

Adaptive decomposition method for multi-modal medical image fusion

ISSN 1751-9659

Received on 30th October 2017

Revised 10th January 2018

Accepted on 16th March 2018

E-First on 17th April 2018

doi: 10.1049/iet-ipr.2017.1067

www.ietdl.org

Jing Wang^{1,2}, Xiongfei Li^{1,2}, Yan Zhang^{1,2}, Xiaoli Zhang^{1,2}✉

¹Key Laboratory of Symbol Computation and Knowledge Engineer of Ministry of Education, Jilin University, No. 2699 Qianjin Street, Changchun, People's Republic of China

²College of Computer Science and Technology, Jilin University, No. 2699 Qianjin Street, Changchun, People's Republic of China

✉ E-mail: zhangxiaoli@jlu.edu.cn

Abstract: In traditional image fusion, source images are separated into a fixed space. The low-frequency part and the high-frequency part are not discriminated according to the nature of the image. Traditional fusion rules often use a fixed proportion, causing colour distortion. In this study, a new adaptive decomposition algorithm is proposed to distinguish high frequency and low frequency of structure image to obtain smoothing layer and texture layer. The smoothing layer of the structural image and the colour information of the function image are fused according to dynamic rules, and then the texture layer is added. On the basis of the objective evaluation metrics, the spectral information evaluation metrics are introduced to evaluate the retention of colour. In the experiments, the proposed method is compared with other six classical image fusion methods. The experiment results show that the proposed method can retain the colour information and structure information very well at the same time. Concerning subjective and objective evaluation, the proposed algorithm is superior to other algorithms.

1 Introduction

With the continuous development of medical imaging engineering and computer technology, advanced imaging devices provide a variety of modal medical images for clinical diagnosis. These images reflect different information about human organs and diseased tissue. In clinical diagnosis, a single modal image cannot provide sufficient information, so multi-modal medical image fusion is necessary. If the multi-modal medical images are properly fused, the structural information and the functional information can be organically combined to express in one fused image. Fusion information from different modalities can help doctors make a more accurate diagnosis or develop a more appropriate treatment.

Multi-modal medical image fusion can make up for the lack of individual inspection imaging. At present, modal medical images include B-ultrasound, digital subtraction angiography, computed radiography, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and single photon emission CT (SPECT). Each modal has its characteristics and advantages, but it also has defects and limitations in the imaging [1]. For example, the resolution of CT is very high, but this is difficult for CT to distinguish the similar density of the organisation; MRI has superior display ability in the soft tissue, but it shows poor in the bone lesions and calcification lesions; PET and SPECT have excellent performance in tissue contrast, but their resolution is low [2]. PET/SPECT are obtained by injecting metabolites labelled with radiopharmaceuticals [3]. The colours on the image display the metabolic status of the human body, showing the difference of the intensity of the different tissues. Various colours represent different organisations and the depth of the colour reflects the activity of the cells. The cancerous parts are often metabolically stronger than the normal parts, and then the disease is diagnosed by detecting physiological metabolic abnormalities. This is important to ensure that the fusion image is undistorted in medical image fusion.

The traditional medical image fusion methods preserved fuzzy structure information [4–6], lacking colour information and even producing colour distortion. There are two main reasons: (i) in image decomposition stage, source images are separated into a fixed space, the low-frequency part and the high-frequency part are not discriminated according to the nature of the image. (ii) In

image fusion stage, fusion rules often use a fixed proportion, causing colour distortion. In this paper, a new adaptive decomposition algorithm is proposed to distinguish high frequency and low frequency of structure image to obtain smoothing layer and texture layer. The smoothing layer and the colour information extracted from the function image are fused according to dynamic rules, and the fused image is obtained by adding the texture layer.

The rest of this paper is organised as follows: in Section 2, related work is reviewed; the proposed fusion algorithm is presented in Section 3; in Section 4, experimental results are displayed and evaluation results are discussed. Conclusions are made in Section 5.

2 Related work

Multi-scale decomposition (MSD) methods are the most popular tools used in multi-modal medical image fusion. Methods based on pyramid and wavelets are the two traditional methodologies of MSD [7]. Laplacian pyramid (LP) [8, 9] is the most representative method of pyramid-based methods. In pyramid-based algorithms, the original image is filtered constantly to form a tower structure. In each layer of the tower, an algorithm is selected for image fusion, resulting in a synthetic tower structure. Then, the synthetic tower structure is reconstructed to obtain the final image. Decomposition algorithms based on pyramid have inherent defects [10–12], and it introduces the up and down sampling operations, which will lead to the lack of translation invariance. Algorithms are based on pyramid lack directional decomposition, so the image orientation information is inadequate.

The wavelet transform is used to process the image on different frequency channels [13]. First, the source image is decomposed by multi-layer wavelet to obtain a series of sub-images. Second, the feature selection is made on the transform domain, and the fusion image is created. Finally, the fusion image is reconstructed by an inverse transform. Discrete wavelet transformation (DWT) [14] preserves different frequency information in a stable form and permits perfect localisation both in time and spatial frequency domain. A simple DWT-based medical image fusion, which follows weighted fusion rule, has been proposed in [15]. However, the method fails to meet the requirements of shift invariance. Contourlet [16, 17] performs better for representing images than

the wavelets with an advantage of owning contour segments to capture the geometrical structures of images. Two stages of the methods are MSD by pyramid transformation and directional decomposition by various directional filter banks. In [16], a novel contourlet contrast measurement focusing on CT–MR is developed and different fusion rules are selected in the low-pass and high-pass sub-bands. A novel image fusion algorithm that combines nonlinear approximation of contourlet transform with image regional features is proposed in [17]. Owing to the use of the LP, the detail image decomposed by contourlet oscillates around the singular point. Pyramid transformation and directional filter banks without subsampled constitute the non-subsampled contourlet transformation (NSCT) method [18–20]. The source medical images are first transformed by NSCT followed by combining low- and high-frequency components. Two different fusion rules based on phase congruency and directive contrast are proposed in [18]. Yang *et al.* [19] present a novel multi-modal medical image fusion method that adopts a multi-scale geometric analysis of the NSCT with type-2 fuzzy logic techniques. Bhateja *et al.* [20] present a two-stage multi-modal fusion framework using the cascaded combination of stationary wavelet transform and NSCT domains for images acquired using two distinct medical imaging sensor modalities. NSCT is a new type of pan-shift, multi-scale, multi-directional rapid transformation. However, the choice of fusion coefficient in NSCT domain affects the quality of the fused images.

Different from the traditional MSD methods, sparse representation (SR) is proposed as a tool for image fusion with the assumption that both the high- and low-frequency images share the same set of sparse coefficients [21–23]. It is necessary to learn the dictionary of the low-frequency sub-band to construct the base atom of the SR. However, the dictionaries have to be designed separately according to the characteristics of the images. Edge-preserving filters, which smooth the images while preserving the edges, have been applied to image processing. Multi-scale edge-preserving decomposition based on weighted least-squares (WLSs) filter [24] is applied in the decomposition stage in the image fusion. The SR-based methods and the edge-preserving-based methods have the outstanding performance in image fusion in recent years. They are chosen as the comparison algorithms applied in multi-modal medical image fusion.

The source images are decomposed into a fixed space by the methods based on LP, DWT and contourlet, leading to the low-frequency part and the high-frequency part are not discriminated according to the nature of the image. The fusion rules in the methods based on NSCT, SR and WLS often use a fixed proportion, causing colour distortion. A new adaptive decomposition algorithm with dynamic rules is proposed in this paper to solve the problems.

3 Image fusion method

A new method is used to divide the structure image into two layers: smoothing layer and texture layer. The high-frequency information of structure image is reserved in texture layer. Texture information plays an important role in disease diagnosis, which indicates structural details. Function image contains only low-frequency information, lacking texture information. The useful information in the function image is usually represented by colour. By detecting the colour information of the pixel, the area where the information amount is large can be separated from the entire image. A new dynamic fusion rule is used to fuse the structure image and the function image. In this paper, the original structure image is represented by I , the original function image is represented by C , and the fused image is represented by F .

