Knowledge-Based Systems 113 (2016) 4–12

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Medical image fusion by combining parallel features on multi-scale local extrema scheme

Jiao Du, Weisheng Li*, Bin Xiao, Qamar Nawaz

Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

ARTICLE INFO

Article history:
Received 5 April 2016
Revised 7 September 2016
Accepted 12 September 2016
Available online 13 September 2016

Keywords:
Local extrema scheme
Parallel features
Edge saliency weighted map
Color saliency weighted map
MRI-CT fusion
MRI-PET fusion

ABSTRACT

Two efficient image fusion algorithms are proposed for constructing a fused image through combining parallel features on multi-scale local extrema scheme. Firstly, the source image is decomposed into a series of smoothed and detailed images at different scales by local extrema scheme. Secondly, the parallel features of edge and color are extracted to get the saliency maps. The edge saliency weighted map aims to preserve the structural information using Canny edge detection operator; Meanwhile, the color saliency weighted map works for extracting the color and luminance information by context-aware operator. Thirdly, the average and weighted average schemes are used as the fusion rules for grouping the coefficients of weighted maps obtained from smoothed and detailed images. Finally, the fused image is reconstructed by the fused smoothed and the fused detailed images. Experimental results demonstrate that the proposed algorithms show the best performances among the other fusion methods in the domain of MRI-CT and MRI-PET fusion.

1. Introduction

Medical image fusion is a task to obtain a single fused image in terms of human visual perception in order to increase the clinical applicability of medical images for diagnosis and assessment of medical problems [1]. Different breakdown frameworks can be found in the literature based on morphology operator [2], knowledge [3–4], artificial neural networks [5–6], and multi-scale analysis (MSA) [7–22].

Instead of working with a single scale [2–6], the fusion methods based on MSA [7–22] are widely used for fusing multi-modal medical image on account of its properties of extracting and combining saliency feature at different scales. Pyramid transform based methods are well-known MSA schemes in medical image fusion [7–10]. They are formed as a difference between successive levels of Gaussian pyramid with filters, such as Laplacian and gradient filters. However, pyramid transform based fusion methods, such as gradient pyramid (GRP) [7], lack of the direction information. Later, wavelet transforms, such as discrete wavelet transform (DWT) [11], contourlet transform (COT) [12] and shearlet transform (ST) [13–15], provide a framework that the source image is decomposed into a series of low-pass and high-pass sub-images at different scales and directions. The limitations and problems of wavelet transforms are: (1) Wavelet-based methods are with the higher computational complexity; (2) The fused images are blurred with low-contrast; And (3) the saliency features are sensitive to shift and noise. Since human beings are the ultimate receivers, it is very necessary for researchers to investigate the human visual system (HVS) in the field of medical image fusion. HVS which simulates the perceptual process a human performs is applied to medical image fusion and then achieved excellent performance [16–19]. In contrast to wavelet transform based fusion methods, HVS based methods aim to decompose the input image into its multi-scale representation to perform exactly the way the optical cells of human being does in spatial domain. The problem of HVS based fusion methods is the requirement of appropriate parameter values for the psychological model. In order to overcome the computational complexity and plenty of parameterization in wavelet transform and HVS based methods, edge-preserving filtering such as bilateral filter [20], guided filer [21], and local extrema [22–23] can separate the input image into smoothed layers and detailed layers efficiently. The input image after edge-preserving smoothing operation is smoothed layers. Then, detailed layers are constructed by the difference between the input image and smoothed layer. However, these methods have only been used to fuse gray scale images rather than color images.

Although many advanced image fusion methods have been proposed in application of multi-modal medical image fusion, there still exists large room for improvement. In this study, we introduce new MSA image fusion methods with parallel features: edge saliency feature (ESF) [24] and color saliency feature (CSF)

* Corresponding author.
E-mail address: liws@cqupt.edu.cn (W. Li).

http://dx.doi.org/10.1016/j.knosys.2016.09.008
0950-7051/

[25] based on local extrema scheme (LES) [23]. There are many different interpretations of the scale in MSA framework [26]. In this study, we assume that the scale of LES is related to the window size of local regions, the scale of ESF is related to the window size of edge detection filter, and the scale of CSF is related to the ratio of color information to position information. In the proposed methods, the first problem to be solved is to determine the MSA decomposition tool. In this study, we chose LES since it defines smoothed layer as mean values of local minima and maxima envelopes rather than to solve a complex optimization problem. The second problem to be solved in the proposed methods is the image fusion rule. Image fusion rule process [27] guarantees that the salient feature information from images is combined into a single image. Each image modal owns its specific properties. For example, structural image provides structural information, and functional image provides more activity of the tissues at the molecular level. Maybe it is unfair to use a single feature extracted from input images. To preserve salient feature information from different medical images, the proposed methods is to generate parallel features [24-25] to explore the properties. The motivation of parallel features is to combine the structural information using edge saliency features (ESF) method [24] and the functional information using color saliency feature (CSF) method [25]. Experimental results show that the proposed methods give an improved performance compared with state-of-the-art multi-modal medical image fusion approaches. The main contributions of the proposed image fusion methods are highlighted in the following.

1) Edge saliency features (ESF) method [24] is proposed as the fusion rule to combine the Canny edge information from the smoothed layer of structural medical images (such as MRI, CT modal).
2) For the detailed layer of functional medical images (such as PET modal), color saliency feature (CSF) [25] method is used as the fusion rule to combine prominent color information.

The rest of this paper is organized as follows. Firstly, it introduces the related work about the multi-scale decomposition analysis tool and the fusion rules with the parallel features of edge and color. Secondly, in Section 3, two algorithms are presented for MRI-CT and MRI-PET image fusions. Then, Section 4 illustrates experimental evaluation. Finally, Section 5 gives the conclusion.

2. Related work

2.1. Local extrema scheme

Compared to the single-scale based methods, MSA has the advantage of extracting and combining image features at different scales. LES is one of the edge-preserving smoothing operators. The input image I can be transformed into two components: smoothed image S and detailed image D. Smoothed image S contains the coarse part of the source image. And detailed image D refers to the texture, luminance and edge information. S and D can be calculated using LES which consists of three steps: (1) To discover the local minima and local maxima in the given image I within a sliding window $w \times w$; (2) To compute minimal extremal envelope and maxima extremal envelope; And (3) to obtain smoothed image S as the medial plane of the local maximal and minimal envelops. In addition, detailed image D can be obtained by subtracting smoothed image from the source image defined as $D=I-S$.

We use LES to construct a multi-scale image representation scheme of the input image I in this study. After getting L recursive smoothing operations using LES, a single input image I is transferred into a series of smoothed images S_l and a series of detailed images D_l at increasing scales of the window size $w=2l-1(l=1, \cdots, L)$,

$$\mathrm{LES}(I) = \sum_{l=1}^{L} (S_l + D_l) \tag{1}$$

where L is the maximum level. In our experiments we choose $l=2$, $l=3$, and $l=4$ as the sizes of the local window for the multi-scale decomposition scheme.

