



Robust multi-modal medical image fusion via anisotropic heat diffusion guided low-rank structural analysis

Qingzheng Wang^a, Shuai Li^{a,*}, Hong Qin^b, Aimin Hao^a

^a State Key Laboratory of Virtual Technology and Systems, Beihang University, China

^b Stony Brook University, USA



ARTICLE INFO

Article history:

Received 12 March 2014

Received in revised form 20 November 2014

Accepted 9 January 2015

Available online 2 February 2015

Keywords:

Multi-modal image fusion

Data-specific filter

Anisotropic heat kernel design

Low-rank analysis

Multi-scale decomposition

ABSTRACT

This paper proposes a novel and robust multi-modal medical image fusion method, which is built upon a novel framework comprising multi-scale image decomposition based on anisotropic heat kernel design, scale-aware salient information extraction based on low-rank analysis, and scale-specific fusion rules. Our framework respects multi-scale structure features, while being robust to complex noise perturbation. First, anisotropic heat kernel is computed by constructing an image pyramid and embedding multi-level image properties into 2D manifolds in a divide-and-conquer way, consequently, multi-scale structure-preserving image decomposition can be accommodated. Second, to extract meaningfully scale-aware salient information, we conduct low-rank analysis over the image layer groups obtained in the first step, and employ the low-rank components to form the scale space of the salient features, wherein the underlying noise can be synchronously decoupled in a natural way. Third, to better fuse the complementary salient information extracted from multi-modal images, we design an S-shaped weighting function to fuse the large-scale layers, and employ the maximum selection principle to handle the small-scale layers. Moreover, we have conducted extensive experiments on MRI and PET/SPECT images. The comprehensive and quantitative comparisons with state-of-the-art methods demonstrate the informativeness, accuracy, robustness, and versatility of our novel approach.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction and Motivation

The objective of data fusion is to integrate the inherent complementary information from two or more resource images into a single one, which is more informative and suitable for human visual perception and further processing tasks [1]. In clinical medicine, magnetic resonance imaging (MRI) and computed tomography (CT) mainly reveal the structural and anatomical information, while positron emission tomography (PET) and single photon emission computed tomography (SPECT) concentrate on the functional information. Thus, multi-modal medical image fusion can provide much more anatomical and physiological characteristics, which has been playing a critical role in image analysis, clinical diagnosis, and treatment planning [2].

As an out-of-the-box tool in image processing and computational vision, recent progresses of Multi-scale Image Decomposition (MID) indicate that MID has promising potential in the improvement of image fusion quality, because it can effectively detach the structure-sensitive information from multi-modal

images and re-integrate them to enhance informative perception via frequency-domain representation [3,4] or intensity domain presentation [5,6]. Meanwhile, since the design of MID kernel function has great influence on the fusion quality, to get the built-in attractive properties such as structure-sensitive, scale-meaningful, and shift-invariant, many anisotropic MID kernels also gain great momentum in image fusion, including curvelet [7], contourlet [8], non-subsampled contourlet [9], and shearlet [10].

Despite the growing success of medical image fusion, certain challenges still exist in terms of feature processing, noise tolerance, and complementary-information combination, etc. [2,11], because good image fusion methods must reasonably and adaptively model the intrinsic structure patterns embedded in data regardless of the impact of artifacts. However, current MID-based methods can only extract limited directional detail information, and it remains difficult to preserve the scale-aware directional structures, because its kernel/basis functions are generic while having little relevance with the images to be fused. Meanwhile, high-quality image fusion should synchronously emphasize salient information, suppress inconsequential information, and decouple artifacts/noise from the source images. However, most of the existing methods tend to preserve the large-scale structure information while discarding

* Corresponding author.

E-mail address: ls@vrlab.buaa.edu.cn (S. Li).

small-scale details, and they are usually sensitive to noise. Another nontrivial problem is fusion rule, which largely determines whether the fusion result could have proper contrast to clearly reflect the complementary information of multi-modal images. For example, the popularly-adopted Average-Maximum fusion rule [12,13] tends to produce either globally darkened or obscured fusion results.

To alleviate the aforementioned problems, as highlighted in Fig. 1, we concentrate on high-quality robust multi-modal medical image fusion by incorporating data-specific filter design and structure-preserving multi-scale image decomposition into the powerful scale-aware low-rank analysis model. It exploits anisotropic heat kernel for multi-scale image decomposition, scale-aware low-rank analysis for salient information extraction, and scale-specific rules for complementary information fusion. The intuitions behind anisotropic heat diffusion are: diffusion is a process with strong statistical meanings in the global scope, which can naturally bridge the gap between the local structural attributes and their global correlations via weighted Laplace-Beltrami operator; the prior knowledge that depends on the image structure allows to trigger the anisotropy property, which also gives rise to a direction that the diffusion is propagated along. Therefore, the filter kernel derived from anisotropic diffusion is non-linear and space-varying, which guarantees to preserve the sharp structure in multi-scale convolution. By using the low-rank component to indicate the salient information commonly occurring across similar-scale smoothed images, we can simultaneously decouple the noise from the source image, which avoids tricky processing for various kinds of noises. Meanwhile, instead of directly using the original source images, we construct a scale space by resorting to the extracted low-rank structures for information fusion. Besides, given the complementary information to be fused, fusion rules may have non-trivial influence on the intensity contrast of the fusion result. Specifically, the intensity of SPECT/PET images directly relates to the saliency and activity of the tissue, higher intensity would indicate greater activity and thus should be visually more important. Therefore, to better preserve the proper image contrast in the fused result, the S-shaped function is finally used to map the intensity of SPECT/PET to the fusion weights.

Specifically, the salient contributions of this paper can be summarized as:

- We formulate a data-specific multi-scale geometrical analysis kernel for image decomposition by conducting anisotropic heat diffusion over 2D manifold embedded in the 3D space, which can well respect the anisotropic directional structure of the images.

- We propose a layer-group low-rank analysis method to extract scale-aware salient structures, which naturally gives rise to the robust representation of the cross-scale salient features, and can synchronously decouple types of noise in a self-taught way.
- We design a scale-specific fusion rule to adaptively determine the fusion weights, which affords to better integrate the complementary information of multi-modal medical images.

2. Related work

2.1. Image fusion methods

Along with the increasing demands for image fusion in clinical applications, a number of techniques have been proposed in the last decade. James and Dasarathy introduced the existing fusion techniques in medical imaging, and analyzed those methods' advantages and drawbacks. For more information, please refer to their works [2,11]. Here, we mainly discuss the substitution method and MID-based method.

For the substitution method, the intensity-hue-saturation (IHS) method [14] is generally applied in various data fusion applications, which can well integrate spectral and spatial features. However, it also gives rise to spectral distortions and spectral information missing due to the arithmetic combination. In contrast, the generalized intensity-hue-saturation (GIHS) framework [14] can be used to suppress such distortion, and thus affords a better visual sense. Recently, Daneshvar and Ghassemian [15] integrated the IHS technique and retina-inspired model to avoid the weak points appearing in IHS and GIHS, hence improving the fusion quality.

Meanwhile, the MID-based technique has also gained growing attention, since it can effectively respect the sharp feature. The central idea of such method can be summarized as: MID method is used to decompose image into multi-scale representation and form a scale space; fusion is then executed to integrate the complementary information for each scale; and the fusion result is finally obtained by the inverse transform. The MID-based methods for fusion are generally classified into two categories: frequency domain based method and intensity domain based method.

As for frequency domain based method, Wang et al. [12] introduced the discrete wavelet transform (DWT) into medical image fusion, which makes the wavelet-based fusion method popular. But the fusion result generally tends to present bad visual perception and usually involves artifacts. That is because, DWT only extracts limited directional information. Later, several multi-scale geometric analysis methods are proposed in succession to further improve the visual perception, such as curvelet [16] and contourlet

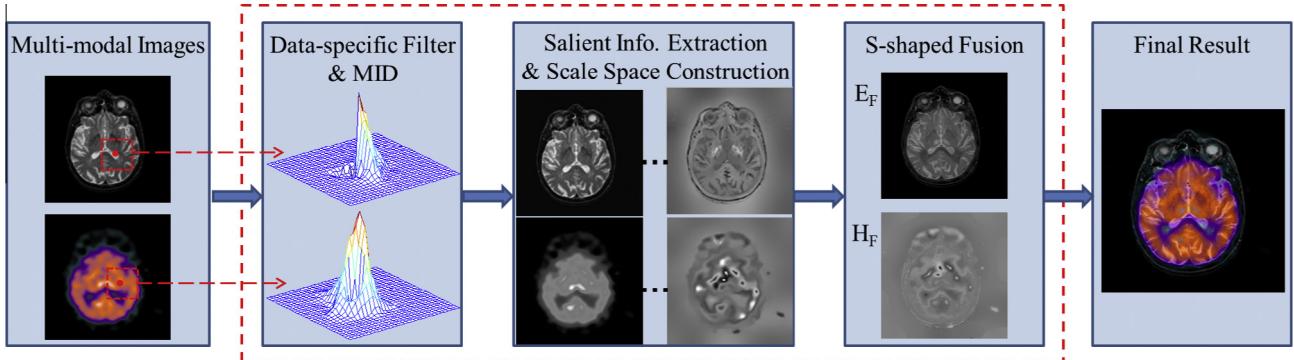


Fig. 1. The flowchart of our method. It consists of three major steps: (1) construct the data-specific filter and conduct multi-scale image decomposition, (2) extract scale-aware salient information via low-rank decomposition, and (3) fuse data with the improved S-shaped fusion rule.