3.1 Adaptive decomposition

In this paper, a new adaptive decomposition algorithm is proposed to distinguish high frequency and low frequency of I to obtain smoothing layer and texture layer.

The ideal smoothing layer \tilde{s} satisfies the following two conditions: (i) the area outside the boundary is as smooth as

possible and (ii) it should have a high similarity with I . It could be described as [25]

$$\tilde{s} = \arg \min_{psp} \left\{ (s(psp) - I(psp))^2 + \mu \left[a_{I,x}(psp) \left[\left(\frac{\partial s}{\partial x} \right)(psp) \right]^2 + a_{I,y}(psp) \left[\left(\frac{\partial s}{\partial y} \right)(psp) \right]^2 \right] \right\} \quad (1)$$

where psp denotes the spatial coordinate of a pixel, μ is used to balance the gradient and grey value differences, a larger μ results in smoother images of s . $[s(psp) - I(psp)]^2$ represents the grey value difference between s and I , $[(\partial s / \partial x)(psp)]^2$ and $[(\partial s / \partial y)(psp)]^2$ represent the horizontal and vertical gradients of s , $a_{I,x}(psp)$ and $a_{I,y}(psp)$ represent the weight of the horizontal and vertical gradients. Minimise the first term $[s(psp) - I(psp)]^2$ enables s to be close to I and minimise the second term $\mu(a_{I,x}(psp)[(\partial s / \partial x)(psp)]^2 + a_{I,y}(psp)[(\partial s / \partial y)(psp)]^2)$ enables s to be smooth. Edge-preserving smoothing requires that s should be as smooth as possible everywhere, except across the edges, while ensuring s be maximally close to I . In (1), the values of $a_{I,x}(psp)$ and $a_{I,y}(psp)$ are related to the horizontal and vertical gradients of the source image I . To smooth I , only the edge portion of I should be reserved in s as much as possible. In other words, $a_{I,x}(psp)$ and $a_{I,y}(psp)$ control the degree of retention of the edge information. The values of $a_{I,x}(psp)$ and $a_{I,y}(psp)$ can be obtained from the source image I . They can be defined as follows:

$$a_{I,x}(psp) = \begin{cases} \sin\left(\frac{\pi}{2}\beta\left(\frac{|\omega_x|}{\omega_1}\right)\right) & \text{if } |\omega_x| \leq \omega_1 \\ 1 & \text{if } |\omega_x| > \omega_1 \end{cases} \quad (2)$$

$$a_{I,y}(psp) = \begin{cases} \sin\left(\frac{\pi}{2}\beta\left(\frac{|\omega_y|}{\omega_2}\right)\right) & \text{if } |\omega_y| \leq \omega_2 \\ 1 & \text{if } |\omega_y| > \omega_2 \end{cases} \quad (3)$$

where ω_x and ω_y can be represented by horizontal gradient and vertical gradient in the source image I : $\omega_x = (\partial I / \partial x)(psp)$ and $\omega_y = (\partial I / \partial y)(psp)$. ω_1 and ω_2 are two thresholds which are the boundaries of the spectral segmentation of the original image. Take the horizontal direction as an example: if $|\omega_x| > \omega_1$, it indicates that the pixel belongs to the high-frequency part. The high-frequency part should be smoothed in (1) in the decomposition process, so $a_{I,x}(psp)$ is given the maximum value 1, $a_{I,x}(psp) = 1$. If $|\omega_x| \leq \omega_1$, it indicates that the pixel belongs to the low-frequency part. As the low-frequency part tends to be smooth, $a_{I,x}(psp)$ is assigned 0–1, which expression is referred to [26]. The function $\beta(\omega)$ is an arbitrary function such that if $\omega \leq 0$, $\beta(\omega) = 0$ or if $\omega \geq 1$, $\beta(\omega) = 1$ and $\beta(\omega) + \beta(1 - \omega) = 1 \quad \forall w \in [0, 1]$. Many functions satisfy these properties, and the most used one is $\beta(\omega) = \omega^4(35 - 84\omega + 70\omega^2 - 20\omega^3)$ [27]. If ω_1 and ω_2 get too large values, smoothing layer s will contain high-frequency information; if they get too small values, the textural layer t will contain low-frequency information. Hence, proper value setting for the two parameters can guarantee full separation of high-frequency and low-frequency information of the original image.

In the proposed algorithm, the edge detection algorithm based on Fourier transform [28] is used to determine the values of ω_1 and ω_2 . The goal of the edge detection algorithm is to find the appropriate boundary to divide the frequency domain into a particular interval. These intervals need to properly reflect the distribution characteristics of the frequency domain. The basic edge detection algorithm needs to specify the number of filters required. In this paper, the filter number is set to two.

Fourier spectrum is obtained in the horizontal and vertical directions while fast Fourier transformation (FFT) [29] is applied to the source image in the horizontal and vertical directions. Take the horizontal direction as an example. Let $\{M_k\}_{k=1}^M$ denote the set of M detected maxima in the magnitude of the Fourier spectrum

[26]. Without any loss of generality, we assume that this set is sorted by decreasing values $M_1 \geq M_2 \dots \geq M_M$. We can formulate it by keeping all maxima larger than the threshold $M_M + \delta(M_1 - M_M)$, where δ corresponds to the relative amplitude ratio. The experiments show that value of $\delta = 0.3$ seems to give consistent results which correspond to a tradeoff between too much detection and a good separation of the information in the Fourier spectrum [26]. Here, the abscissa value of the horizontal line represented by the threshold and the first intersection of the spectrum is assigned to ω_1 . ω_2 can be got in the same way. Fig. 1 shows a horizontal Fourier spectrum obtained from FFT and the horizontal blue dotted line represents the value obtained by $M_M + \delta(M_1 - M_M)$. The first intersection of this line and the spectrum is A . The abscissa value of A is assigned to ω_1 .

By using the matrix notation, (1) can be rewritten as equation below:

$$\tilde{s} = \arg \min_{psp} \sum \left\{ (s - I)^T (s - I) + \mu (s^T D_x^T A_x D_x s + s^T D_y^T A_y D_y s) \right\} \quad (4)$$

where D_x and D_y denote the discrete differential operators, A_x and A_y represent the diagonal matrixes and the elements on the main symmetry axis are $a_{l,x}(psp)$ and $a_{l,y}(psp)$. The vector s represents the unique solution of the following linear system:

$$(E + \mu L_g)s = I \quad (5)$$

where $L_g = D_x^T A_x D_x + D_y^T A_y D_y$.
 s can be obtained from (5)

$$s = (E + \mu L_g)^{-1} I \quad (6)$$

In practical applications, $a_{l,x}(psp) = a_{l,y}(psp) = a$, (6) can be simplified as

$$s = (E + \mu a L)^{-1} I \quad (7)$$

where $L = D_x^T D_x + D_y^T D_y$ is the ordinary Laplacian matrix.

Texture layer t can be obtained while subtracting the smoothing layer from the source image:

$$t = I - s \quad (8)$$

3.2 Dynamic fusion

While reconstructing the low-frequency sub-band, the traditional selections of fusion coefficient are the absolute maximum method, weighted average method or regional energy method [30]. These methods have a limited efficiency in the retention of the edge, leading edge artefacts. To solve the above problem, a new dynamic fusion rule is proposed to fuse the low-frequency layer. In the new fusion method, the smoothing layer of the structure image and the colour information of the function image are fused, and then the texture layer is added.