2.2. Image fusion rules

Fusion rules refer to algorithms that seek to highlight the interested features in images and restrain the features of insignificance. The main contributions of the rules are the combination with multiple original images into a single image. An effective fusion rule plays a significant role in affecting the objective quality assessments of the fused image. A fusion rule normally includes three components: activity-level measurement, coefficient grouping and coefficient combination.

(1) Activity-level measurement: The activity-level scheme measures the saliency of each coefficient at different scales. ESF [24] and CSF [25] are selected as the activity-level measurements in this study.
The ESF uses the Canny edge detector to extract the edge and margin information from the two-dimension (2D) images by scale multiplication [24]. The scale multiplication function is used as the responses of the edge detection filter at different scales. Firstly, two filters for detecting edges in vertical and horizontal directions are constructed:

$$f_s^x(x,y) = -xe^{-x^2/(2s^2)}/s^2,\ f_s^y(x,y) = -ye^{-y^2/(2s^2)}/s^2 \tag{2}$$

where s is the scale of the input image $I(x, y)$. Secondly, the responses to the edge detection filters $f_s^x(x,y)$, $f_s^y(x,y)$ at two scales s_1, s_2 are $H_{s_1}^x(x,y), H_{s_2}^x(x,y), H_{s_1}^y(x,y), H_{s_2}^y(x,y)$. Thirdly, the scale multiplication of the responses is:

$$P_I^x(x,y) = H_{s_1}^x(x,y) \cdot H_{s_2}^x(x,y);\ P_I^y(x,y) = H_{s_1}^y(x,y) \cdot H_{s_2}^y(x,y) \tag{3}$$

where s_1 is a constant, $s_2=2s_1$. Fourthly, edge points are obtained by computation of the local maximal of $M_I(x, y)$ in the direction of $A_I(x, y)$:

$$M_I(x,y) = \sqrt{P_I^x(x,y) + P_I^y(x,y)} \tag{4}$$

$$A_I(x,y) = \arctan\left(\frac{\mathrm{sgn}(H_y^{s_1}(x,y)) \cdot \sqrt{P_I^y(x,y)}}{\mathrm{sgn}(H_x^{s_1}(x,y)) \cdot \sqrt{P_I^x(x,y)}}\right) \tag{5}$$

Finally, the multi-scale product is denosing through a threshold T

$$\mathrm{ESF}(I) = M_I(x,y) - T \tag{6}$$

The CSF aims at detecting the saliency feature which contains both the prominent object and some of background that convey the context [25]. The CSF measurement is driven by low-level saliency features, such as color, orientation and intensity. The CSF follows four principles: (1) Local low-level feature is of the considerations. (2) Global feature maintains the feature deviated from the norm. (3) The approach should consist with human visual organization rules. And (4) the saliency regions are apt to center. The CSF algorithm contains local-global single-scale saliency, multi-scale saliency and center prior. Firstly, the single-scale saliency F_i^I of pixel i at scale r is obtained as:

$$s_i^r = 1 - \exp\left\{-\frac{1}{K}\sum_{k=1}^{K} d\left(q_i^r, q_k^r\right)\right\} \tag{7}$$

where K is the sum of image patches p. And the dissimilarity measurement between a pair of patches $d(q_i, q_k)$ is defined as:

$$d(q_i, q_k) = \frac{d_{color}(q_i, q_k)}{1 + c \cdot d_{position}(q_i, q_k)} \tag{8}$$

where $d_{color}(q_i, q_k)$ is the Euclidean distance between the patches p_i, q_k in Lab color space, $d_{position}(q_i, q_k)$ is the Euclidean distance between the positions of q_i, q_k. Secondly, the multi-scale saliency is constructed by decreasing the saliency of pixels from the background to improve the contrast between saliency and non-saliency region as:

$$\hat{f}_i = \frac{1}{M} \sum_{r \in R} s_i^r (1 - d_{foci}^r(i)) \tag{9}$$

where $R=\{r_1,\ldots, r_M\}$ denotes the different patch sizes r of pixels, and $d_{foci}^r(i)$ denotes the Euclidean positional distance between pixel i and the closets focus of attentions pixel at scale r. Thirdly, CSF is obtained by incorporating a center prior to the multi-scale saliency:

$$\mathrm{CSF}(I) = \hat{f}_i G_i \tag{10}$$

where $G_i(\sigma_x, \sigma_y)$ is a 2D Gaussian function positioned at the center of the images.

(2) Coefficient grouping: The coefficient grouping scheme roughly contains no-grouping (NG), single-scale grouping (SG) and multi-scale grouping (MG) [27]. NG denotes that each coefficient is fused independently. And SG scheme means that the corresponding coefficients at the same scale are fused by the same strategy. In this study, MG scheme is selected for grouping the corresponding coefficients at different scales.

(3) Coefficient combination: Two common combination schemes are the average (AVG) and the weighted average (WA). In AVG, the average value function is applied in the corresponding coefficients. In WA, the combined coefficients are obtained by a linear weighted function. The variance of luminance image (VLI) and entropy of image (EI) functions are selected as our weighed function.

VLI evaluates the variance of luminance in the input image that is the expectation of the input image I. It could be obtained by applying the following expression:

$$\mathrm{VLI}(I) = |L(x, y) - \bar{L}(x, y)|^2 \tag{11}$$

where luminance $L(x, y)$ is computed by averaging the red, green, blue (RGB) channels. And $\bar{L}(x, y)$ is computed by $\sqrt{1/3[(R-L)^2 + (G-L)^2 + (B-L)^2]}$.

EI measures the entropy information of the image. Based on Shannon's Entropy Theory, EI is directly computed according to the image signal probability p_i as follows:

$$\mathrm{EI}(I) = -\sum p_i \log_2(p_i) \tag{12}$$

3. Image fusion methods

The main process of Algorithms 1 and 2 proposed for MRI-CT and MRI-PET images fusion are summarized in Figs. 1-2. Let A and B are two input images with MRI modal and CT modal used in Algorithm 1, and $\bar{A}$ and $\bar{B}$ are two input images with MRI modal and PET modal used in Algorithm 2. Firstly, LES is utilized to get the multi-scale representations of two input images. Secondly, since MRI and CT modals belong to the structural medical image, PET modal belongs to the functional medical image, two different fusion rules based on CSF and ESF functions are adopted to combine the decomposed images, respectively. Finally, the fused image is reconstructed by the fused smoothed image and the fused detailed image.

3.1. Algorithm 1 of MRI-CT images fusion

Algorithm 1 proposed in this study is for fusing MRI and CT images and it contains three steps. From Fig. 1, the two input images A and B are firstly transformed into a series of smoothed images $S_l^A, D_l^A, l \in \{1, \cdots, L\}$ and detailed images $S_l^B, D_l^B, l \in \{1, \cdots, L\}$ at different scales based on LES scheme by Eq. (1).