[17]. Although such methods have obtained quite good fusion results, they are still suffering from the lack of shift-invariant property, and the color distortion may still occur. The Non-Subsampling Contourlet Transform (NSCT) [9] and shearlet [18] are then proposed to overcome such limitation and suppress artifacts, whose improved versions further improve the image fusion quality [19–23]. However, the fusion results tend to lose some small details. The main reason is that, their kernel/basis functions are general while having little relevance with specific images, and thus they are hard to faithfully preserve the intrinsic directional structures during image fusion. Therefore, it makes sense to investigate the data-specific filter design.

Apart from frequency domain based approaches, the intensity domain based methods also gain great momentum. For example, Laplacian pyramid [24] is one of the simple and efficient methods that can be applied to multi-focus image fusion [25] and multi-spectral image fusion [26]. However, such methods tend to involve other irrelevant information, such as artifacts and noise. Thus, Hu and Li [5] introduced the directional bilateral filter into image fusion, and the L_0 gradient minimization method is also proposed most recently [6]. The experimental results demonstrate that such data-specific filter based method can make the fused image have more detail information. Relatively speaking, bilateral filter based method usually involves artifacts around the edge, while L_0 method tends to lose small-scale details. That is because, L_0 method, which is a global optimization algorithm, can only preserve the salient detail. Thus, the solving of such problems needs to design a structure sensitive filter that can strongly respect the intrinsic directional structures.

2.2. Low-rank analysis models

Low-rank decomposition is an active research topic in matrix completion [27] and background modeling [28,29]. Recently, low-rank modeling has achieved great success in medical image analysis, such as image denoising and image reconstruction, please refer to Zhou's work [30] for the comprehensive review. In image denoising, Nguyen et al. [31] took a dynamic image sequence as data matrix, and proposed a maximum a posterior (MAP) framework to couple low-rank model to remove noise. By iteratively performing the singular value thresholding (SVT) operation on the data, Candès et al. [32] used the low-rank model to achieve more robust results. In image reconstruction, Christodoulou et al. [33] integrated the low-rank structure into the framework of the partial separability model to reconstruct the cardiac image. Similarly, there are also other low-rank modeling methods for CT and PET image reconstruction [34,35]. Although low-rank modeling has been widely used and achieved great success, fusion application of low-rank modeling is only just a beginning. Most recently, Wan et al. [36] proposed a multi-focus fusion framework based on the low-rank modeling. The method fuses the multi-focus images by choosing the significant features from the sparse component. But the framework is difficult to respect scale-aware, intrinsic anisotropic features. Thus, it motivates us to further explore the integration of low-rank model in medical image fusion.

3. Method overview

In this paper, we respectively denote the SPECT/PET, MRI, and the fused image as \mathbf{C} , \mathbf{G} , and \mathbf{F} , while $\mathbf{I}_{(\cdot)}$ is the intensity component of image “ \cdot ”. The base and detail layers resulted from MID are denoted by \mathbf{E} , \mathbf{D} . We also assume that the source images to be fused have been perfectly co-aligned. We adopt GIHS method [14] to handle the fusion between the gray and the color images. Since color images are commonly regarded as the combination of three

monochrome channels called RGB, they can easily be converted to the IHS color space based on triangular planes.

We give a brief description of the proposed algorithm in **Algorithm 1**: Steps 1–5 generate anisotropic multi-scale image decomposition; Steps 6–8 extract the scale-aware salient feature and intrinsically decouple the underlying noise; and Steps 9–11 conduct the scale-specific fusion based on the extracted multi-scale salient features. All the technical elements will be detailed in the following sections one-by-one.

Algorithm 1. Multi-modal medical image fusion

input: Original images \mathbf{I}_C , \mathbf{I}_G , β , θ , α , d , k , the block size.

output: Fused image \mathbf{F} .

- 1: Divide image pyramid into r overlapping blocks;
 - FOR each block
 - 2: Generate 3D triangular surface;
 - 3: Compute matrices \mathbf{M} and \mathbf{A} ;
 - 4: Solve Λ and Φ that satisfy $\mathbf{M}\Phi = \Lambda\mathbf{A}\Phi$;
 - 5: Conduct multi-scale decomposition;
 - END
 - 6: Group the decomposed multi-scale layers into k groups;
 - 7: Extract salient information via group-wise low-rank analysis;
 - 8: Form a scale space for salient features;
 - 9: Intra-layer fusion with the improved fusion rule;
 - 10: Inter-layer reconstruction \mathbf{I}_F ;
 - 11: Synthesize fused image \mathbf{F} .
-

4. Data-specific filter design based on anisotropic heat diffusion

To fuse much more complementary information, based on heat diffusion theory, this section focuses on the data-specific and anisotropic filter design and its application in MID, and describes the implementation details step-by-step. Fig. 2 illustrates the basic idea of the proposed data-specific filter design.

4.1. Heat kernel theory

The heat kernel [38] has achieved great success over 2D manifold for geometry analysis [39,40] benefiting directly from its inherent properties of being structure sensitive and multi-scale. In this subsection, we briefly describe the heat kernel theory, and address some issues yet to be solved for the application of heat kernel in image space.

The heat diffusion process over the manifold \mathcal{M} is governed by the heat equation:

$$\Delta_{\mathcal{M}} u(x, t) = -\frac{\partial u(x, t)}{\partial t}, \quad (1)$$

where $\Delta_{\mathcal{M}}$ is the Laplace–Beltrami operator of \mathcal{M} . The generic solution of Eq.(1) is known as heat kernel $h(x, y, t)$, which can be defined as:

$$h(x, y, t) = \sum_{i=0}^{\infty} e^{-\lambda_i t} \phi_i(x) \phi_i(y), \quad (2)$$

where $\{\lambda_i\}_{i=0}^{\infty}$ ($0 = \lambda_0 < \lambda_1 \leq \dots$) and $\{\phi_i\}_{i=0}^{\infty}$ are respectively the eigenvalue of the Laplacian matrix and its corresponding eigenvector. Its matrix form can be rewritten as:

$$\mathbf{H}(t) = (\Psi \odot \Phi)\Phi'. \quad (3)$$

Here “ \odot ” represents the element-wise Hadamard product, $\Phi = \{\phi_i\}_{i=0}^{\infty}$, $\Psi = \{\gamma, \dots, \gamma'\}$, and $\gamma = \{e^{-\lambda_1 t}, e^{-\lambda_2 t}, \dots, e^{-\lambda_{\infty} t}\}$. It may

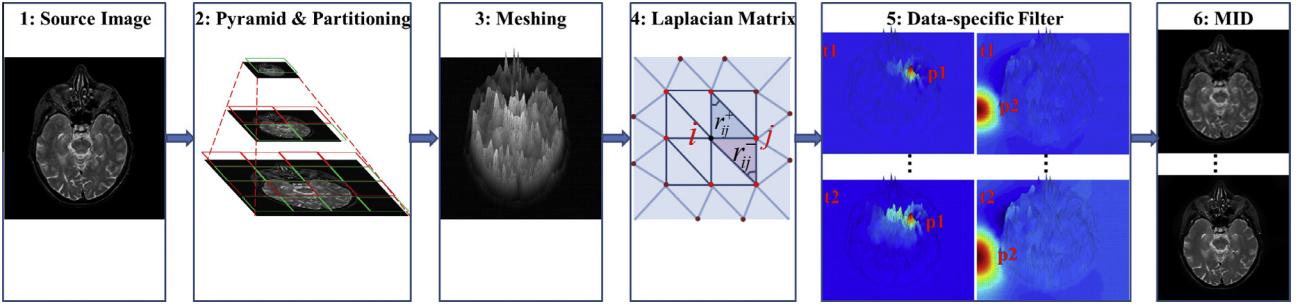


Fig. 2. Overview of the data-specific and anisotropic filter design with its application in MID. We first build an image pyramid and divide the images into overlapping blocks. Then, we map the image block to manifold mesh via Delaunay triangulation [37], and compute its Laplacian matrix over the manifold mesh. Based on the spectral analysis, the heat kernel is introduced to construct a data-specific and anisotropic filter, which is then employed to conduct MID. In Step five, please pay attention to the variance of the proposed filter at different locations (e.g., $\{p_1, p_2\}$) and scales (e.g., $\{t_1, t_2\}$).

be noted that, the spectral decomposition only needs to be computed once for multi-scale heat diffusion. Given t , the heat kernel can be directly computed using Eq. 3. Thus, the heat diffusion at $t + 1$ scale is independent of the heat diffusion of the adjacent t scale.