The valid information in medical function image is represented by colour. Areas with large amounts of information can be segmented from the entire image by detecting the pixel information. To facilitate the calculation, this paper uses the grey image instead of an original image when extracting valid information from function image. The traditional method to segment function image is described as below:

$$w(psp) = \begin{cases} 1 & \text{if } G_f(psp) \geq \tau \\ 0 & \text{if } G_f(psp) < \tau \end{cases} \quad (9)$$

where $G_f(psp)$ represents the grey value of psp in C , τ is a pending threshold. In (9), when $G_f(psp) \geq \tau$, the pixel belongs to the valid information, so we set $w(psp) = 1$, and it can be reserved in fusion

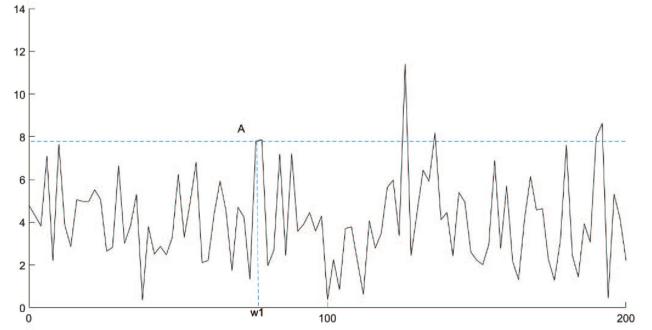


Fig. 1 Fourier spectrum obtained from FFT

image. When $G_f(psp) < \tau$, the pixel belongs to the invalid information, so we set $w(psp) = 0$, and it cannot be reserved in fusion image.

The structure images contain high-frequency information and low-frequency information. However, the function images usually contain only low-frequency information, lacking high-frequency information. The low-frequency information of function images refers to the valid colour information. Therefore, the focus of image fusion is on the fusion of smoothing layer and valid colour information. $w(psp)$ represents the probability of the pixel belongs to valid colour information in C , that is, the weight of the pixel at the low-frequency fusion layer.

The fusion rule of smoothing layers s and C at the low-frequency fusion layer is

$$F'(psp) = w(psp) \times C(psp) + (1 - w(psp)) \times s(psp) \quad (10)$$

A function image C is given, and the image size is $M \times N$, τ is used to divide C into two categories $\{D_i\}_{i=1,2}$. $\omega_i (i = 1, 2)$ represents the proportion of the pixels contained in D_i , $\omega_1 = \sum_{j=0}^{\tau} n_j / (M \times N)$, $\omega_2 = \sum_{j=\tau+1}^{255} n_j / (M \times N)$. n_j denotes the number of pixels which grey value are j . μ_i represents the mean value of C_i , $\mu_1 = (\sum_{j=0}^{\tau} j \times n_j) / \sum_{j=0}^{\tau} n_j$, $\mu_2 = (\sum_{j=\tau+1}^{255} j \times n_j) / \sum_{j=\tau+1}^{255} n_j$. The interclass variance of C is defined below:

$$\sigma_{\tau}^2 = \sum_{i=1}^2 \omega_i (\mu_i - \mu)^2 \quad (11)$$

where μ represents the mean value of all pixels in C , $\mu = \sum_{j=0}^{255} j \times (n_j / (M \times N))$.

According to Otsu *et al.* [31], $\tau = \arg \max \{\sigma_{\tau}^2\}$ is the optimal threshold, which can maximise interclass variance σ_{τ}^2 .

The fused image will produce a pseudo-edge near the boundary, which is due to the unintentional value of $w(psp)$. In (9), when $G_f(psp) < \tau$, $w(psp) = 0$, it indicates that the pixel is the invalid information. This will result in the loss of some colour information and the artefacts. This paper adjusts $w(psp)$ by considering the grey information of smoothing layer and colour information of function image.

Given a sigmoid function $S(x)$

$$S(x) = \frac{1}{1 + e^{-x}} \quad (12)$$

A sigmoid function is a mathematical function having an ‘S’-shaped curve, and it is a monotonically increasing smooth curve. It has finite limits at negative infinity and infinity, most going from 0 to 1, which corresponds to the range of weights. This feature can be used to measure the probability of one pixel belonging to the valid information. It can avoid artefacts in weights calculation in the image fusion due to its continuous differentiable. By introducing the sigmoid function, the probability can be assigned a lot of values instead of only 0 or 1. In fact, the probability of one pixel belonging to the valid information is proportional to the grey value of the pixel itself.

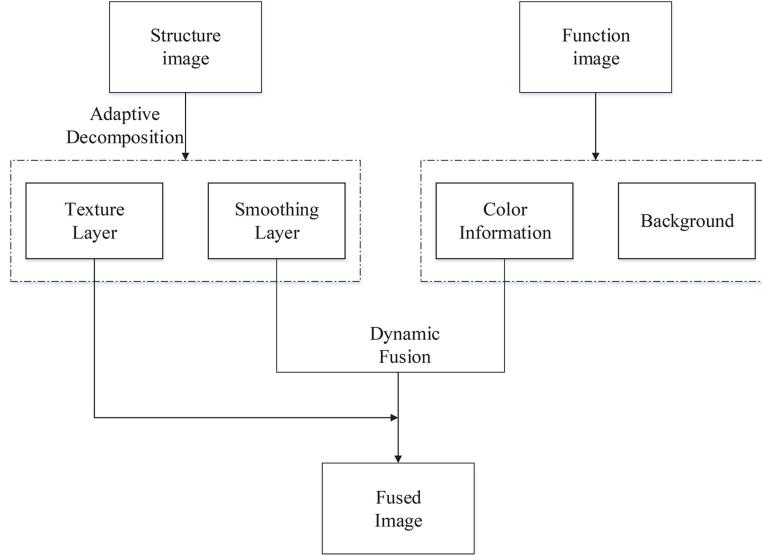


Fig. 2 Fusion framework

$$\alpha(psp) = \frac{p_c(psp) - \min(p_c(psp_n))}{\max(p_c(psp_n)) - \min(p_c(psp_n))} \quad (13)$$

where $p_c(psp)$ represents the grey value of psp in the function image; $\max[p_c(psp_n)]$ represents the maximum grey value in the function image; and $\min[p_c(psp_n)]$ represents the minimum grey value in the function image

$$\beta(psp) = \frac{p_s(psp) - \min(p_s(psp_n))}{\max(p_s(psp_n)) - \min(p_s(psp_n))} \quad (14)$$

where $p_s(psp)$ represents the grey value of psp in the smoothing layer; $\max[p_s(psp_n)]$ represents the maximum grey value in smoothing layer, and $\min[p_s(psp_n)]$ represents the minimum grey value in the smoothing layer

$$w(psp) = \frac{1}{1 + e^{\lambda(\alpha(psp) - \theta\beta(psp))}} \quad (15)$$

where λ and θ are parameters, and here set $\lambda = -10$, $\theta = 0.3$.

Rewrite (9) as

$$w(psp) = \begin{cases} 1 & \text{if } G_f(psp) \geq \tau \\ \frac{1}{1 + e^{\lambda(\alpha(psp) - \theta\beta(psp))}} & \text{if } G_f(psp) < \tau \end{cases} \quad (16)$$

In (16), if $G_f(psp) \geq \tau$, the probability of psp belonging to the colour information area is set to 1; if $G_f(psp) < \tau$, the probability of psp belonging to the colour information area is related to the pixel value of psp in s and C .

The fused image F is obtained by the formula as below:

$$F(psp) = w(psp) \times C(psp) + (1 - w(psp)) \times s(psp) + t(psp) \quad (17)$$

$s = I - t$ can be obtained by (8). Convert (17) to colour space as below:

$$F(psp) = w(psp) \times (C(psp) + [I_g^l, I_g^l, I_g^l](psp)) + (1 - w(psp)) \times I(psp) \quad (18)$$

where $[I_g^l, I_g^l, I_g^l]$ represents the colour version of the texture layer, its red, green and blue three channels are filled by I_g^l .

3.3 Fusion framework

The new medical image fusion framework is shown in Fig. 2. It mainly includes the following four steps:

Step 1: Transform I into grey space, $I_g = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B$, where R, G, B represents the red, green and blue three channels.

Step 2: I_g is decomposed into texture layer t and smoothing layer s using (1).