Secondly, the specific fusion rules are applied to combine smoothed and detailed images of MRI and CT modals, respectively. The fusion rules for the images are illustrated in detail as follows:

(1) ESF+MG+WA+EI based fusion rule is adopted to combine the structural information of MRI and CT images. The fused smoothed images of the two input images are obtained as follows:

$$S_F^A = \frac{\sum_{l=1}^L \mathrm{EI}(E_l^A) \cdot E_l^A}{\sum_{l=1}^L \mathrm{EI}(E_l^A)}, S_F^B = \frac{\sum_{l=1}^L \mathrm{EI}(E_l^B) \cdot E_l^B}{\sum_{l=1}^L \mathrm{EI}(E_l^B)} \tag{13}$$

where E_l^A, E_l^B denote the responses of smoothed images S_l^A, S_l^B in the ESF function by Eqs. (2)–(6) and EI function by Eq. (12) is the weighted function used in WA scheme.

(2) MG+AVG based fusion rule is adopted to combine the detailed image D_l^A

$$D_F^A = \frac{1}{L} \sum_{l=1}^L D_l^A \tag{14}$$

(3) MG+WA+VLI based image fusion is adopted to combine detailed images of the CT image B owing to the variation of luminance in CT image

$$D_F^B = \frac{\sum_{l=1}^L \mathrm{VLI}(D_l^B) \cdot D_l^B}{\sum_{l=1}^L \mathrm{VLI}(D_l^B)} \tag{15}$$

Finally, the fused image F is computed by the addition of the fused smoothed image S_F and the fused detailed image D_F

$$F = S_F + D_F \tag{16}$$

where the fused smoothed image and detailed image is obtained by,

$$S_F = (S_F^A + S_F^B)/2, D_F = (D_F^A + D_F^B)/2 \tag{17}$$

3.2. Algorithm 2 of MRI-PET images fusion

Algorithm 2 consists of three main parts: multi-scale image decomposition, image fusion rules and image reconstruction in Fig. 3. In the first part of Algorithm 2, the original images $\bar{A}$ and $\bar{B}$ are decomposed into multi-scale representations of smoothed images $S_l^{\bar{A}}, D_l^A, l \in \{1, \cdots, L\}$ and detailed images $S_l^{\bar{B}}, D_l^{\bar{B}}, l \in \{1, \cdots, L\}$ based on LES by Eq. (1).

Secondly, image fusion rules are used to combine smoothed images and detailed images at each decomposition level. The fusion rules are constructed as follows.

(1) ESF+MG+WA+EI based fusion rule is applied in combining the fused smoothed image of the input image $\bar{A}$. In order to combine the structural information of the original images, ESF function is selected as the activity-level measurement. The fused image $S_F^{\bar{A}}$ can be easily obtained as follows:

$$S_F^{\bar{A}} = \frac{\sum_{l=1}^L \mathrm{EI}(E_l^{\bar{A}}) \cdot E_l^{\bar{A}}}{\sum_{l=1}^L \mathrm{EI}(E_l^{\bar{A}})} \tag{18}$$

where $E_l^{\bar{A}}$ *is* equal to the response of the smoothed images $S_l^{\bar{A}}$ in the ESF function by Eqs. (2–6) and EI function by Eq. (12) is used as the weighted function in WA scheme.

Algorithm 1 Basic steps of the MRI-CT fusion scheme.

input: MRI image A, CT image B
output: fused image F
Step 1: Multi-scale image decomposition
1. Apply L-level LES algorithm on decomposing A and B into smoothed image and detailed image at various levels (S_i^A, D_i^A), (S_i^B, D_i^B) with $i=1, \cdots, L$ (Eq. (1))

Step 2: Image fusion rules
1. Apply ESF+MG+WA+EI fusion rule on S_i^A to get S_F^A, and MG+AVG fusion rules on D_i^A to get D_F^A (Eqs. (13) and (14))
2. Apply ESF+MG+WA+EI fusion rule on S_i^B to get S_F^B, and MG+WA+VLI fusion rules on D_i^B to get D_F^B (Eqs. (13) and (15))

Step 3: Image reconstruction
1. Estimate the fused smoothed image: $S_F = (S_F^A + S_F^{B)})/2$
2. Estimate the fused detailed image: $D_F = (D_F^A + D_F^{B)})/2$
3. Obtain the fused image: $F = S_F + D_F$

Algorithm 2 Basic steps of the MRI-PET fusion scheme.

input: MRI image $\bar{A}$, PET image $\bar{B}$
output: fused image $\bar{F}$
Step 1: Multi-scale image decomposition
1. Apply L-level LES algorithm on decomposing $\bar{A}$ and $\bar{B}$ into smoothed image and detailed image at various levels $(S_i^{\bar{A}}, D_i^{\bar{A}})$, $(S_i^{\bar{B}}, D_i^{\bar{B}})$ with $i=1, \cdots, L$ (Eq. (1))

Step 2: Image fusion rules
1. Apply ESF+MG+WA+EI fusion rule on $S_i^{\bar{A}}$ to get $S_{\bar{F}}^{\bar{A}}$, and MG+AVG fusion rules on $D_i^{\bar{A}}$ to get $D_{\bar{F}}^{\bar{A}}$ (Eqs. (18) and (19))
2. Apply MG+AVG fusion rule on $S_i^{\bar{B}}$ to get $S_{\bar{F}}^{\bar{B}}$, and CSF+MG+WA+VLI fusion rules on $D_i^{\bar{B}}$ to get $D_{\bar{F}}^{\bar{B}}$ (Eqs. (19) and (20))

Step 3: Image reconstruction
1. Estimate the fused smoothed image: $S_{\bar{F}} = (S_{\bar{F}}^{\bar{A}} + S_{\bar{F}}^{\bar{B})})/2$
2. Estimate the fused detailed image: $D_{\bar{F}} = (D_{\bar{F}}^{\bar{A}} + D_{\bar{F}}^{\bar{B})})/2$
3. Obtain the fused image: $\bar{F} = S_{\bar{F}} + D_{\bar{F}}$

