6) (2007) 1433–1447.
- [63] G. Piella, H. Heijmans, A new quality metric for image fusion, *IEEE Int. Conf. Proc. Image Process.* 2003 3 (2) (2003) (173–6).
- [64] C.S. Xydeas, V.S. Petrović, Objective pixel-level image fusion performance measure, *Int. Soc. Opt. Photon.* 2000 (2000) 89–98.
- [65] S. Li, B. Yang, J. Hu, Performance comparison of different multi-resolution transforms for image fusion, *Inf. Fusion* 12 (2) (2011) 74–84.
- [66] C.S. Xydeas, V. Petrović, Objective image fusion performance measure, *Electron. Lett.* 36 (4) (2000) 308–309.
- [67] A. Mittal, R. Soundararajan, A.C. Bovik, Making a "completely blind" image quality analyzer, *IEEE Trans. Signal Process. Lett.* 20 (3) (2013) 209–212.
- [68] Johnson, K.A., Becker, J.A. 2016. The Whole Brain Altas [Online], (Available): <<http://www.med.harvard.edu/aanlib/>>.
- [69] S. Li, B. Yang, Multifocus image fusion by combining curvelet and wavelet transform, *Pattern Recognit. Lett.* 29 (9) (2008) 1295–1301.
- [70] Q. Zhang, B. Guo, Multifocus image fusion using the nonsubsampled contourlet transform, *Signal Process.* 89 (7) (2009) 1334–1346.
- [71] F. Nencini, A. Garzelli, S. Baronti, et al., Remote sensing image fusion using the curvelet transform, *Inf. Fusion* 8 (2) (2007) 143–156.
- [72] K. Subr, C. Soler, F. Durand, Edge-preserving multiscale image decomposition based on local extrema, *ACM Trans. Graph.* 28 (5) (2009) 147.
- [73] K. He, J. Sun, X. Tang, Guided image filtering, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (6) (2013) 1397–1409.
- [74] F. Durand, J. Dorsey, Fast bilateral filtering for the display of high-dynamic-range images, *ACM Trans. Graph.* 21 (3) (2002) 257–266.
- [75] N. Nill, A visual model weighted cosine transform for image compression and quality assessment IEEE Trans. Communications 33 (6) (1985) 551–557.
- [76] M. Li, W. Cai, Z. Tan, A region-based multi-sensor image fusion scheme using pulse-coupled neural network, *Pattern Recognit. Lett.* 27 (16) (2006) 1948–1956.



**Jiao Du** was born in Chongqing, P.R. China, in 1988. She received M.S. degree from Chongqing University of Posts and Telecommunications, Chongqing, P.R. China. Currently, she is a Ph.D. candidate in computer science and technology, with the Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications. Her research interests include pattern recognition and image processing.



**Weisheng Li** graduated from School of Electronics & Mechanical Engineering at Xidian University, Xi'an, China in July 1997. He received M.S. degree and Ph.D. degree from School of Electronics & Mechanical Engineering and School of Computer Science & Technology at Xidian University in July 2000 and July 2004, respectively. Currently he is a professor of Chongqing University of Posts and Telecommunications. His research focuses on intelligent information processing and pattern recognition.



**Ke Lu** graduated from Ningxia University, Yinchuan, China in 1993. He received M.S. degree and Ph.D. degree at Northwest University, Xi'an, China in 1999 and July 2003. Currently he is a professor of University of Chinese Academy of Sciences. His research focuses on image processing and pattern recognition.



**Bin Xiao** was born in 1982. He received his B.S. and M. S. degrees in Electrical Engineering from Shaanxi Normal University, Xian, China in 2004 and 2007, received his Ph. D. degree in computer science from Xidian University, Xi'an, China. He is now working as an associate professor at Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include image processing, pattern recognition and digital watermarking.