Step 3: Transform C into grey space, $w(psp)$ can be obtained using (16).

Step 4: The fused image is obtained by using (18).

4 Empirical evaluation

4.1 Experimental settings

The experiments are performed on the Whole Brain Atlas database (<http://www.med.harvard.edu/AANLIB/home.html>). The Whole Brain Atlas 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. In this paper, nine pairs of images were given, of which three were MRI-SPECT, three were MRI-PET, and the other three were CT-SPECT. Each pair of images contains one medical structure image and one medical function image. We compared the performance of our proposed method with six existing image fusion methods including the methods based on LP, discrete wavelet transformation (DWT), NSCT, guided filtering-based fusion (GFF), LP-SR, and WLSs. The codes of above-compared algorithms are from http://xudongkang.weebly.com/uploads/1/6/4/6/16465750/gff_1.0.7z (password: LiVoJeaimnniTgain) and <http://www.escience.cn/people/liuyu1/englishing.html>. The decomposition scale of the source image is four. The parameter is set to the default value in the code.

4.2 Subjective evaluation of experimental results

In our experiments, 30 people with experience in image fusion research are organised to make the subjective evaluation of the fusion image. They are in the same working environment and equipped with the same 23 inch liquid crystal display monitor and the screen resolution is 1920×1080 . Subjects need to enter an integer between 1 and 10 as the score of the image according to the visual quality of each fused image. The larger the value is, the better quality the image is. For each fusion image, the highest and lowest scores are removed, and the average of the remaining scores is taken as the final score for the image.

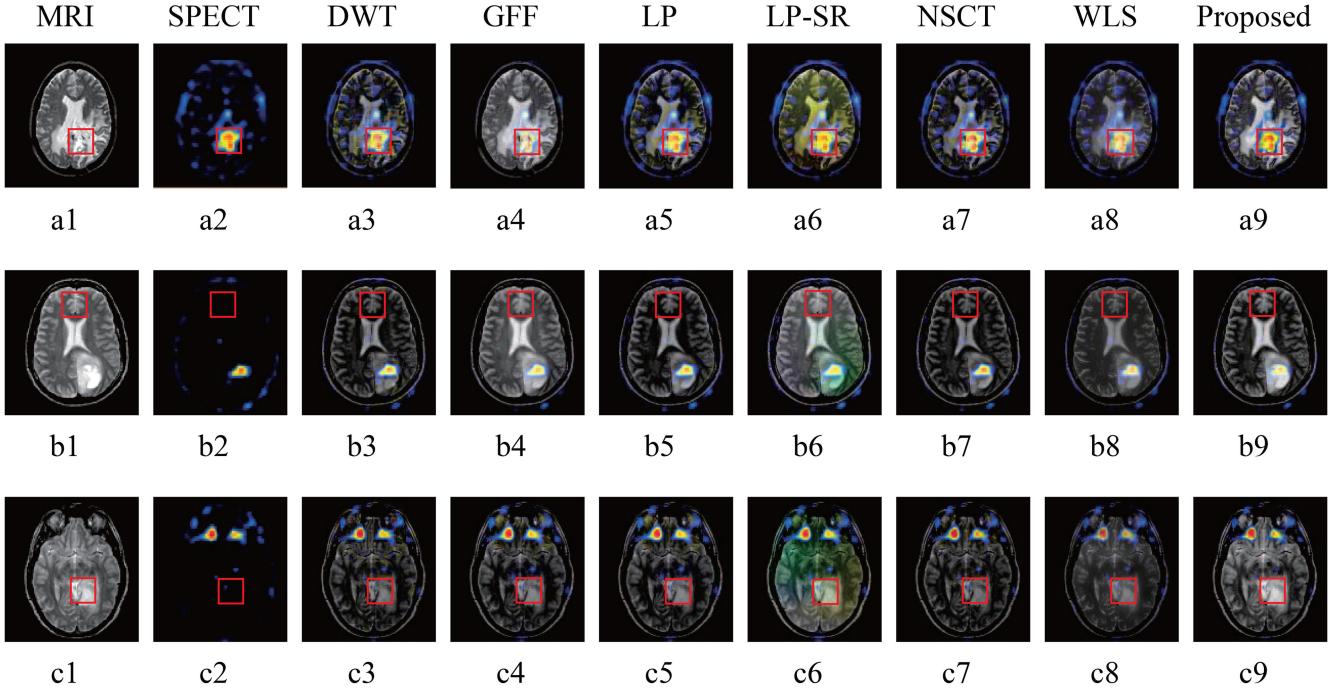


Fig. 3 MRI-SPECT fusion image of seven different algorithms

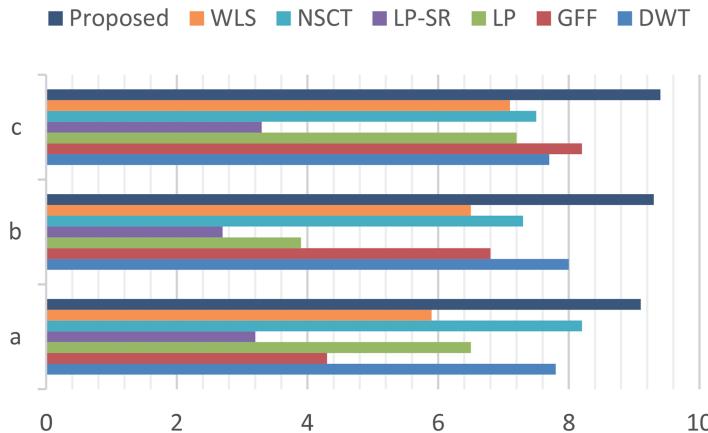


Fig. 4 MRI-SPECT fusion images of subjective scores

The MRI-SPECT fusion images fused by seven different algorithms are shown in Fig. 3. In Fig. 3, the first two columns represent the structure and function images to be fused, and the third column to the ninth column represents the images fused by different methods. We can see that image contrasts of a3, b3 and c3 fused by DWT is reduced. a4, b4 and c4 fused by GFF contain little colour information from function images. a6, b6 and c6 fused by LP-SR generate colours that do not exist in the source image, resulting in significant colour distortion. a5, b5, c5 fused by LP, a7, b7, c7 fused by NSCT and a8, b8, c8 fused by WLS capture little spatial details. a9, b9, c9 fused by the proposed method retain the colour information and spatial details very well. The images fused by the proposed method are of high visual quality. With more careful observation, the red frame shows the details of the fusion results. In pair a, only a7 and a9 remain the rich colour information well from a2, but a9 retains clearer spatial details than a7. In pair b, only b4 and b9 capture the clear spatial details from b1, but b4 lost a lot of colour information. In pair c, only c9 remains the clear spatial details. The corresponding subjective scores are shown in Fig. 4. It can be seen that the images fused by the proposed method got the highest score. The subjective scores further illustrate good performance of the proposed method on MRI-SPECT.

The MRI-PET fusion images of seven different algorithms are shown in Fig. 5. PET image reflects the metabolic situation. Tissue or lesions in the PET with the high metabolic rate shows a bright light signal, and the low metabolic rate shows a dark signal.

Therefore, the remaining different colours and the brightness of the fused image are important. Fig. 5 indicates that images fused by DWT, GFF, LP, NSCT, and WLS remain little colour information from the PET images, while LP-SR and the proposed method do better. In the red frame of pair e, the fused images produce different degrees of colour distortion, except for e9. There is one bright spot in the red frame of f2, but it disappears in f4. The bright spot becomes dark in f3, f5, f7 and f8. Its brightness is preserved better in f6 and f9, but f9 is closer to f2. The corresponding subjective scores are shown in Fig. 6. It can be seen that the images fused by the proposed method got the highest score. This result further illustrates that the proposed method is in high visual quality on MRI-PET.