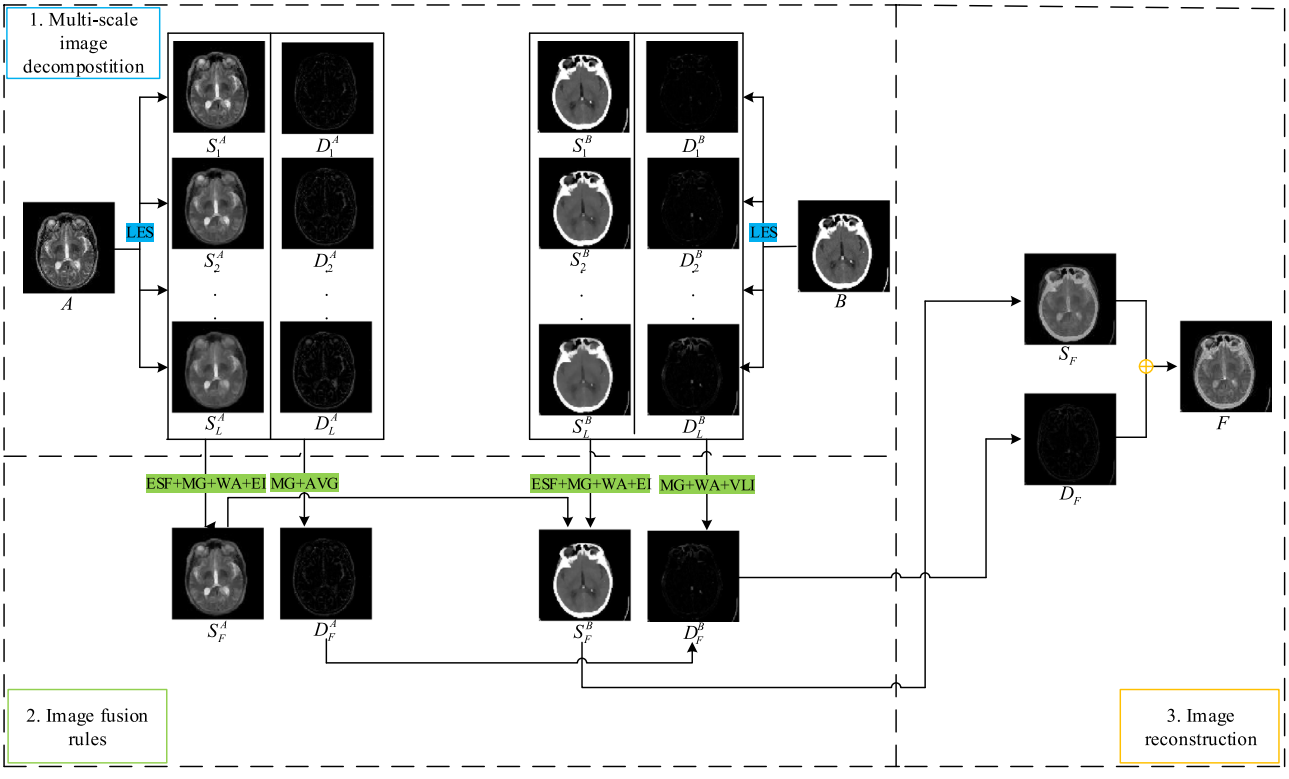

Fig. 1. Schematic diagram of Algorithm 1 of MRI-CT images fusion by the multi-scale analysis tool.

(2) MG+AVG based fusion rule is applied to get the fused detailed image $D_{\bar{F}}^{\bar{A}}$ and the fused smoothed image $S_{\bar{F}}^{\bar{B}}$

$$D_{\bar{F}}^{\bar{A}} = \frac{1}{L}\sum_{l=1}^{L} D_l^{\bar{A}}, S_{\bar{F}}^{\bar{B}} = \frac{1}{L}\sum_{l=1}^{L} S_l^{\bar{B}} \tag{19}$$

(3) CSF+MG+WA+VLI based fusion rule is applied to get the fused detailed image of the input image $\bar{B}$. Due to the properties of the PET image, CSF function is used to extract the color and luminance information from the source image. The fused detailed image $D_{\bar{F}}^{\bar{B}}$ can be easily computed as follows:

$$D_{\bar{F}}^{\bar{B}} = \frac{\sum_{l=1}^{L} \mathrm{VLI}(C_l^{\bar{B}}) \cdot C_l^{\bar{B}}}{\sum_{l=1}^{L} \mathrm{VLI}(C_l^{\bar{B}})} \tag{20}$$

where $C_l^{\bar{B}}$ is equal to the response of detailed images $D_l^{\bar{B}}$ in the CSF function by Eqs. (7–10) and VLI function by Eq. (11) is used as the weight in WA scheme.

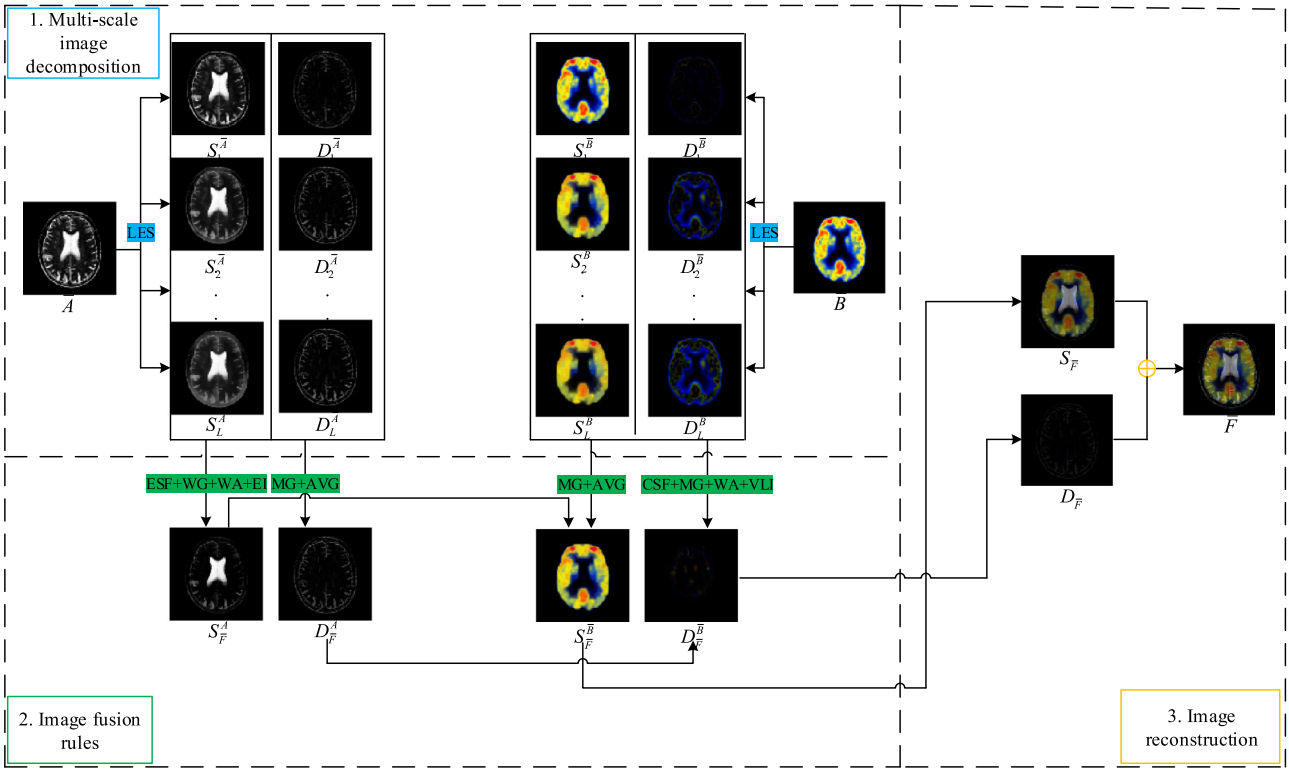

Fig. 2. Schematic diagram of Algorithm 2 of MRI-PET images fusion by the multi-scale analysis tool.