Since Laplacian matrix encodes the structure property of the underlying manifold, heat kernel naturally stores all of the properties that a good multi-scale image decomposition method should have, such as being multi-scale, anisotropic-structure sensitive, shift-invariant, Gaussian-decay, and geometry-specific. In this paper, we want to take those advantages for the multi-scale analysis of image. However, there are several issues to be addressed: (1) The construction of heat kernel is based on the topological structure of 2D manifold, while 2D image consists of pixel lattice that provides little geometrically-meaningful topology information; and (2) Naively transplanting heat diffusion to 2D image is not trivial due to the extremely high computational cost and storage requirements caused by massive pixels. Therefore, it strongly motivates us to design a technically-sound algorithm by exploring the underlying manifold mesh of image in a divide-and-conquer way. In the following subsections, we will address these issues in details.

4.2. Manifold construction and Laplacian analysis

To respect the inherent structure in image space, we represent a gray-scale image as a 2D manifold embedded in 3D Euclidean space with vertex coordinate $(x, y, I(x, y))$, where x and y are the pixel coordinates, $I(x, y)$ is its corresponding intensity value. The corresponding manifold mesh can be created via Delaunay triangulation [37]. This in fact serves as a precondition (topological structure) to define heat kernel over 2D image, which also gives rise to the solving of the first problem described in Section 4.1. Meanwhile, we replace the third dimensional coordinate $I(x, y)$ with $\beta I(x, y)$ to enhance the anisotropic structures of the image and control structure-sensitive heat diffusion. Specifically, an example is listed in Step three of Fig. 2, while the effect of parameter β will be detailed in Section 8.1.

Based on the constructed mesh, as shown in [41], its discrete Laplacian matrix can be defined using the cotangent scheme [42]. Denote $\{\nu_i\}_{i=1}^N$ as the vertex set, its discrete Laplacian matrix is $\mathbf{P} = \mathbf{A}^{-1}\mathbf{M}$. The mass matrix \mathbf{A} is a diagonal matrix and \mathbf{A}_{ii} is proportional to the area of all triangles sharing the vertex i . The stiffness matrix \mathbf{M} encodes the local structure of the image, which can be defined as

$$\mathbf{M}(i, j) = \begin{cases} \sum_{j \neq i} m(i, j) & \text{if } i = j \\ -m(i, j) & \text{if } \nu_i \text{ and } \nu_j \text{ are adjacent,} \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where $m(i, j) = \cot(r_{ij}^+) + \cot(r_{ij}^-)$, and r_{ij}^+ and r_{ij}^- are the two opposite angles sharing the edge connecting the vertices i and j (shown in Step 4 of Fig. 2). The eigenvalues and eigenvectors used in Eq. (3) can then be obtained by solving the generalized eigensystem $\mathbf{M}\Phi = \Lambda\mathbf{A}\Phi$, where $\Lambda = \{\lambda_i\}_{i=0}^\infty$.

However, it is still impractical to perform global eigen-decomposition for image comprising several hundred of thousand of pixel lattice. In the next subsection, we further elaborate a divide-and-conquer strategy to solve this problem.

4.3. Divide-and-conquer computation of data-specific filter

To deal with the second problem described in Section 4.1, we approximate the heat kernel in a local supporting area based on the characteristics of Gaussian decay by dividing the source image into overlapping blocks, which can greatly reduce the data scale, and avoid conventional time-consuming global spectral decomposition. Using overlapping blocks, we are able to effectively restrain block effects when stitching each smoothed block together to recover the smoothed image.

However, it is worth noting that the supporting areas may span the block boundaries for larger t , as shown in Fig. 3(f), the block effects are still inevitable in the smoothed image. A simple way is to increase the block overlapping width. But it gives rise to a great increase in the computational burden. Meanwhile, it is also against our original intention of using partitioning strategy. To eliminate the block effects without increasing the computational burden, we design an image pyramid based divide-and-conquer strategy. By taking the heat kernel pyramid into account, its quadtree based hierarchical structure can intrinsically form a scale space, wherein a block corresponding to the upper layer is divided into four sub-blocks in adjacent lower layer. Therefore, the heat diffusion of one block in the upper layer can be regarded as an effective approximation for that of the corresponding four sub-blocks. As shown in Fig. 3(g) and (h), the block effects can be effectively eliminated in this way. Specifically, another asset of our pyramid model is that it involves the local property resulted from partitioning strategy and global property encoded in the pyramid scheme. In order to facilitate the parallelized implementation of the proposed method (to be discussed in Section 7), the height and width of the block are both set to be an integral multiple of 16, and the overlapping width is set to be 16. Moreover, as illustrated in Fig. 4, we analyze the time-accuracy relationship for various block sizes by defining a multiple-indicator embedded average metric:

$$\zeta = \frac{\sum_{i=1}^5 Ind_{w,h}^i / \max(Ind^i)}{Tim_{w,h} / \min(Tim)}, \quad (5)$$

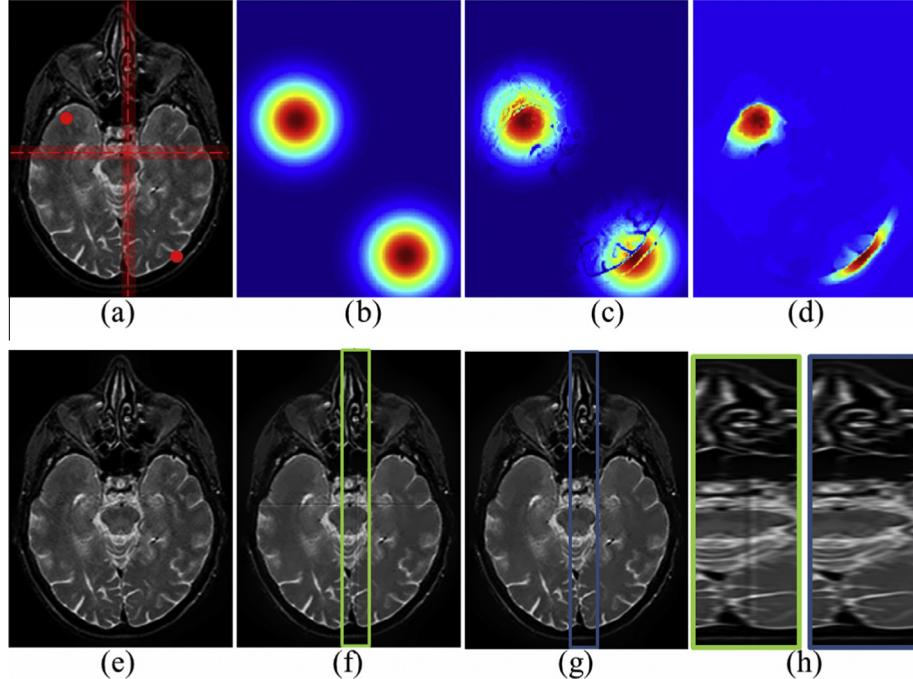


Fig. 3. Anisotropic heat diffusion and divide-and-conquer strategy based on the image pyramid. (a) The source image is divided into four overlapping blocks. The red region denotes the overlapping area between blocks. The red point represents the anchor pixel. (b–d) Kernels of Gaussian Filter, Bilateral Filter (BF), and anisotropic heat diffusion. Gaussian filter is isotropic. The anisotropy of BF is too local and weak and cannot entirely respect the global structure. Our kernel can well respect the global anisotropic structure. (e) Smoothed image corresponding to smaller t . (f) For larger t , naive block division inevitably gives rise to block-edge response (highlighted by red box). (g) By introducing the image pyramid, the block-edge response is eliminated. (h) Zoom-in effects. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

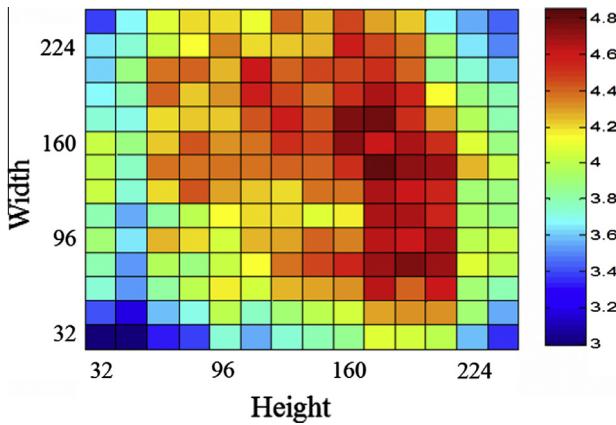


Fig. 4. Illustration of the time-accuracy relationship analysis for various block sizes. The relatively optimal size is 160×160 for 256×256 image.

where $Ind_{w,h}^i$ means the score of the i -th evaluation indicator with block width w and height h , $Tim_{w,h}$ denotes the corresponding time cost. From Fig. 4, we can observe that the relatively optimal block size is 160×160 for a 256×256 image.