SPECT-CT and PET-CT fusion images can provide functional metabolic information and anatomical location information of the lesion. Researchers have developed PET-CT equipment, which is the most advanced integration equipments of PET and CT. The device has been put into use, mainly used in the early detection and diagnosis of major diseases such as cancer and brain tumours. Therefore, we only give the fusion images of SPECT and CT in this paper. The CT-SPECT fusion images of seven different algorithms are shown in Fig. 7. g3, h3 and i3 produce significant block effects. g4, h4, i4 and g6, h6, i6 do not keep the colour information of SPECT well, producing significant colour distortion. The spatial details of g7, h7, i7 and g8, h8, i8 are not clear. They do not preserve the high resolution of CT. The fusion

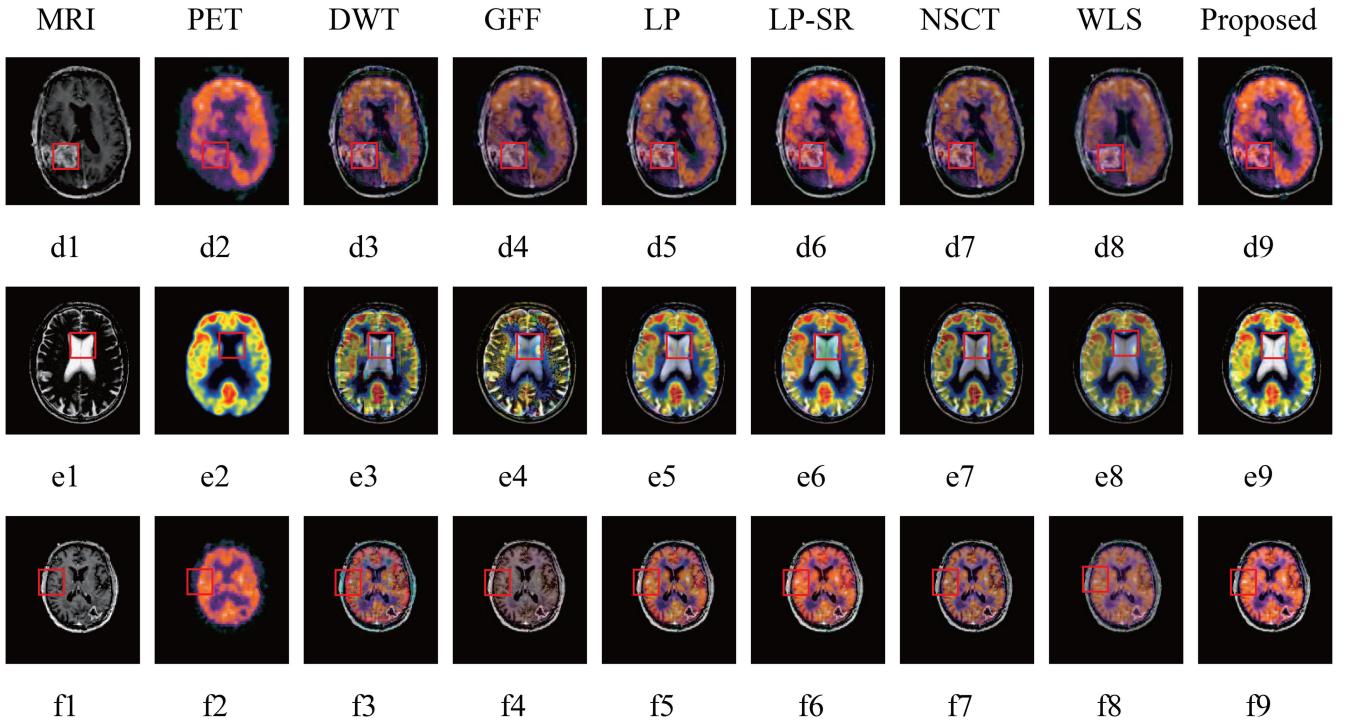


Fig. 5 MRI-PET fusion image of seven different algorithms

■ Proposed ■ WLS ■ NSCT ■ LP-SR ■ LP ■ GFF ■ DWT

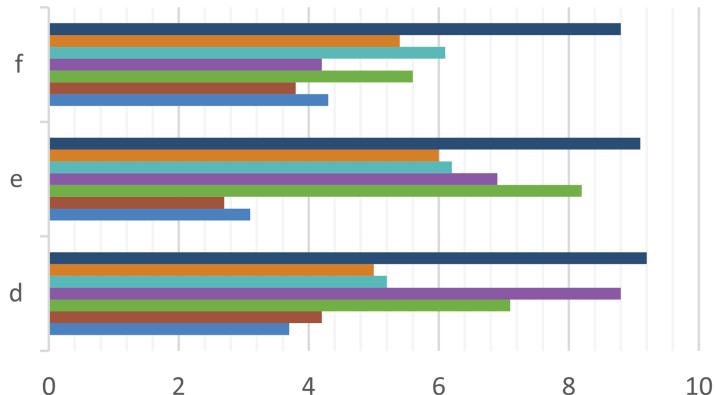


Fig. 6 MRI-PET fusion images of subjective scores

images fused by LP and our proposed methods remain the clear structure information and rich colour information. The corresponding subjective scores are shown in Fig. 8. It can be seen that the images fused by the proposed method got the highest score. This result further illustrates good performance of the proposed method on CT-SPECT.

4.3 Objective fusion quality metrics

Preserving the spectral characteristics is necessary as different colours on function image represent different organisations and the brightness of the colour reflects the activity of the cells. In previous studies, retention of spatial details is often evaluated, while the evaluation of retention of colour information is neglected. In this paper, we evaluate the retention of spatial details and colour information. Objective evaluation metrics include:

- The root mean square error (RMSE) [32], which is the RMS difference between the reference and the fused image.
- The relative average spectral error (RASE) [32], which reflects fusion image spectral quality.
- Spectral angle mapper (SAM) [33], which calculates the angle between the corresponding pixels of the fused and the reference image.

- Mutual information (MI) [34], which reflects the information of the fused image obtained from the source image.
- Quality of visual information (QFAB) [35] is used to evaluate the amount of information that is transferred from the source image to the fused image.
- Q0, weighted function quality index (QW) and edge-dependent fusion quality index (QE) [36] are constructed according to the theory of structural similarity [37, 38], and the similarity between the source image and the fusion image is measured from three aspects: correlation, brightness and contrast.

4.4 Objective evaluation of experimental results

The objective evaluation results of the fusion images in Fig. 3 are shown in Table 1, and the best results are marked in bold. For pair c, the objective evaluation metrics of the proposed algorithm are the best among all. For pair a, QFAB, QW score of GFF are higher than the proposed method. For pair b, Q0 score of GFF is higher. The reason is that GFF preserves the edge information well from the structure. However, the retention of GFF on colour information is the weakest. It almost lost the important colour information of the function image. The proposed method is the best in the three spectral metrics, RMSE, RASE and SAM. In addition, the proposed method achieves the highest score on MI, which indicates