Finally, the fused image $\bar{F}$ is obtained by combining the fused smoothed image $S_{\bar{F}}$ and the fused detailed image $D_{\bar{F}}$

$$\bar{F} = S_{\bar{F}} + D_{\bar{F}} \tag{21}$$

where the fused smoothed image $S_{\bar{F}}$ is the average of the smoothed images of different source images $S_{\bar{F}}^{\bar{A}}$ and $S_{\bar{F}}^{\bar{B}}$, the fused detailed image $D_{\bar{F}}$ is the average of the detailed images $D_{\bar{F}}^{\bar{A}}$ and $D_{\bar{F}}^{\bar{B}}$.

4. Empirical evaluation

4.1. Experimental settings

To evaluate the quality of the proposed methods, we implement experiments on the platform of MATLAB R2010b. The proposed MSA methods consist of three major stages [28] which are (1) multi-scale image decomposition, (2) image fusion rules, and (3) image reconstruction. In the first stage, the input images are decomposed into their multi-scale representations by LES scheme. In the second stage, parallel features based fusion rules are used to combine multi-scale representations. In the third stage, the fused image is constructed by the inversed LES scheme. Their flow-graphs are given in Figs. 1–2 and pseudo-codes are given in Algorithms 1 and 2.

The experiments are performed on the Whole Brain Atlas database (available at [29]). The Whole Brain Atlas [29] 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. In our experiments, 60 pairs of source images are employed to be the testing images for verifying the effectiveness of the proposed methods. Among them, there are 30 pairs of MRI-CT image fusion, and 30 pairs of MRI-PET image fusion. For each pair, the input medical images are assumed to be co-aligned. The compared methods include GRP [7], DWT [11], COT [12], ST [13], SVT [9], ND [10] and LES [22].

4.2. Objective fusion quality metrics

In order to assess the fused images by different methods, six fusion quality metrics are adopted. The following image quality metrics are used to evaluate the performance with and without reference image. For each medical image fusion method, the six quality metrics (structural similarity (*SSIM*) [30], difference of entropy (*DEN*) [13] mutual information (*MI*) [31], a visual saliency-induced index (*VSI*) [32], gradient similarity scheme (*GSM*) [33], and natural image quality evaluator (*NIQE*) [34] metrics), are firstly computed for each experimental data. The readers can refer to these references for more detail information about how to compute them. Then, the total score of the fused results using different fusion methods, *Sum*, is obtained according to its performance on six objective quality metrics

$$Sum = \sum_{t=1}^{T} r_t \tag{22}$$

where r_t refers to the number of algorithms that performs better in terms of the selected objective metric t compared with other

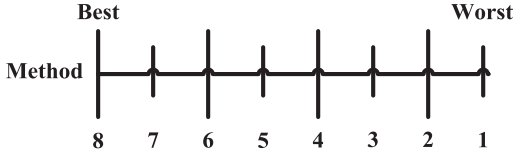

Fig. 3. Scores for the objective image quality metrics.

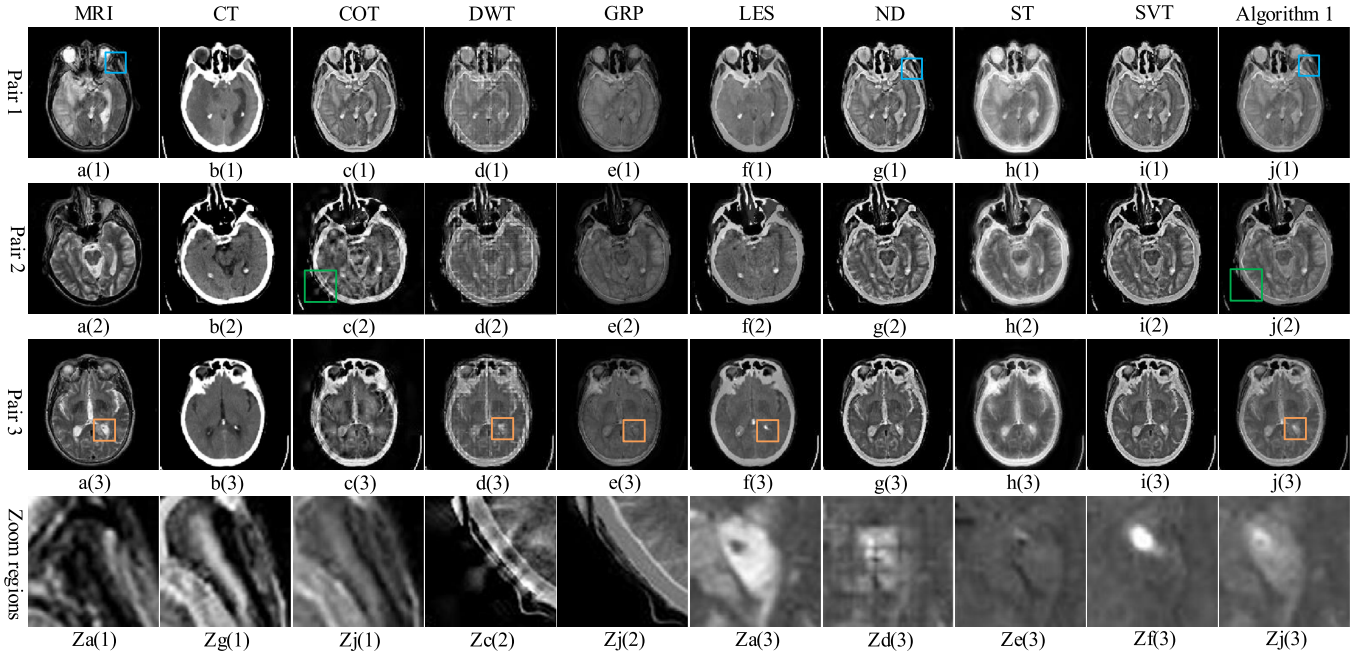

Fig. 4. Fused images from MRI and CT source images by different methods.