4.4. Anisotropic multi-scale image decomposition

The defined heat kernel over 2D image can be regarded as a data-specific low-pass filter. Suppose a unit heat source located on the vertex x , the heat kernel $h(x,y,t)$ intrinsically presents the average heat amount diffused from x to y within time t . Thus, given an image organized as a matrix $\mathbf{I} \in \mathbb{R}^{N \times 1}$, the heat kernel based image smoothing can be given in an analytical form as:

$$\mathbf{I}(t) = \mathbf{H}(t)\mathbf{I} = ((\Psi \odot \Phi')\Phi')\mathbf{I}. \quad (6)$$

Here the parameter t is related to scale. A larger t means the smoothed image contains fewer details. Meanwhile, as illustrated in Fig. 3(d), by comparing with Gaussian filter (Fig. 3(b)) and bilateral filter (Fig. 3(c)), heat diffusion is shown to better respect the structure, and Fig. 3(e)–(g) correspondingly demonstrate the structure-preserving property of our filter when smoothing the image. It provides strong evidence that heat diffusion based multi-scale analysis strongly relates to the intrinsic geometrical structure of the image. Besides, directly inherited from heat kernel [38], our filter is multi-scale, anisotropic, shift-invariant, and data-specific. As incrementally and gradually increasing the values of t , multi-scale smoothed images can be independently generated by convolving the source image with the heat kernel, which will be used for low-rank analysis further to form a scale space of salient structures (discussed in Section 5).

5. MID-based low-rank analysis for scale-aware salient information

To analyze the frequency-relevant structures, the most common way is to compute the difference of adjacent layers, wherein the convolutional result with the largest scale means the base layer and the difference of adjacent layers denotes the detail layer at different scales. Although this method can guarantee that the original image itself can be perfectly reconstructed by decomposed layers, the data in base and detail layers cannot indicate the importance of fusion. In addition, the choice of the scale parameter requires the user to manually tune, and if the base layer contains fewer information, it gives rise to the loss of complementary information. Thus, it motivates us to incorporate the data-specific multi-scale decomposition into a low-rank analysis model to extract scale-aware salient information. The benefits of introducing

low-rank analysis include: (1) it makes the scale parameter t easy to choose; and (2) the low-rank part serves as common information across some images of a similar scale, which means the salient (important) information could be preserved in the fused result. We will detail the low-rank analysis model in the following subsections.

5.1. Low-rank decomposition theory

From the point of view of matrix decomposition, given a matrix \mathbf{B} , it can be divided into three components: low-rank component \mathbf{L} , sparse component \mathbf{S} , and noise component \mathbf{G} . The formulation of low-rank decomposition is defined as:

$$\mathbf{B} = \mathbf{L} + \mathbf{S} + \mathbf{G}, \quad s.t. \quad \text{rank}(\mathbf{L}) \leq r, \text{card}(\mathbf{S}) \leq c, \quad (7)$$

wherein r, c are respectively the rank constraint and the cardinality constraint.

Given an $m \times n$ dense matrix B and its r bilateral random projections (BRP) [43], its fast rank- r approximation can be obtained as:

$$L = Y_1(A_2^T Y_1)^{-1} Y_2^T, \quad (8)$$

where $Y_1 = BA_1, Y_2 = B^T A_2, A_1 \in \mathbb{R}^{n \times r}, A_2 \in \mathbb{R}^{m \times r}$ are independent Gaussian/SRFT random matrices [44]. Eq. 8 tends to perform poorly when the singular values of B decay slowly. By introducing the power scheme [45], we replace B with $\tilde{B} = (BB^T)^q B$ to accelerate the low-rank approximation and thus improve the accuracy [27], in which q is the power iteration strength. The BRP of \tilde{B} is $Y_1 = \tilde{B}A_1, Y_2 = \tilde{B}^T A_2$. The rank approximation of \tilde{B} is:

$$\tilde{L} = Y_1(A_2^T Y_1)^{-1} Y_2^T. \quad (9)$$

By calculating the QR decomposition of Y_1 and Y_2 , we can obtain the rank approximation of B as:

$$L = (\tilde{L})^{\frac{1}{2q+1}} = Q_1[R_1(A_2^T Y_1)^{-1} R_2^T]^{\frac{1}{2q+1}} Q_2^T. \quad (10)$$

Then, the sparse part S of B can be calculated as:

$$S = \mathcal{P}_\Omega(B - L), \quad (11)$$

where Ω is the nonzero subset of the first c largest entries of $|B - L|$.

Therefore, a dense matrix B can be decomposed into low-rank, sparse, and noise components, which can be efficiently and robustly estimated by GoDec method [27]. Here, we use the low-rank component \mathbf{L} to indicate the salient information commonly occurring in the image group with a similar scale, and further construct a scale space of salient features.

5.2. Scale-aware salient information extraction

In principle, let $\mathbf{I}(0)$ denotes the source image, and the smoothed image is represented as $\mathbf{I}(t)$ at time t . As shown in *Matrix Group* of Fig. 5, assuming each image pyramid level has $d \times k$

smoothed images and each group comprises d smoothed images. For each group, we suppose that there is a parameter sequence $[t_0^k, t_1^k, \dots, t_d^k]$ with a smaller step size for the smoothed images, where superscript k indicates the group index. Taking the original source image into account, we can reorganize $(d+1)$ image vectors to form a matrix as $\mathbf{B}_p^k = [\mathbf{l}_p(0), \mathbf{l}_p(t_1^k), \dots, \mathbf{l}_p(t_d^k)]$ for the p -th level of pyramid. Here, in order to extract the multi-scale intrinsic low-rank structure, we need a sufficient number of time scales to get enough smoothed images that have higher correlation and similar scale. Given the maximum value of scale parameter t , it will be equally divided into $d \times k$ segments to generate the smoothed images. Considering an enough large value of t , we can cast the scale parameter selection as a selection problem of k and d . Therefore, in this sense it makes the scale parameter t selection become easy. The selection and insensitivity analysis of scale parameter t are discussed later.

Fig. 5 describes the pipeline of salient feature extraction and scale-space construction based on low-rank analysis. Given the image matrix \mathbf{B}_p^k , the low-rank decomposition is employed to obtain the common information \mathbf{L}_p^k of image matrix. We then extract the first column $\mathbf{L}_p^{k,1}$ of matrix \mathbf{L}_p^k as the salient features of corresponding scale (the k -th group). The right column of Fig. 6 shows the extracted low-rank component of a noise-perturbed image, which well captures the intrinsic salient information of source image while automatically removing the noise. Therefore, coinciding with the theory of low-rank analysis, our method can effectively decouple the noise from the source image in a “built-in” way. By group-wisely executing the low-rank analysis with specific rank constraint, k salient features of source image can be obtained, denoted as $\{\mathbf{L}_p^{1,1}, \mathbf{L}_p^{2,1}, \dots, \mathbf{L}_p^{k,1}\}$. Since each image group corresponds to a different similar scale, the extracted information thus can be used to construct a scale space for each pyramid level.

Besides, in our method, instead of directly using the original source images, we construct a scale space by resorting to the extracted low-rank structures, which are salient and common occurring across similar-scale smoothed images. And we decomposes them into different layers by computing the difference of adjacent-scale low-rank structures. On that basis, the salient information could then be effectively fused into the result. Let $\mathbf{D}_p^i = \mathbf{L}_p^{i,1} - \mathbf{L}_p^{i+1,1}$ denotes the i -th small-scale detail layer at the p -th pyramid level, and $\mathbf{E}_p = \mathbf{L}_p^{k,1}$ as the large-scale base layer. Therefore, the *intra-level reconstruction* of the decomposed image at the p -th pyramid level can be formulated as:

$$\mathbf{L}_p^{1,1} = \sum_{i=1}^{k-1} \mathbf{D}_p^i + \mathbf{E}_p. \quad (12)$$

For the *inter-level reconstruction* of the pyramid, let $\mathbf{D}_p = \mathbf{L}_p^{1,1} - U(\mathbf{L}_{p+1}^{1,1})$ denotes the p -th detail level of the pyramid,

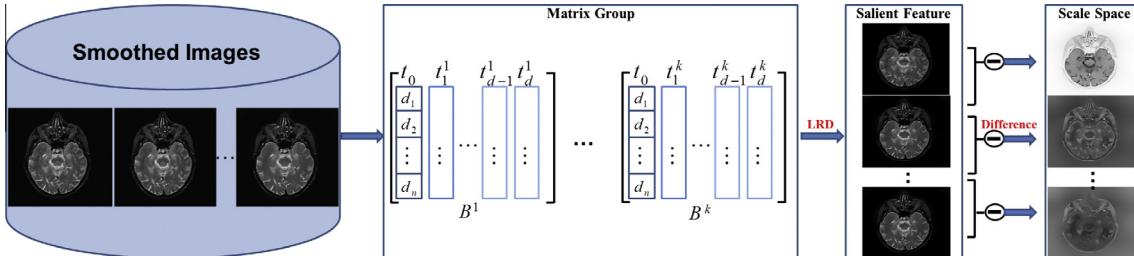


Fig. 5. Pipeline of MID-based low-rank analysis for scale-aware features. We conduct multi-scale image decomposition (MID) based on anisotropic heat kernel to obtain multi-scale images. Then, we group the multi-scale images into k groups and extract the salient features from each groups via low-rank analysis, and compute the differences of adjacent groups to form a scale space of salient information.