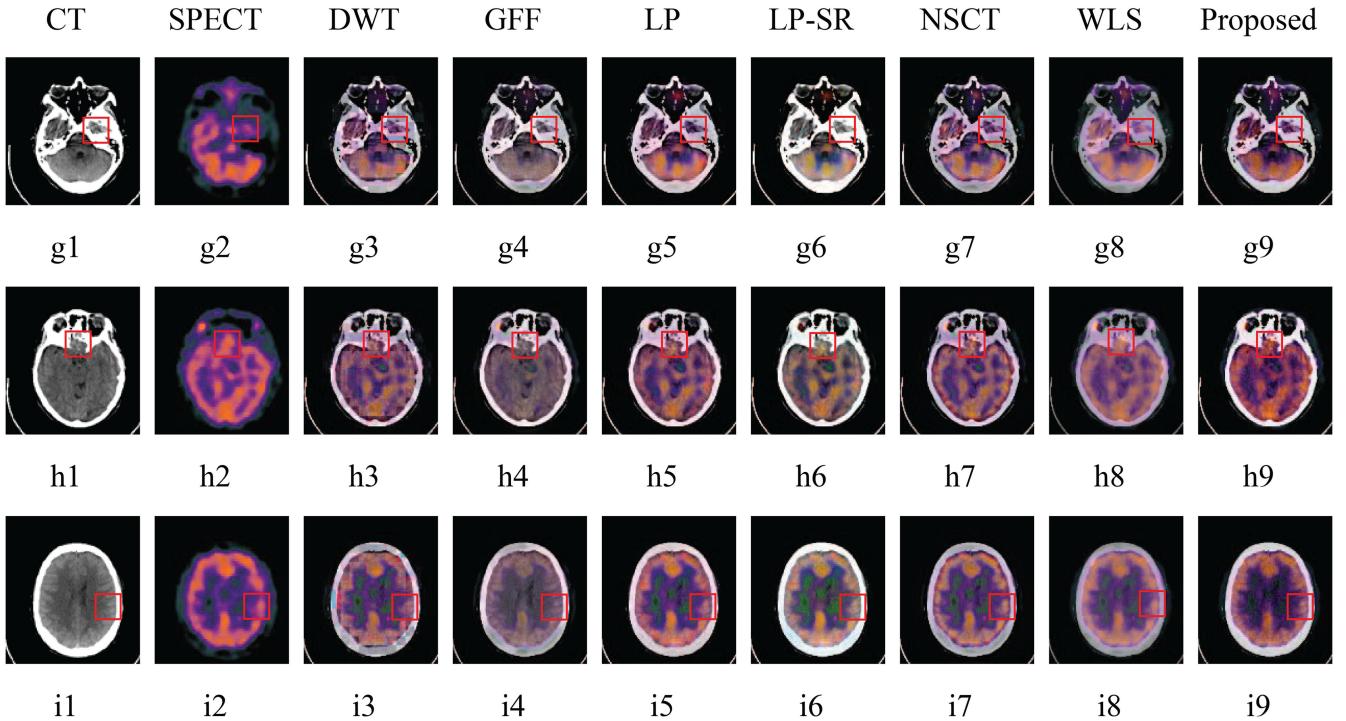


Fig. 7 CT-SPECT fusion image of seven different algorithms

■ Proposed ■ WLS ■ NSCT ■ LP-SR ■ LP ■ GFF ■ DWT

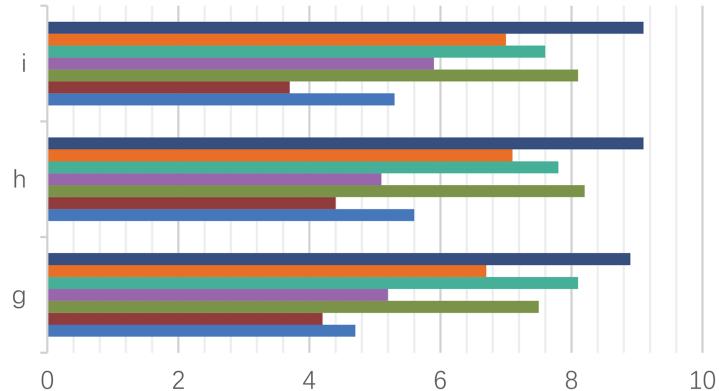


Fig. 8 CT-SPECT fusion images of subjective scores

that it preserves the largest amount of information from the source image. The experimental results indicate that the proposed method performs better on preserving the spatial details from MRI images and the colour information from SPECT images.

The objective evaluation results of the fusion images in Fig. 5 are shown in Table 2, and the best result is marked in bold. For pair d, the objective evaluation metrics of this algorithm are the best among all. In pair e, QFAB score of LP-SR is higher than that of the proposed method. For pair f, QW score of GFF is higher than the proposed method. However, the images e6 and f4 have notable colour distortion. The proposed method is also the best on RMSE, RASE and SAM. The proposed method also achieves the highest score on MI. Experimental results show that the proposed method preserves the spectral information better.

The objective evaluation results of the fusion images in Fig. 7 are shown in Table 3, and the best result is marked in bold. For pair i, the objective evaluation metrics of this algorithm are the best among all. For pair g, QFAB score of NSCT is a little higher than the proposed method. For pair h, h6 performance is better in terms of indicator QE. However, the colour information in the fused image is better preserved by the proposed method. The proposed algorithm has been superior to other algorithms in terms of RMSE, RASE, SAM and MI.

By comparing Fig. 3 and Table 1, Fig. 5 and Table 2 and Fig. 7 and Table 3, subjective evaluation results are sometimes inconsistent with objective evaluation results. Objective evaluation metrics are not sufficient to adequately reflect the quality of the fusion image. The subjective evaluation metrics consider not only the colour and texture information, but also the personal preferences and aesthetic differences. Whether it is a subjective or objective evaluation, the proposed method is superior to other methods.

4.5 Comparison of running time

The average running time of different image fusion algorithms on the above nine pairs of images is shown in Table 4. All the experiments were done on a 2.66 Ghz Xeon central processing unit with 3.25 GB random access memory. All algorithms were implemented in MATLAB R2010b. The source image size is 512×512 . As shown in Table 4, the time overhead of the proposed algorithm is greater than DWT, LP, LP-SR and WLS. To better distinguish between high-frequency and low-frequency informations, the edge detection algorithm is used to determine the boundary of the spectral segmentation on Fourier spectrum and the dynamic fusion coefficient is calculated in the image fusion. These two steps are the major time-consuming steps in the proposed

Table 1 Objective evaluation results of seven algorithms on MRI-SPECT

Set	Metrics	DWT	GFF	LP	LP-SR	NSCT	WLS	Proposed
A	RMSE	39.3960	39.9032	39.2345	40.3827	39.2170	39.9363	39.2075
	RASE	140.3089	143.3423	132.0920	123.2948	137.7638	123.3850	112.1189
	SAM	0.6136	0.5992	0.5981	0.6570	0.5956	0.5957	0.5940
	MI	2.9215	4.0717	2.6638	2.6519	2.6734	3.6022	4.1327
	QFAB	0.6416	0.8149	0.7209	0.7195	0.7180	0.4986	0.8112
	Q0	0.7680	0.8871	0.6899	0.7569	0.7041	0.7867	0.8924
	QW	0.6778	0.9044	0.7214	0.7279	0.7306	0.6831	0.8574
	QE	0.7415	0.8516	0.7869	0.7883	0.7966	0.7307	0.85862
b	RMSE	35.6368	34.3817	35.2144	36.6184	34.6120	36.2816	34.1786
	RASE	147.3261	128.8817	138.2780	170.9329	144.0610	142.0625	116.6881
	SAM	0.5967	0.5739	0.5808	0.6627	0.5772	0.6157	0.5704
	MI	3.6147	4.3154	3.4071	4.1533	3.4000	4.2948	5.8822
	QFAB	0.6643	0.7803	0.7494	0.6957	0.7436	0.5809	0.8130
	Q0	0.7841	0.8957	0.7038	0.7225	0.6958	0.7964	0.8689
	QW	0.7113	0.8536	0.7507	0.6751	0.7543	0.7347	0.8762
	QE	0.7770	0.8543	0.8127	0.7568	0.8200	0.7814	0.8759
c	RMSE	41.5222	41.9729	41.7347	44.4192	41.8488	39.2060	41.1058
	RASE	142.9310	115.6579	135.0847	120.6007	141.0801	127.6795	109.5998
	SAM	0.7471	0.7375	0.7334	0.7661	0.7365	0.7351	0.7322
	MI	3.1741	4.4143	3.0625	3.1875	3.0691	4.1914	5.6144
	QFAB	0.7323	0.8271	0.8120	0.8218	0.7955	0.5480	0.8384
	Q0	0.7991	0.7897	0.8459	0.8751	0.6364	0.7981	0.9068
	QW	0.7445	0.8387	0.8042	0.8304	0.7902	0.7104	0.8802
	QE	0.7928	0.8543	0.8440	0.8616	0.8361	0.7582	0.8642

The best results are marked in bold.