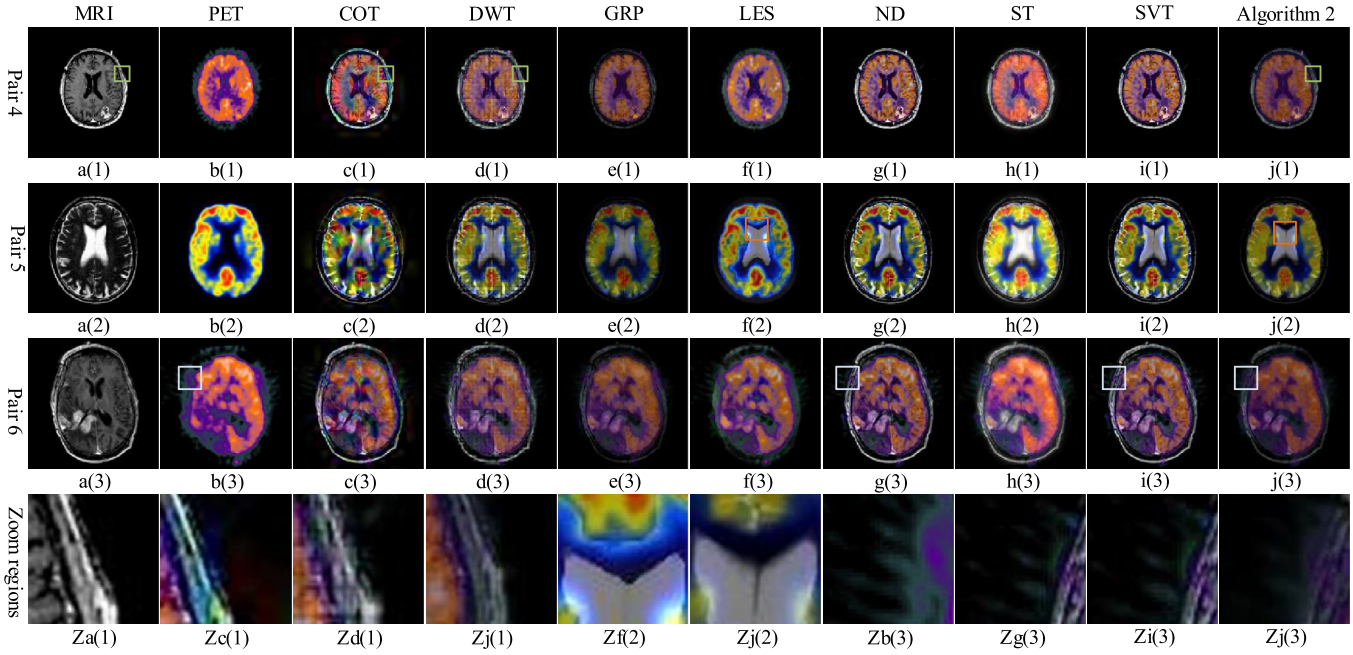

Fig. 5. Fused images from MRI and PET source images by different methods.

fusion methods. As shown in Fig. 3, eight fusion methods are compared in our experimental part. The scores with the range between 1 and 8 are used to distinguish the psychological levels, denoting that the score with the value of 8 is assigned to the method only if its performance is the best and the score with the value of 1 is assigned to the method only if its performance is the worst.

4.3. Subjective evaluation of experimental results

Figs. 4 and 5 display the fused images of three pairs for MRI-CT, MRI-PET image fusion using eight methods, respectively. From Fig. 4, the first three rows display the results of pairs 1–3, and the last row displays the zoom regions of the first three rows. It can be clearly observed that Algorithm 1 can perfectly integrate much more edge information and suppress the color distortion to the greatest extent. COT method introduces serious artifacts along the edges (Fig. 4(c1)-(c3)). DWT method produces block artifacts in the result images (Fig. 4(d1)-(d3)). ST method gets low contrast in the result images (Fig. 4(h1)-(h3)). Furthermore, zoom regions of the result images are shown in Fig. 4z. Fig .4Za(1), Zg(1), and Zj(1) display the zoom regions of a(1), g(1), and j(l) in the first row from which we can observed that ND method get much more edge information from the input MRI image than Algorithm 1. As shown in Fig. 4Zc(2) and Zj(2), Algorithm 1 produces less artifacts along the edge compared to COT method. Furthermore, Algorithm 1 is better at preserving both anatomical and functional information

Table 1
Objective assessments of the experimental results for MRI-CT, where the best results are highlighted in bold.

Image	Metrics	COT	DWT	GRP	LES	ND	ST	SVT	Algorithm 1
Case 1	*SSIM*	7.9368	8.4573	8.4640	8.6052	8.5601	8.3655	8.6248	**8.7715**
	DEN	15.4809	4.6520	**4.3274**	4.3496	4.3316	15.4411	4.3295	4.4303
	MI	19.2274	19.4181	17.8169	21.4236	20.6205	**22.2054**	20.6015	21.5986
	VSI	8.5326	8.7595	8.8248	8.8283	8.7180	8.7243	8.7365	**8.9360**
	GSM	9.8265	9.8799	9.8675	9.8189	9.8745	9.8637	9.8807	**9.8942**
	NIQE	**40.1868**	67.0410	49.8721	53.7500	52.8271	47.4757	53.5075	49.0303
Case 2	*SSIM*	7.2993	7.5619	7.5105	7.7778	7.7920	7.6141	7.8648	**7.9871**
	DEN	14.1155	9.1513	**4.8313**	6.1004	6.8650	19.0452	6.8238	7.7368
	MI	16.7829	17.1755	14.9277	19.9943	18.4757	20.0346	18.6525	**20.2433**
	VSI	8.5118	8.6794	8.5141	8.6460	8.7861	8.5823	8.6091	**8.8723**
	GSM	9.7306	9.8214	9.8352	9.7606	9.7886	9.8063	9.7991	**9.8525**
	NIQE	42.2856	62.6824	42.1967	48.9901	46.6977	**41.7405**	47.6363	46.2418
Case 3	*SSIM*	7.9777	8.4141	8.4637	8.5496	8.5454	8.3681	8.6091	**8.7330**
	DEN	16.4243	6.5514	4.5119	**4.2699**	4.3588	18.1312	4.3472	4.5072
	MI	18.4659	19.6439	17.7244	21.9003	20.8800	21.3232	21.0058	**22.4870**
	VSI	8.5297	8.7386	8.6760	8.8470	8.6837	8.7423	8.6780	**8.8663**
	GSM	9.8210	9.8751	9.8558	9.8038	9.8695	9.8688	9.8776	**9.8886**
	NIQE	**42.7265**	60.2125	48.2254	53.7430	46.5946	47.8748	45.4028	49.7287

Table 2
Objective assessments of the experimental results for MRI-PET.

Image	Metrics	COT	DWT	GRP	LES	ND	ST	SVT	Algorithm 2
Case 4	*SSIM*	8.8224	9.0971	9.1089	9.1749	9.2060	9.0391	**9.2359**	9.2080
	DEN	17.8041	2.9188	2.7485	1.7229	2.0882	10.7900	1.8882	**0.4540**
	MI	15.9030	16.4767	16.9984	17.5653	**17.5935**	17.4137	17.5398	17.1476
	VSI	8.6666	8.7244	8.7371	8.7631	8.7339	**8.8302**	8.7617	8.8225
	GSM	**9.8821**	9.8378	9.8396	9.8389	9.8448	9.8537	9.8502	9.8504
	NIQE	52.4962	78.0409	59.7543	72.3385	56.1247	49.6559	53.7899	**49.2096**
Case 5	*SSIM*	7.4923	7.8560	7.9268	7.8555	8.0695	7.6849	**8.1048**	7.9114
	DEN	17.5519	6.5861	**3.0550**	6.5049	5.0996	16.0159	4.9279	3.5541
	MI	16.3636	17.6265	18.4399	**19.8528**	18.6065	19.2036	18.7759	19.2655
	VSI	8.4934	8.5748	8.5421	8.6169	8.6110	**8.6552**	8.6360	8.5270
	GSM	9.6700	9.6813	9.6705	9.6891	9.6971	9.6977	**9.7041**	9.6804
	NIQE	**46.7610**	79.2734	60.4485	82.1489	52.6458	50.4497	49.2801	52.8956
Case 6	*SSIM*	8.5756	8.8953	8.9005	8.9056	8.9361	8.7099	**8.9700**	8.8962
	DEN	13.7092	4.7713	2.3597	3.9215	3.9934	11.7797	3.8956	**1.9875**
	MI	17.7872	19.8186	20.3629	**22.0930**	20.4095	21.5030	20.5888	21.9195
	VSI	8.8183	8.8847	8.8490	8.8512	8.8818	**8.9070**	8.8996	8.8771
	GSM	9.8065	9.8274	9.8291	9.8277	9.8291	**9.8395**	9.8348	9.8306
	NIQE	80.7528	83.9598	80.7969	78.0511	71.4832	72.7597	**60.8290**	72.2454