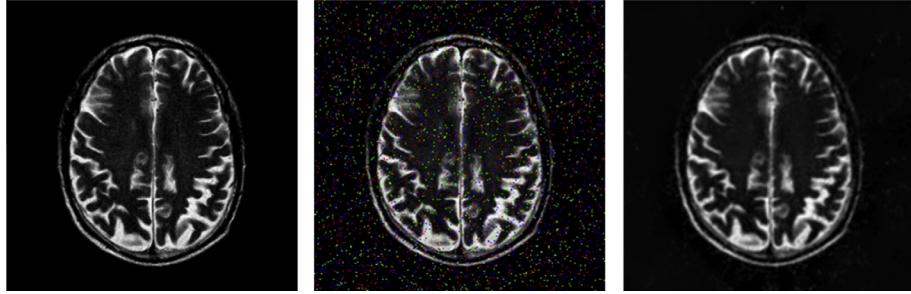


Fig. 6. Illustration of low-rank analysis. Left: The clean image. Middle: Salt and pepper noise-perturbed image. Right: The extracted low-rank component well captures the intrinsic salient information of original image while automatically removing the noise.

Table 1
Quantitative comparison for different maximum values of scale parameter t .

t	PSNR	MI	MSSIM	FSIM	FSIMC
56	42.3472	3.0722	0.7747	0.9838	0.9509
58	42.3479	3.0718	0.7747	0.9838	0.9509
60	42.3505	3.0736	0.7748	0.9838	0.9510
62	42.3483	3.0716	0.7747	0.9838	0.9509
64	42.3482	3.0715	0.7747	0.9837	0.9509
66	42.3482	3.0712	0.7747	0.9837	0.9509

where $U(\cdot)$ means up-sampling operation. Here, the bicubic interpolation method is used to generate the high-resolution image. The *inter-level reconstruction* can be formulated as:

$$\mathbf{L}_i^{1,1} = (1 - \alpha)(\mathbf{D}_i + U(\mathbf{L}_{i+1}^{1,1})) + \alpha\mathbf{L}_i^{1,1}, \quad (13)$$

where $i = \{1, 2, \dots, p - 1\}$, α is a constant to balance the detail and the noise.

As for the selection of t , we measure its rationality according to MSSIM metric, which can evaluate the salient edge-preserving

quality on the entire image. It is expected to have the better performance at scale t , if $MSSIM \geq 0.8$ [46]. By extensively testing experiments over the MRI, SPECT, and PET images, we empirically find that $t = 60$ can meet the requirement of $MSSIM \geq 0.8$. By fixing other parameters, we show the insensitivity analysis of scale parameter t in Table 1, wherein the changes of five metrics are extremely small for different maximum values of parameter t . It indicates that the low-rank analysis helps reduce the sensitivity during the selection of parameter t , and makes the scale parameter t easy to choose.

6. Improved frequency-specific fusion rule

Since MRI images are of high-resolution and primarily reveal the anatomical structures, the anatomical information should be effectively displayed for the accurate location of lesion. The SPECT/PET images usually directly reflect the saliency and activity of the tissue, whose fusing weights should be a monotonically increasing function.

For large-scale image layers, the base layer weights of MRI image are set as 2 to reduce the harmful influence of black back-

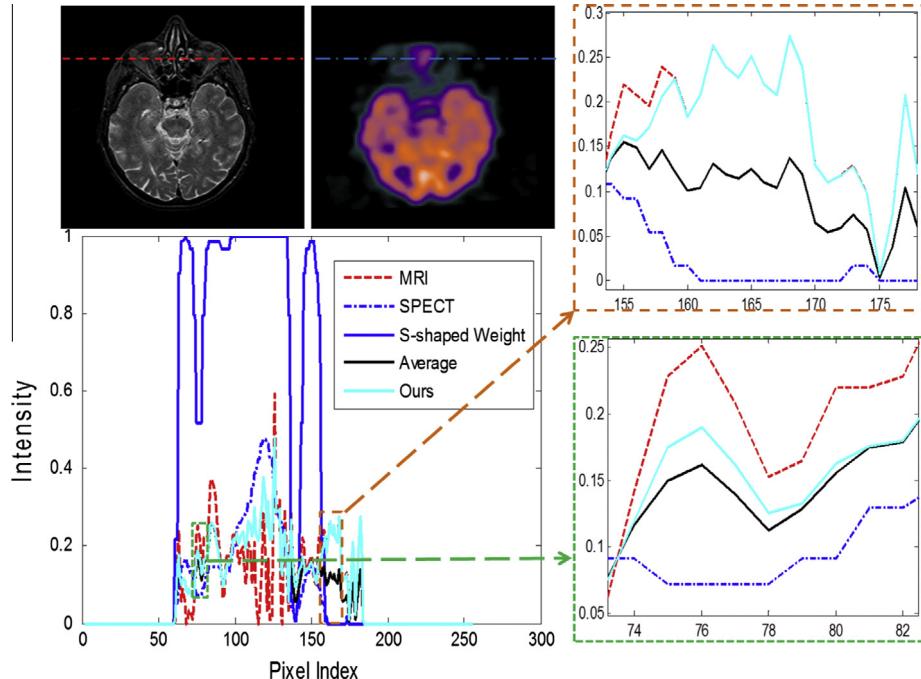


Fig. 7. Illustration of S-shaped mapping function. Top Left: Source images. The red dashed and blue dash-dot lines indicate the input data in source images. Bottom Left: The fusion comparison. Blue line indicates the fusion weight mapped by S-shaped function ($\theta = 0.15$). Top Right: There is no complementary information in SPECT image (the intensity is zero) at the edge region of MRI image. The Average-Maximum fusion rule tends to reduce intensity and blur the edge (indicated by black color). The result from the S-shaped fusion rule is quite similar to the original MRI data. Bottom Right: For the MRI edge regions where SPECT image contains complementary information, the Average-Maximum fusion rule still produces similar blurred result, while our result achieves a better performance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ground in SPECT/PET image; the weights for functional images are in consistent with an S-shaped weighting function $f(x)$, which should start from the point $(0, 0)$ and end at the point $(1, 1)$,

$$f(x) = \begin{cases} ax^3 & \text{if } x \leq \theta \\ 1 - b(1 - \theta)^c & \text{if } x > \theta \end{cases} \quad (14)$$

Here θ is the threshold of the inflection point. In order to better balance the visual perception between anatomic and functional images, we assume that the weight of inflection point is 0.5. Considering the C^1 continuity constraint at the inflection point, the parameters a, b , and c can be obtained by

$$\begin{cases} f(\theta^-) = \frac{1}{2} \\ f(\theta^+) = \frac{1}{2} \\ f'(\theta^-) = f'(\theta^+) \end{cases} \Rightarrow \begin{cases} a = \frac{1}{2\theta^3} \\ b = \frac{1}{2(1-\theta)^c} \\ c = \frac{3(1-\theta)}{\theta} \end{cases} \quad (15)$$

Given the mapping function $f(x)$, the weight for the large-scale layer coefficient $\mathbf{E}_p(i, j)$ of SPECT/PET image is defined as:

$$\mathbf{W}_{\mathbf{E}_p}(i, j) = f(\mathbf{E}_p(i, j)). \quad (16)$$

As for the small-scale detail layers, the larger coefficients should be retained to preserve the salient details. We thus adopt the maximum selection principle to fuse the high-frequency coefficients. Fig. 7 shows an example of S-shaped mapping function. For our S-shaped rule, the fusion weight is shown with the blue line, wherein relatively larger value is assigned to the higher intensity (shown in bottom left figure of Fig. 7). The top right figure shows SPECT image does not contain complementary information (the intensity is zero) corresponding to the edge region of MRI image. The Average-Maximum fusion rule tends to reduce intensity and blur the sharp edge (indicated by black color) while our result effectively preserves the sharp edge (represented by cyan color). The bottom right figure shows that the SPECT image contain the complementary information corresponding to another edge region of MRI image. Our result achieves the better contrast than that resulted from the Average-Maximum rule. Thus, the S-shaped function is better to fuse the complementary information and adaptively preserve the amplitude of the structure. The choice of parameter θ and the structure-preserving capability analysis of new fusion rule will be discussed in Sections 8.1 and 8.6.

7. CUDA-based parallel implementation

Algorithm 2. Parallel generation of multi-scale images

input: Image block \mathbf{B} , β , and the number of the smallest magnitude (sm) eigenvalue n .

output: Multi-scale images $\mathbf{B}(t)$.