Table 2 Objective evaluation results of seven algorithms on MRI-PET

Set	Metrics	DWT	GFF	LP	LP-SR	NSCT	WLS	Proposed
d	RMSE	31.4456	30.5639	32.1820	34.5207	30.1099	30.2955	30.0510
	RASE	82.5639	75.6018	81.4176	77.6467	79.2771	75.7624	75.2902
	SAM	0.4828	0.4613	0.4791	0.4781	0.4603	0.4712	0.4569
	MI	2.3344	2.4931	2.4323	2.4093	2.3635	2.5164	2.6979
	QFAB	0.6205	0.7348	0.7338	0.7290	0.7101	0.4996	0.7351
	Q0	0.5766	0.5578	0.6398	0.6236	0.5005	0.6205	0.6635
	QW	0.7304	0.8014	0.7865	0.7734	0.7897	0.7012	0.8070
	QE	0.7770	0.8401	0.8403	0.8324	0.8351	0.7226	0.8412
e	RMSE	58.6220	62.1301	61.6092	62.4823	57.9646	54.1873	51.6824
	RASE	131.9837	139.5413	116.5767	113.9637	127.1727	113.5962	112.6263
	SAM	0.6591	0.6089	0.6244	0.6279	0.6274	0.6025	0.6005
	MI	2.0792	2.8781	2.2058	2.2311	2.1603	2.8740	2.8839
	QFAB	0.4830	0.6553	0.6227	0.6038	0.6160	0.4997	0.6352
	Q0	0.6380	0.7784	0.7187	0.7404	0.5871	0.7632	0.7851
	QW	0.6013	0.6736	0.6881	0.6681	0.6949	0.6528	0.7238
	QE	0.6631	0.6353	0.7497	0.7336	0.7527	0.7147	0.7573
f	RMSE	27.7432	27.4181	28.0444	29.3733	27.4275	27.7185	27.4176
	RASE	106.7849	99.7407	106.6224	103.9925	101.9915	99.7748	99.3670
	SAM	0.4899	0.4817	0.4799	0.4826	0.4815	0.4774	0.4748
	MI	2.0728	2.3592	2.1684	2.2183	2.0871	2.0088	2.3800
	QFAB	0.6839	0.8473	0.8003	0.79708	0.7766	0.4969	0.8512
	Q0	0.8516	0.8660	0.9066	0.9066	0.7212	0.8699	0.9095
	QW	0.8087	0.8670	0.8742	0.8652	0.8678	0.7482	0.8771
	QE	0.8434	0.9049	0.8741	0.8834	0.8944	0.8035	0.8851

The best results are marked in bold.

algorithm. However, the time overhead of the proposed algorithm is less than GFF and NSCT. This result also indicates that the proposed algorithm is doable to do almost real-time fusions.

5 Conclusion

MRI/CT structure image can produce clear texture information and SPECT/PET function image can reflect the activity of the cells. MRI-SPECT, MRI-PET, CT-SPECT and CT-PET fusion images should remain both structure and colour informations. This paper proposed a new adaptive decomposition algorithm. One of the attractive features of the decomposition algorithm is that it can automatically distinguish high frequency and low frequency of

Table 3 Objective evaluation results of seven algorithms on CT-SPECT

Set	Metrics	DWT	GFF	LP	LP-SR	NSCT	WLS	Proposed
g	RMSE	46.5883	44.6865	46.0994	46.5278	45.6356	43.5395	42.7706
	RASE	102.3246	91.2690	97.2067	92.3262	100.7501	95.5757	90.0339
	SAM	0.5149	0.4842	0.5035	0.4756	0.4994	0.4603	0.4507
	MI	2.8083	2.8997	2.6901	2.4452	3.6276	2.4921	2.9299
	QFAB	0.8572	0.9245	0.8915	0.8595	0.9452	0.4697	0.9269
	Q0	0.7622	0.7192	0.8916	0.5863	0.8799	0.7211	0.8966
	QW	0.8993	0.9526	0.9485	0.9047	0.9830	0.6986	0.9874
	QE	0.9153	0.9488	0.9533	0.9232	0.9833	0.7868	0.9867
h	RMSE	40.5521	38.2817	39.9318	40.1469	39.2548	37.3108	36.9219
	RASE	82.9390	75.4216	79.5183	72.3125	80.1219	74.1582	72.1255
	SAM	0.4664	0.4394	0.4540	0.4428	0.4492	0.4192	0.4175
	MI	3.0746	3.4182	3.0723	3.5535	2.9020	2.8355	3.5920
	QFAB	0.8189	0.9107	0.8493	0.9002	0.8184	0.4527	0.9283
	Q0	0.7650	0.7511	0.8692	0.8809	0.5900	0.7293	0.8926
	QW	0.8985	0.9625	0.9466	0.9671	0.9120	0.7098	0.9675
	QE	0.9262	0.9595	0.9551	0.9778	0.9368	0.8058	0.9510
i	RMSE	39.8918	37.8623	39.6387	40.9904	38.3080	36.6466	41.2035
	RASE	85.9391	79.3295	79.7111	73.1157	82.4632	77.0349	86.1283
	SAM	0.4657	0.4390	0.4498	0.4428	0.4435	0.4150	0.4680
	MI	2.9700	3.4953	3.0076	3.2663	2.8518	2.9739	3.5099
	QFAB	0.7234	0.8365	0.7639	0.8161	0.7382	0.4278	0.8408
	Q0	0.7280	0.7461	0.8308	0.8390	0.5612	0.7275	0.8395
	QW	0.8158	0.9044	0.9016	0.9274	0.8444	0.6664	0.9328
	QE	0.8758	0.9286	0.9323	0.9567	0.9023	0.7833	0.9544

The best results are marked in bold.

Table 4 Average running time of different image fusion algorithms

Algorithms	DWT	GFF	LP	LP-SR	NSCT	WLS	Proposed
time, s	0.1015	0.1648	0.0036	0.0328	3.0427	0.1069	0.1436

images based on Fourier spectrum analysis. In addition, we investigate its application on structure image and function image in the medical field.

To justify the performance of the proposed fusion algorithm, performance of the proposed method is compared with other six classical image fusion methods on nine pairs of images in our experiments. The proposed algorithm is superior to other algorithms in terms of subjective and objective evaluations. The comparison results show that the proposed method is the best in the three spectral metrics, RMSE, RASE and SAM, and it also makes good performance on other metrics. Experimental results indicate that the proposed method can preserve both the colour information and structure information very well. Extending the proposed method to other areas is our future work such as remote sensing image fusion and industrial precision parts detection. Moreover, we will proceed to grade-distinguished diagnostic images of pathological situations.

6 Acknowledgments

The work was supported by Projects in the National Science & Technology Pillar Program, China (2012BAH48F02), National Science Foundation of China (61272209), Technology Development Plan of Jilin Province (201105017), China Postdoctoral Science Foundation funded project (2017M611323), Natural Science Foundation of Jilin Province (20180101055JC), Outstanding Young Talent Foundation of Jilin Province (20180520029JH) and China Postdoctoral Science Foundation (2017M611323). We would also like to thank <http://www.med.harvard.edu/aanlib/home.html> for providing us the source medical images.