from the inputs than DWT, GRP, LES methods in Fig. 4Za(3), Zd(3), Ze(3), Zf(3), and Zj(3). Fig. 5 lists three pairs of MRI-PET fusion results and zoom regions. Due to the properties of different modality medical images, Algorithm 2 has computed the ESF saliency weighted map from MRI image and the CSF saliency weighted map from PET image. By observing the input functional PET images, it can be observed that there exists motion artifacts in Fig. 5b(1) and b(3). Especially, motion artifacts concentrates on the edges of the input PET image in Fig. 5b(3) and its zoom region in Fig. 5Z(b3). In contrast to the other fusion methods, Algorithm 2 can effectively reduce motion artifacts in Fig. 5j(1) and j(3). Besides, by comparing the zoom regions of result images using ND, SVT, and Algorithm 2 methods in Fig. 5Zg(3), Zi(3), and Zj(3), Algorithm 2 not only contains color information from the input PET images but also produces clearly edge information. In pair 4, COT and DWT methods represent color distortion and block artifacts, respectively. Similarly, LES method brings serious color distortion in Fig. 5f(2). The regions in blue and white color are distorted in the result image using LES method by comparing the zoom region of result image in Fig. 5f(2).

4.4. Objective evaluation of experimental results

Objective assessments of the fused images from different methods are conducted using six metrics *SSIM, DEN, MI, VSI, GSM*, and *NIQE*. The bigger the metrics *SSIM, MI, VSI*, and *GSM*, the better the fusion result is. Furthermore, the score of the image for the metric *DEN* and *NIQE* has a value between 0 and 100 (0 presents the best quality, 100 the worst). In Table 1, each value signifies the sum of results using different fusion methods for 10 pairs. We can see that the proposed Algorithm 1 provides the best fusion results by considering the metrics *SSIM, VSI*, and *GSM* in all the clinical cases for MRI-CT. It means that Algorithm 1 can well preserve the structural information, visual saliency information, and gradient information from input images. From the value of the metric *MI*, it can be seen that Algorithm 1 gets the best performance for the result images in case 2 and 3. That is, Algorithm 1 can well preserve the original information of input images with different modalities. However, ST method provides the highest value in terms of *MI* metric in case 1. Table 2 shows the sum value of ten pairs in each case for MRI-PET. It can be easily observed that SVT method wins the first prize in terms of *SSIM* metric in all the cases for MRI-PET. Moreover, ST method gets the largest value in terms of *VSI* metric in all the cases for MRI-PET. By carefully observing the values in Table 2, it can be easily found that SVT method performs better than the other fusion methods because of five metric value in boldface. As for case 4, Algorithm 2 gets the best values in terms of two metrics. However, for case 5 and 6, the proposed Algorithm 2 performs worse compared to ST and SVT methods. Especially, there is no value marked in bold using Algorithm 2 for case 5. Although six metrics for evaluating fusion performance have been illustrated

Table 3
Objective assessments of the experimental results using scoring scheme.

Image		Metrics	COT	DWT	GRP	LES	ND	ST	SVT	Algorithm 1
MRI-CT	Case 1	*Sum*	15	21	28	27	27	25	32	**41**
	Case 2	*Sum*	13	19	31	29	26	27	30	**41**
	Case 3	*Sum*	15	19	23	30	28	24	37	**40**
										Algorithm 2
MRI-PET	Case 4	*Sum*	18	12	21	29	30	31	35	**40**
	Case 5	*Sum*	13	19	25	27	32	31	**41**	28
	Case 6	*Sum*	8	17	23.5	29	31.5	31	**41**	35

in Tables 1 and 2, none of them is definitely better that the others. All of these metrics can only measure the quality of fused images from a limited perspective. Therefore, it is necessary to evaluate the fusion methods using scoring scheme described in Section IV.B. Table 3 gives the scores of fused images using different fusion methods. The bigger the metric *Sum* is, the better the method is. The proposed methods in this study get the best results among 40 pairs of input images. However, as for case 5 and 6 for MRI-PET, SVT method gets the largest value in the terms of *Sum* metric. It can be concluded that the proposed Algorithm 1 is better at fusing MRI and CT image than the other existed methods, including COT, DWT, GRP, LES, ND, ST, and SVT. And for MRI-PET, SVT performs better than the proposed Algorithm 2. In total, the proposed methods: Algorithms1 and 2 gets the highest value of 225 among 60 pairs of the testing medical images.

4.5. Discussions

The proposed methods: Algorithms 1 and 2 present new fusion methods using parallel features. From Tables 1–3, Figs. 4 and 5, the discussion can be illustrated as follows:

(1) Fig. 4, Tables 1, and 3 show the subjective and objective evaluation for MRI-CT image fusion. Both MRI and CT images yield anatomical information of human brain in gray-level. It is interesting to discuss the integration of MRI and CT images using different fusion methods. The edge information in CT image is highlighted with high intensity. MRI image provides information about the tissue type of human brain in high resolution. Serious artifacts can be obviously observed in result images using COT method in Fig. 4c(1)-c(3). But, COT method performs best in the metric *NIQE*. That is, COT method introduces less distortions in natural scene statistic features. Visually, ND, SVT methods, and Algorithm 1 present high clear and contrast result images. Objectively, Algorithm 1 performs best in Tables 1 and 3.
(2) The results of MRI-PET image fusion are shown in Fig. 5, Tables 2, and 3. In contrast to CT image, PET image in pseudo-color is usually influenced by motion artifacts in Fig. 5b(1) and b(3). It can be clearly observed that the result images using COT method introduce color distortion in Fig. 5c(1)-c(3). Accordingly, COT method gets very low values in terms of *SSIM, MI* and *VSI* metrics in Table 2 and Table 3. ND and SVT methods perfectly preserve the structural information form the input MRI image. Additionally, SVT method gets the largest value in terms of *Sum* metric for case 5 and case 6 in Table 3. However, motion artifacts in the PET image do not weak using ND and SVT methods in Fig. 5. Algorithm 2 proposed in this study not only preserve the color information but also reduces motion artifacts in Fig. 5j(1)-j(3). From Table 3, Algorithm 2 gets the first prize for the metric *Sum* for case 4.