Pixel-level kernel for each vertex i in a block:

Step I.

for $j = 1 : 6$

if v_j and v_{j-1} exist

1: $s1 = v_i - v_{j-1}; s2 = v_j - v_{j-1};$

2: $\mathbf{A}_{ii} += \text{norm}(\text{cross}(s1, s2));$

3: $\text{fac} = \text{norm}(s1) \times \text{norm}(s2);$

4: $m_{ij} += 1.0 / \tan(\text{acos}(\text{dot}(s1, s2) / \text{fac}));$

end

if v_j and v_{j+1} exist

compute \mathbf{A}_i and m_{ij} as steps 1 ~ 4;

end

$m_{ii} += m_{ij}; \text{stif}_{ij} = -m_{ij};$

$\text{rowIdx}_{ij} = \text{Idx}_i; \text{colIdx}_{ij} = \text{Idx}_j;$

end

$\text{stif}_{ii} = m_{ii}; \text{area}_{ii} = \mathbf{A}_{ii};$

$\text{rowIdx}_{ii} = \text{Idx}_i; \text{colIdx}_{ii} = \text{Idx}_i;$

5: $[\Phi, \Lambda] = \text{eigs}(\mathbf{M}, \mathbf{A}, n', \text{'sm'});$

Block-level kernel for each block:

Step II. Compute fixed component for each block:

6: $\Gamma = \Phi' \mathbf{B};$

Step III. Generate multi-scale images for i -th pixel:

for $j = 1:n$

7: $\mathbf{B}_{ij}(t) += \exp(-\lambda_j t) \times \phi_{ij} \times \Gamma_j;$

end

To alleviate the time-consuming computation, as illustrated in Fig. 8, we implement our method via well-designed pixel-level and block-level CUDA algorithms, which allow us to efficiently generate Laplacian matrix and multi-scale images for each block in a parallel way.

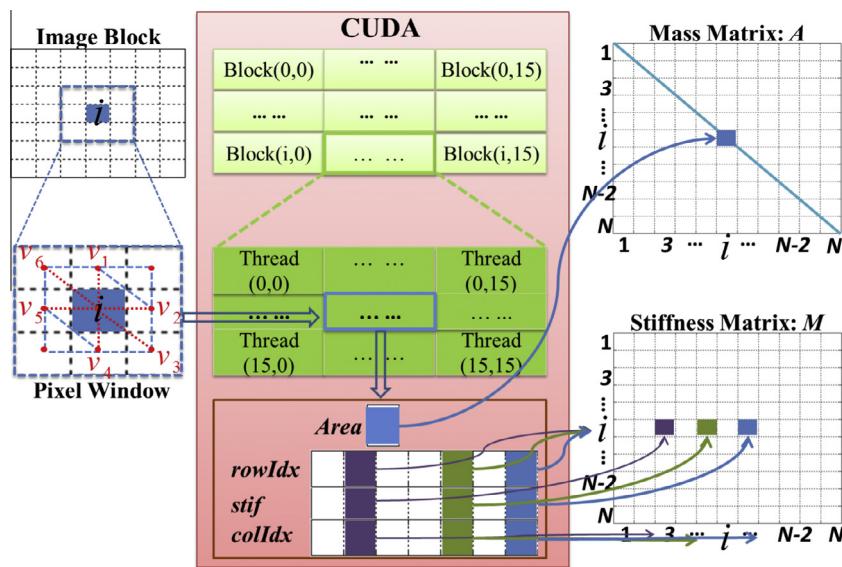


Fig. 8. Illustration of CUDA-based Laplacian matrix computation in a block-parallel way. Left: image block and pixel window. Middle: algorithmic architecture on CUDA. Right: Laplacian matrix construction via mapping area , rowIdx , colIdx , stif to \mathbf{M} and \mathbf{A} .

7.1. Pixel-level parallelized Laplacian matrix construction

Since the pixel-wise computation involved in Laplacian matrix construction is independent of each other, we compute Laplacian matrix in parallel by invoking a CUDA thread for each pixel to compute its entry $m(i,j)$ and \mathbf{A}_{ii} (Step I in [Algorithm 2](#)). Meanwhile, each thread has a lower computational burden, because there are at most six edges to be calculated (denoted by red-dotted line in [Fig. 8](#)) for each pixel.

Considering the sparsity of stiffness matrix \mathbf{M} , we need a 3-tuple ($rowIdx$: row index, $colIdx$: column index, $stif$: data value) to store the non-zero entries. As for diagonal matrix \mathbf{A} , a unique tuple ($area$: area value) is enough, because the row index and column index are the same as the corresponding pixel index.

For each block, our parallelized method consists of the following steps. First, we transfer the image block containing N pixels from CPU to GPU, and compute the Laplacian matrix using N CUDA threads. Second, in each thread corresponding to pixel i , we extract its eight one-ring neighboring pixels and denote the no-existing neighbors with negative values. Third, we compute the area $area_{ii}$ and the affinities $stif_{ij}$ between pixel i and other six pixels (refer to Eq. 4). Finally, Laplacian matrix is constructed on GPU. For a 256×256 image, the whole process can be finished within 0.47 s with the help of GPU implementation, which is only 1/50 time cost of the CPU-based serial processing (23.71 s). Then, the eigenvalues Λ and eigenvectors Φ of each Laplacian matrix can be obtained by solving the generalized eigen-system with Matlab function **eigs**.

7.2. Block-level parallelized multi-scale image generation

Given the eigen-decomposition of block-wise Laplacian Matrix, it seems that we can easily design block-level CUDA kernels to generate multi-scale images in parallel. However, it is impractical to directly compute the anisotropic heat kernel and conduct image convolution due to the extremely-high GPU memory cost (e.g., $\mathbf{H}(t) \in \mathbb{R}^{N \times N}$). To address this problem, we rewrite Eq. (6) as $\mathbf{I}(t) = (\Psi \odot \Phi)(\Phi' \mathbf{I})$ ($\Psi \in \mathbb{R}^{N \times n}$, $\Phi \in \mathbb{R}^{N \times n}$), where N and n are respectively the image pixel number and the eigenvector number ($n \ll N$). Besides, $\Gamma = \Phi' \mathbf{I} \in \mathbb{R}^{n \times 1}$ is constant during multi-scale convolution ([Algorithm 2](#) Step III), which only needs to be computed once for each block ([Algorithm 2](#) Step II).

Meanwhile, we can make quantitative analysis on the memory cost and computational burden to measure the performance improvement of the rewritten equation. [Fig. 9](#) compares the calculation processes and accompanying costs of space allocation and floating-point operations (FLO) between the naive calculating method and our practically-adopted method. Since $\Psi \odot \Phi$ is equally used by the two methods, we do not consider its influence on memory cost and computational burden. For the naive calculating method, $T \times N^2 \times (2n - 1)$ FLOs and $T \times N \times (2N - 1)$ FLOs are respectively needed for the computing tasks shown in [Fig. 9\(a\)](#), where T denotes the number of the multi-scale images. And $N \times N$ additional storage cells are needed for heat kernel $\mathbf{H}(t)$ computation. In our practically-adopted method ([Fig. 9\(b\)](#)), $n \times (2N - 1)$ FLOs are performed and n additional storage cells are needed to compute Γ . And we need perform $T \times N \times (2n - 1)$ times FLO to get T smoothed images in Step III. Benefiting from this strategy, it not only can reduce about 99% storage cost but also can greatly reduce the total operational burden ($n \times (2N - 1) + T \times N \times (2n - 1)$ FLOs vs. $T \times N \times (2N - 1) + T \times N^2 \times (2n - 1)$ FLOs), which guarantees us to achieve all the multi-scale images within 0.37 s.

Therefore, as detailed in [Algorithm 2](#), our CUDA-based multi-scale images generation method can be summarized as follows. Given Λ and Φ , we first compute the constant component Γ , and

store it for later use. And then we invoke one CUDA thread for each image scale to conduct image convolution operation. Once all the CUDA threads complete the calculation, we copy the multi-scale images from GPU to host memory for further low-rank analysis.

8. Experiments and evaluation

We evaluate and validate our newly-proposed method through various multi-modal image fusion experiments wherein the multi-modal image benchmark contains 86 image pairs of normal brain, 144 image pairs of neoplastic disease, and 24 image pairs of degenerative disease. And all the experimental images are from Harvard University site,¹ which had already been registered and have a uniform size of 256×256 . Meanwhile, to verify and evaluate the robustness of our fusion method, we also perturb each image with white Gaussian, salt and pepper noise, which further provides 13,970 image pairs. For white Gaussian noise, the adopted noise power (in dB) is $\{1 \text{ dB}, 2 \text{ dB}, \dots, 40 \text{ dB}\}$. For salt and pepper noise, the percentages of noise-perturbed pixel are $\{1\%, 2\%, \dots, 15\%\}$.