7 References

- [1] James, A.P., Dasarathy, B.V.: ‘Medical image fusion: a survey of the state of the art’, *Inf. Fusion*, 2014, **19**, pp. 4–19
- [2] Feger, S., Rief, M., Zimmermann, E., et al.: ‘Patient satisfaction with coronary CT angiography, myocardial CT perfusion, myocardial perfusion MRI, SPECT myocardial perfusion imaging and conventional coronary angiography’, *Eur. Radiol.*, 2015, **25**, (7), pp. 2115–2124
- [3] Tang, T., Garcia, J., Louie, A.Y.: ‘PET/SPECT/MRI multimodal nanoparticles’, in Bulte, J.W.M., Modo, M.M.J. (Eds.): ‘Design and applications of nanoparticles in biomedical imaging’ (Springer International Publishing, Cham, 2017), pp. 205–228
- [4] Bhatnagar, G., Wu, Q.M.J., Raman, B.: ‘Real time human visual system based framework for image fusion’. Int. Conf. Image and Signal Processing, Berlin, Germany, 2010, pp. 71–78
- [5] Alfano, B., Ciampi, M., De Pietro, G.: ‘A wavelet-based algorithm for multimodal medical image fusion’. Semantic Multimedia, Berlin, Germany, 2007, pp. 117–120
- [6] Yuanyuan, K., Bin, L., Lianfang, T., et al.: ‘Multi-modal medical image fusion based on wavelet transform and texture measure’. IEEE Control Conf., Zhangjiajie, China, 2007, pp. 697–700
- [7] Pritika, , Budhiraja, S.: ‘Multimodal medical image fusion based on guided filtered multi-scale decomposition’, *Int. J. Biomed. Eng. Technol.*, 2016, **20**, (4), pp. 285–301
- [8] Burt, P., Adelson, E.: ‘The Laplacian pyramid as a compact image code’, *IEEE Trans. Commun.*, 1983, **31**, (4), pp. 532–540
- [9] Ravishankar, P., Sharmila, T.S., Rajendran, V.: ‘Acoustic image enhancement using Gaussian and Laplacian pyramid – a multiresolution based technique’, *Multimedia Tools Appl.*, 2018, **77**, (5), pp. 5547–5561
- [10] Paris, S., Hasinoff, S.W., Kautz, J.: ‘Local Laplacian filters: edge-aware image processing with a laplacian pyramid’, *Commun. ACM*, 2015, **58**, (3), pp. 81–91
- [11] Cui, G., Feng, H., Xu, Z., et al.: ‘Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition’, *Opt. Commun.*, 2015, **341**, pp. 199–209
- [12] Liu, Z., Chai, Y., Yin, H., et al.: ‘A novel multi-focus image fusion approach based on image decomposition’, *Inf. Fusion*, 2017, **35**, pp. 102–116
- [13] Daza, R.J.M., Gonzalez, L.C.B., Leon, J.C.M.: ‘Improving the spatial resolution of Ikonos images based on orthophotos: an application of image fusion with wavelet transform’ IEEE 2016 11th Iberian Conf. Information Systems and Technologies (CISTI), Gran Canaria, Canary Islands, Spain, 2016, pp. 1–6

- [14] Singh, R., Khare, A.: ‘Fusion of multimodal medical images using Daubechies complex wavelet transform – a multiresolution approach’, *Inf. Fusion*, 2014, **19**, pp. 49–60
- [15] Cheng, S., He, J., Lv, Z.: ‘Medical image of PET/CT weighted fusion based on wavelet transform’. IEEE the Second Int. Conf. Bioinformatics and Biomedical Engineering, Shanghai, China, 2008, pp. 2523–2525
- [16] Yang, L., Guo, B.L., Ni, W.: ‘Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform’, *Neurocomputing*, 2008, **72**, (1), pp. 203–211
- [17] Huang, H., Feng, X., Jiang, J.: ‘Medical image fusion algorithm based on nonlinear approximation of contourlet transform and regional features’, *J. Electr. Comput. Eng.*, 2017, **2017**, pp. 1–9
- [18] Bhatnagar, G., Wu, Q.M.J., Liu, Z.: ‘Directive contrast based multimodal medical image fusion in NSCT domain’, *IEEE Trans. Multimed.*, 2013, **15**, (5), pp. 1014–1024
- [19] Yang, Y., Que, Y., Huang, S., et al.: ‘Multimodal sensor medical image fusion based on type-2 fuzzy logic in NSCT domain’, *IEEE Sens. J.*, 2016, **16**, (10), pp. 3735–3745
- [20] Bhateja, V., Patel, H., Krishn, A., et al.: ‘Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains’, *IEEE Sens. J.*, 2015, **15**, (12), pp. 6783–6790
- [21] Yin, H., Li, S., Fang, L.: ‘Simultaneous image fusion and super-resolution using sparse representation’, *Inf. Fusion*, 2013, **14**, (3), pp. 229–240
- [22] Yang, B., Li, S.: ‘Pixel-level image fusion with simultaneous orthogonal matching pursuit’, *Inf. Fusion*, 2012, **13**, (1), pp. 10–19
- [23] Liu, Y., Wang, Z.: ‘Simultaneous image fusion and denoising with adaptive sparse representation’, *IET Image Process.*, 2014, **9**, (5), pp. 347–357
- [24] Jiang, Y., Wang, M.: ‘Image fusion using multiscale edge-preserving decomposition based on weighted least squares filter’, *IET Image Process.*, 2014, **8**, (3), pp. 183–190
- [25] Farbman, Z., Fattal, R., Lischinski, D., et al.: ‘Edge-preserving decompositions for multi-scale tone and detail manipulation’, *ACM Trans. Graph. (TOG)*, 2008, **27**, (3), p. 67
- [26] Gilles, J.: ‘Empirical wavelet transform’, *IEEE Trans. Signal Process.*, 2013, **61**, (16), pp. 3999–4010
- [27] Daubechies, I.: ‘Ten lectures on wavelets’ (Society for industrial and applied mathematics, 1992)
- [28] Asano, T., Ishiguro, K., Kondoh, T., et al.: ‘Image quality evaluation for electronic displays based on Fourier spectrum analysis’. IEEE 2016 11th France–Japan & Ninth Europe-Asia Congress on Mechatronics (MECATRONICS)/17th Int. Conf. Research and Education in Mechatronics (REM), 2016, pp. 050–053
- [29] Letourneau, P.D., Langston, M.H., Lethin, R.: ‘IA sparse multi-dimensional fast Fourier transform with stability to noise in the context of image processing and change detection’. 2016 IEEE High Performance Extreme Computing Conf. (HPEC), 2016, pp. 1–6
- [30] Cheng, X., Shao, M., Zhang, H., et al.: ‘Fusion method of visible and infrared images based on calibration information and regional energy optimization’. Fourth Int. Conf. Optical and Photonics Engineering Int. Society for Optics and Photonics, 2017, pp. 102500I–102500I-6
- [31] Otsu, N., Ostu, N., Nobuyuki, O.: ‘A threshold selection method from gray-level histograms. New Afr. 9:62–66’, *IEEE Trans. Syst. Man Cybern.*, 1979, **9**, (1), pp. 62–66
- [32] Jagalingam, P., Hegde, A.V.: ‘A review of quality metrics for fused image’, *Aquatic Procedia*, 2015, **4**, pp. 133–142
- [33] Yuhas, R.H., Goetz, A.F.H., Boardman, J.W.: ‘Discrimination among semi-arid landscape end members using the spectral angle mapper (SAM) algorithm’, 1992
- [34] Petrović, V., Xydeas, C.: ‘Objective evaluation of signal-level image fusion performance’, *Opt. Eng.*, 2005, **44**, (8), pp. 087003–087003-8
- [35] Xydeas, C.S., Petrović, V.: ‘Objective image fusion performance measure’, *Electron. Lett.*, 2000, **36**, (4), pp. 308–309
- [36] Piella, G., Heijmans, H.: ‘A new quality metric for image fusion’. 2003 Int. Conf. Image Processing 2003 ICIP 2003 Proc. IEEE, 2003, vol. **3**, p. II-173
- [37] Prashanth, H.S., Shashidhara, H.L., KN, B.M.: ‘Image scaling comparison using universal image quality index’. IEEE Int. Conf. Advances in Computing, Control, & Telecommunication Technologies 2009 ACT09, 2009, pp. 859–863
- [38] Wang, Z., Bovik, A.C.: ‘Modern image quality assessment’, *Synth. Lect. Image Video Multimed. Process.*, 2006, **2**, (1), pp. 1–156