5. Conclusion

Two MSA methods for MRI-CT and MRI-PET image fusion are presented in this study based on LES scheme. In the proposed methods, parallel features based fusion rules are used to combine the multi-scale image representations. ESF and CSF saliency maps are firstly used to measure the activity level of the anatomical and functional images to obtain the initial map, and then the final maps are refined with EI, VLI, and AVG. The proposed fusion methods are compared with seven representative methods in terms of both subjective and objective evaluations. Experiments show that Algorithms 1 and 2 can well preserve the structure and luminance information from multiple input images.

The interest of this study is that it demonstrates ESF and CSF features can extract the salient information from anatomical image: MRI and PET, and functional image: PET for respectively. However, this study has some defects. In the future, we will carry out experiments in four aspects. Firstly, we find the feature descriptor for measuring the activity level of SPECT image. Secondly, we can use bio-inspired swarm intelligence algorithms to solve a complex optimization problem [35–36], such as monarch butterfly optimization (MBO), earthworm optimization algorithm (EWA), and elephant herding optimization (EHO). Thirdly, we construct fusion methods which is capable of fusing more than two images with different imaging techniques. Finally, it is necessary to design a new objective image quality metric originated from prior knowledge from the clinical organizations.

Acknowledgement

This work was supported in part by Natural Science Foundation of China (No.61272195, 61201383, 61472055, U1401252), 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. Fus. 19 (2014) 4–19.
[2] S. Marshall, G.K. Matsopoulos, J.N. Brunt, Fusion of MR and CT Images of the human brain using multiresolution morphology, Math. Morphol. Appl. Image Process. (1994) 317–324.
[3] G.L. Rogova, P.C. Stomper, Information fusion approach to microcalcification characterization, Inf. Fus. 3 (2) (2002) 91–102.
[4] N. Walker, J. Fox, Knowledge based interpretation of images: a biomedical perspective, Knowl. Eng. Rev. 2 (4) (1987) 249–264.
[5] Z. Wang, Y. Ma, Medical image fusion using m–PCNN, Inf. Fus. 9 (2) (2008) 176–185.
[6] S. Das, M.K. Kundu, A Neuro-fuzzy approach for medical image fusion, IEEE. Trans. Biomed. Eng. 60 (12) (2013) 3347–3353.
[7] E.H. Adelson, C.H. Anderson, J.R. Bergen, P.J. Burt, J.M. Ogden, Pyramid methods in image processing, RCA Eng. 29 (6) (1984) 33–41.
[8] Z. Liu, K. Tsukada, K. Hanasaki, Y.K. Ho, Y.P. Dai, Image fusion by using steerable pyramid, Pattern Recognit. Lett. 22 (9) (2001) 929–939.
[9] Z. Sheng, Z.S. Wen, J. Liu, G.X. Zhu, J.W. Tian, Multisource image fusion method using support value transform, IEEE Trans. Image Process. 16 (7) (2007) 1831–1839.
[10] H. Zhao, Z Z.Shang, Y.Y. Tang, Multi-focus image fusion based on the neighbor distance, Pattern Recognit. 46 (3) (2013) 1002–1011.
[11] Q. Guihong, Z. Dali, Y. Pingfan, Medical image fusion by wavelet transform modulus maxima, Opt. Express 9 (4) (2001) 184–190.
[12] 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.

[13] Q. Miao, C. Shi, P. Xu, A novel algorithm of image fusion using shearlets, Opt. Commun. 284 (6) (2010) 1540–1547.
[14] 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. Fus. 19 (2013) 29–37.
[15] L. Wang, B. Li, L.F. Tian, Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients, Inf. Fus. 19 (2012) 20–28.
[16] 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) 3390–3479.
[17] G. Bhatnagar, Q.J. Wu, Z. Liu, Human visual system inspired multi-modal medical image fusion framework, Expert Syst. Appl. 40 (5) (2013) 1708–1720.
[18] X. Zhang, X. Li, Y. Feng, Image fusion with internal generative mechanism, Expert Syst. Appl. 42 (5) (2015) 2382–2391.
[19] S. Daneshvar, H. Ghassemian, MRI and PET image fusion by combining IHS and retina-inspired models, Inf. Fus. 11 (2) (2010) 114–123.
[20] J. Hu, S. Li, The multiscale directional bilateral filter and its application to multisensor image fusion, Inf. Fus. 13 (3) (2012) 196–206.
[21] S. Li, X. Kang, J. Hu, Image fusion with guided filtering, IEEE Trans. Image Process. 22 (7) (2013) 2864–2875.
[22] Z. Xu, Medical image fusion using multi-level local extrema, Inf. Fus. 19 (2013) 38–48.
[23] K. Subr, C. Soler, F. Durand, Edge-preserving multiscale image decomposition based on local extrema, ACM Trans. Graph. 28 (5) (2009) 147–155.
[24] P. Bao, D. Zhang, X. Wu, Canny edge detection enhancement by scale multiplication, IEEE Trans. Pattern Anal. Mach. Intell. 27 (9) (2005) 1485–1490.
[25] S. Goferman, L. Zelnik-Manor, A. Tal, Context-aware saliency detection, IEEE Trans. Pattern Anal. Mach. Intell. 34 (10) (2012) 1915–1926.
[26] C. Lopez-Molina, B.D. Baets, H. Bustince, Multiscale edge detection based on Gaussian smoothing and edge tracking, Knowl. Based Syst. 44 (1) (2013) 101–111.
[27] R. Shen, I. Cheng, A. Basu, Cross-scale coefficient selection for volumetric medical image fusion, IEEE. Trans. Biomed. Eng. 60 (4) (2013) 1069–1079.
[28] B. Yang, S. Li, Pixel-level image fusion with simultaneous orthogonal matching pursuit, Inf. Fus. 13 (2012) 10–19.
[29] K.A. Johnson, J.A. Becker, The Whole Brain Altas, 2011 [Online]. Available: http://www.med.harvard.edu/aanlib/.
[30] Z. Wang, A.C. Bovik, H.R. Sheikh, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[31] Q. Xiao-Bo, Y. Jing-Wen, X. Hong-Zhi, Z. Zi-Qian, Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain, Acta Autom. Sin. 34 (12) (2008) 1508–1514.
[32] L. Zhang, Y. Shen, H. Li, VSI: a visual saliency-induced index for perceptual image quality assessment, IEEE Trans. Image Process. 23 (10) (2014) 4270–4281.
[33] L. Anmin, L. Weisi, Image quality assessment based on gradient similarity, IEEE Trans. Image Process. 21 (4) (2012) 1500–1512.
[34] A. Mittal, R. Soundararajan, A.C. Bovik, Making a “completely blind” image quality analyzer, IEEE Sig. Process Lett. 20 (3) (2012) 209–212.
[35] W. Cai-Ge, D. Suash, S.C. Leandro, Elephant herding optimization, International Symposium on Computational and Business Intelligence, 2015.
[36] D. Yongsheng, S. Hongbin, H. Kuangrong, A bio-inspired emergent system for intelligent Web service composition and management, Knowl. Based Syst. 20 (2007) 457–465.