The evaluation includes comprehensive comparison with 5 state-of-the-art methods, including wavelet-based method, Laplacian pyramidal-based method (LP), curvelet-based method, contourlet-based method, and NSCT-based method (NSCT). To quantitatively evaluate the fusion quality, we resort to 5 commonly-used indicators, including peak signal-to-noise ratio (PSNR), mutual information (MI), mean structural similarity (MSSIM) [46], feature similarity (FSIM) [47], and FSIM for the chrominance information (FSIMC). Specifically, MI measures the degree of dependence; MSSIM can evaluate the salient edge preserving quality on the whole image; FSIM and FSIMC can measure the feature similarity between the source image and the fused image [47]. The PSNR, MI, MSSIM, FSIM, and FSIMC between the fused image \mathbf{F} and the referenced image \mathbf{R} , are respectively defined as:

$$PSNR(F, R) = 10 \times \log \left(\frac{255^2}{MSE(F, R)} \right), \quad (17)$$

$$MSE(F, R) = \sum_{i=0}^N (F(i) - R(i))^2, \quad (18)$$

$$MI(F, R) = \sum_{f \in F} \sum_{r \in R} P_{FR}(f, r) \log \left(\frac{P_{FR}(f, r)}{P_F(f)P_R(r)} \right), \quad (19)$$

where $P_{FR}(f, r)$ is the joint distribution, $P_F(f)$ and $P_R(r)$ are the marginal probability distribution,

$$MSSIM(F, R) = \frac{1}{|W|} \sum_{w=1}^W SSIM(F_w, R_w), \quad (20)$$

where W means the family of the sliding windows, $|W|$ is the number of the sliding windows, $SSIM(F_w, R_w)$ denotes the structural similarity metric for the corresponding regions and is defined as:

$$SSIM(F_w, R_w) = \frac{(2\bar{w}_F\bar{w}_R + C_1)(2\sigma_{w_F}\sigma_{w_R} + C_2)}{(\bar{w}_F^2 + \bar{w}_R^2 + C_1)(\sigma_{w_F}^2 + \sigma_{w_R}^2 + C_2)} \cdot \frac{(\sigma_{w_F w_R} + C_2)}{(\sigma_{w_F}\sigma_{w_R} + C_3)}. \quad (21)$$

where C_1, C_2, C_3 are constants, \bar{w} and σ_w are respectively the mean and the variance of the region w , $\sigma_{w_F w_R}$ means the covariance of the regions w_F and w_R ,

$$FSIM(F, R) = \frac{\sum_{i \in \Omega} S_{PC}(i)S_G(i)PC(i)}{\sum_{i \in \Omega} PC(i)}, \quad (22)$$

¹ <http://www.med.harvard.edu/AANLIB/home.html>.

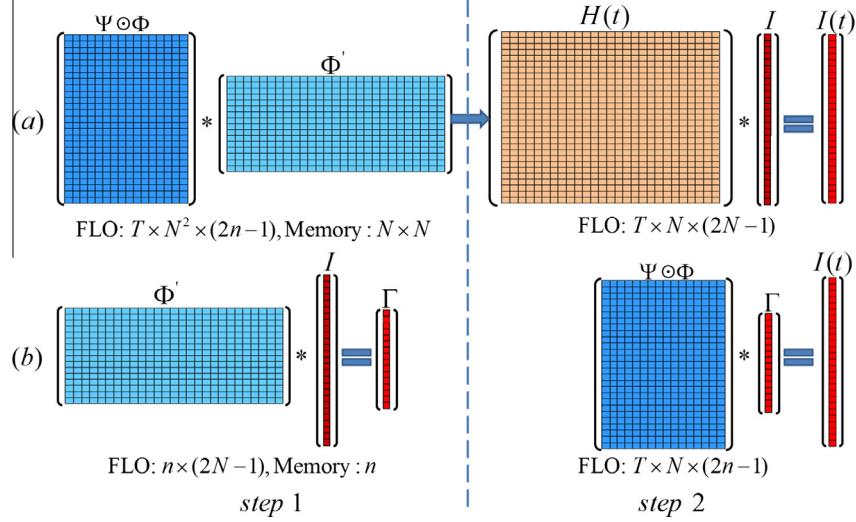


Fig. 9. Illustration and comparison of multi-scale image convolution methods. (a) The naive calculating method. Step 1: compute the anisotropic heat kernel $\mathbf{H}(t)$; Step 2: compute multi-scale image $\mathbf{I}(t)$. (b) Our practically-adopted method. Step 1: compute the constant component Γ ; Step 2: compute multi-scale image $\mathbf{I}(t)$. It only needs $n \times (2N - 1) + T \times N \times (2n - 1)$ times FLO and n additional storage cells, while $T \times N \times (2N - 1) + T \times N^2 \times (2n - 1)$ FLOs and $N \times N$ additional storage cells are needed in Method (a). T denotes the number of the multi-scale images, N and n are respectively the pixel number and the eigenvector number ($n \ll N$).

$$FSIMC(F, R) = \frac{\sum_{i \in \Omega} S_{PC}(i) S_G(i) [S_I(i) S_Q(i)]^v PC(i)}{\sum_{i \in \Omega} PC(i)}, \quad (23)$$

$$PC(i) = \max(PC_F(i), PC_R(i)), \quad (24)$$

$$S_{PC}(i) = \frac{2PC_F(i)PC_R(i) + T_1}{PC_F(i)^2 + PC_R(i)^2 + T_1}, \quad (25)$$

$$S_G(i) = \frac{2G_F(i)G_R(i) + T_2}{G_F(i)^2 + G_R(i)^2 + T_2}, \quad (26)$$

$$S_I(i) = \frac{2I_F(i)I_R(i) + T_3}{I_F(i)^2 + I_R(i)^2 + T_3}, \quad (27)$$

$$S_Q(i) = \frac{2Q_F(i)Q_R(i) + T_4}{Q_F(i)^2 + Q_R(i)^2 + T_4}. \quad (28)$$

It may be noted that, Ω denotes the entire image spatial domain, i is the spatial coordinate, T_1, T_2, T_3, T_4 are constants, $v > 0$ is a balance parameter. PC measures the significance of local structure, and G is the gradient magnitude. I and Q are the corresponding chromatic channels in the YIQ color space. The higher the values of PSNR, MI, MSSIM, FSIM, and FSIMC are, the better fusion result could be achieved.

Given the two source images C (SPECT/PET), G (MRI) and the fused result F , the first two indicators ($PSNR(F, C, G)$ and $MI(F, C, G)$) are computed by summing up their values resulted from (F, C) and (F, G) . As for the last three indicators, we respectively compute their average values of the whole image. To guarantee an objective, unbiased quality assessment, all the indicator's testing algorithms have been implemented by a third-party Objective Evaluation Criteria Package.² Meanwhile, we construct 3-level image pyramid for all experiments of our method. And we uniformly perform 4-level image decomposition for other methods. For the contourlet transform based method, the decomposition level is set to be [2,2–4], and the directional filter is set to be 'pkva' [8]. The pyramid filter used in wavelet and contourlet methods is set to be '9–7' [8]. For

NSCT method, the directional and pyramid filters are respectively set to be 'dmaxflat7' and 'maxflat'.

Moreover, Table 2 documents the time statistics by running the 256×256 image on a computer with 8 GB RAM, Intel i7-3770 CPU, 3.4 GHz and NVIDIA GeForce GTX 660 Ti, wherein we detail the time cost of our method, including Laplacian matrix computation (Lap), eigen-decomposition (Eig-Decom), smoothing operators, low-rank decomposition (LRD), and image fusion. It shows that the dominating time cost of our method is mainly expended on the Eigen-decomposition operation. Note that our method costs about 8.34 s to complete the fusion, which is faster than NSCT-based method (36 s).

8.1. Parameter selection

There are the parameters ($\beta, \theta, \alpha, d, k, c$, and r) that may affect the final fusion result. Here we discuss how to select these parameters.

First, small β gives rise to isotropic heat diffusion and tends to drastically weaken the sharp structures. Otherwise, although large β can enhance anisotropy, it fails to effectively decouple noise from the image. Let us pay attention to the green curve shown in Fig. 10, the value of MI improves with the increasing of β , and approaches the steady constant in the range [1100, 1300]. Therefore, in the experiments, we usually set the value of β to be in this range.

Second, the parameter θ determines the amount of functional information contained in the final fused image. For larger θ , only the most important functional information can be fused into the final result. The red curve in Fig. 10 illustrates the influence of θ . We find that the best fusion result can be obtained when the value of θ is set to be in the range [0.1, 0.3].

Third, the parameter α is used to control the block-edge response. Small α oftentimes gives rise to the obscured fusion result with fewer details. The blue curve in Fig. 10 depicts the influence of α . It shows that higher MI can be obtained when the value of α is greater than 0.5. Specifically, when α is around 0.9, MI reaches a maximum.

Furthermore, we investigate the influence of the parameters k and d . Given the maximum value of scale parameter t , the step size is equal to $t/(k \times d)$. By taking the sum of five metrics as a function of group number k and image number d , Fig. 11 analyzes the fusion

² http://www.quxiaobo.org/index_software.html.

quality. It shows that, when $d = 7$, the fused results achieve relatively higher quantitative performance. And the best result can be obtained when $k = 17$ and $d = 7$.

Meanwhile, taking Eq. 5 as an evaluation function, we analyze the influence of the parameters c and r . r_1, r_2 , and r_3 denote the rank constraints corresponding to the three levels of the image pyramid. According to Fig. 12, the better fusion effect can be achieved when $r_1 = 1$. As r_1 increases, the performance gradually deteriorates, since a larger r_1 tends to bring the noise into the low-rank component. As down-sampling goes, the noises gradually decrease at each coarse layer of the image pyramid, and thus the changes of r_2 and r_3 have a relatively small influence on the result. From the highlighted row (indicated by deep red box), our method performs better when $c = 2.1e + 3$. Therefore, considering the whole performance, we assign $c = 2.1e + 3$ and $r = \{1, 2, 3\}$ as an optimal setting. In all our following experiments, we select the parameters according to the aforementioned analysis, unless otherwise indicated.

8.2. Noise-free image fusion experiments and evaluation

Fig. 13(a) lists the MRI-SPECT fusion results from different methods, whose quantitative evaluation is shown at the bottom-left of Fig. 13. Comparing with other methods, it shows that our method can integrate more functional and structural information, and suppress the color distortion to the greatest extent, because our data-specific filter can better respect the distinct structures and the S-shaped weighting function can elegantly compromise the multi-source complementary information extracted by our low-rank analysis model. For example, in our results the anatomical brain structures from MRI (indicated by the red arrow) can be faithfully preserved, and the functional information from SPECT is also completely integrated (especially for the region boundaries indicated by the green box).

In sharp contrast, wavelet, LP, and contourlet methods fail to adaptively reduce the side effect caused by black background on the entire images, whose results appear to be too dark and lead to color distortion. The curvelet-based method introduces obvious artifacts nearby the edge (shown in the green arrow). Although NSCT-based method can produce good visual results due to its non-subsampling nature, however, it still misses certain sharp feature (shown in the red and green boxes). In addition, Fig. 13(b) lists another set of MRI-PET fusion results, which also demonstrates the same advantages of our method (refer to the regions illustrated by the magenta box). It should be noted that, curvelet-based method tends to produce halo artifacts nearby the edges (indicated by the magenta arrow). Moreover, based on the quantitative evaluation shown at the bottom-right of Fig. 13, all the five indicators again prove that our method outperforms other methods.

Meanwhile, to verify the generic advantages of our method, Table 3 further lists the quantitative comparison of average fusion quality (with five indicators: PSNR, MI, MSSIM, FSIM, and FSIMC)

Table 2

Time performance (in seconds). The image size is 256×256 , and the block size is 160×160 . In the Eig-Decom column, the first value is the eigen-decomposition cost for one block, while the second one is the total number of the blocks in the image pyramid.

Method (s)	Wavelet	LP	Curvelet	Contourlet	NSCT	Ours					
						Lap	Eig-Decom	Smoothing	LRD	Fusion	Sum
Fig. 13(a)	0.03	0.02	3.98	2.34	37.25	0.48	1.04×6	0.38	0.70	0.09	7.90
Fig. 13(b)	0.04	0.03	2.79	1.92	36.02	0.46	1.00×6	0.39	0.69	0.08	7.65
Fig. 14(a)	0.05	0.03	2.50	1.75	36.08	0.48	1.11×6	0.37	0.70	0.09	8.32
Fig. 14(b)	0.03	0.02	2.54	1.67	36.71	0.47	1.02×6	0.36	0.71	0.09	7.79
Fig. 17(a) and (b)	0.03	0.02	2.49	1.69	35.87	0.48	1.41×6	0.38	0.68	0.10	10.10
Fig. 19(a) and (b)	0.03	0.02	2.44	1.64	36.02	0.46	1.12×6	0.36	0.69	0.09	8.33
Average	0.035	0.023	2.79	1.835	36.325	0.472	1.05×6	0.373	0.695	0.09	8.34

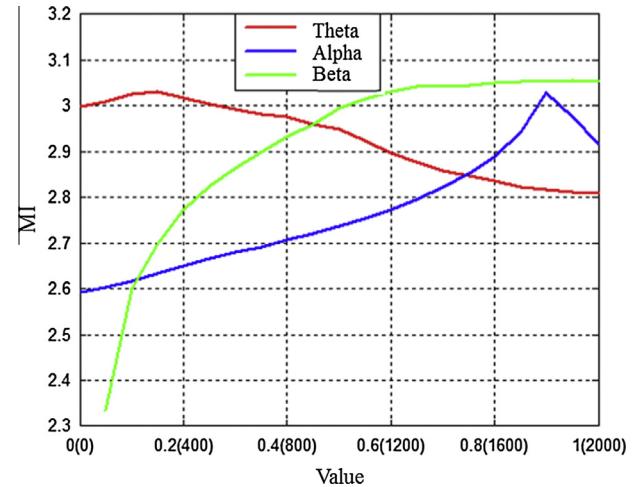


Fig. 10. Illustration of statistics-based parameters' analysis. Green curve: we set the value of the parameter β to be in the range [1100, 1300]. Red curve: the best fusion result can be obtained when the value of θ is in the range [0.1, 0.3]. Blue curve: if α equals 0.9, MI value reaches a maximum. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

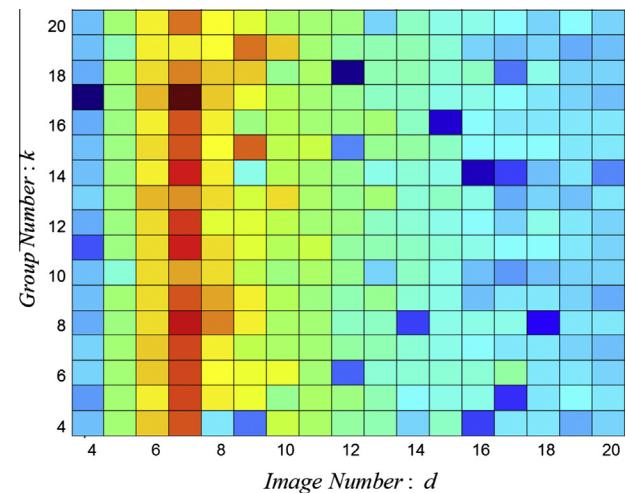


Fig. 11. The statistical analysis of selecting different group number parameter k and smoothed image number parameter d for sharp feature extraction. It depicts the effects resulted from different combinations of these two parameters, where the color represents the MI value of the fusion result. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

over 254-pair noise-free image fusion results. From Table 3, we can observe that our method fully outperforms the competing methods for all quality indicators. Therefore, it proves that our

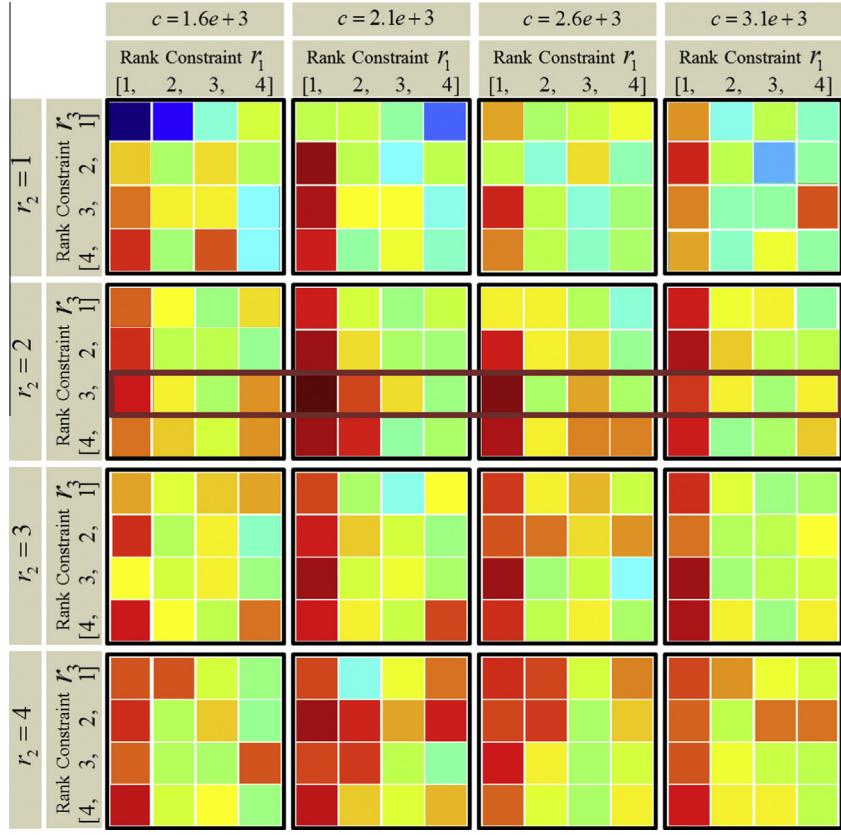


Fig. 12. Illustration of r and c of low-rank analysis. r_i denotes the rank constraint in the i -th level of image pyramid. The color from red to blue indicates the performance from better to worse. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

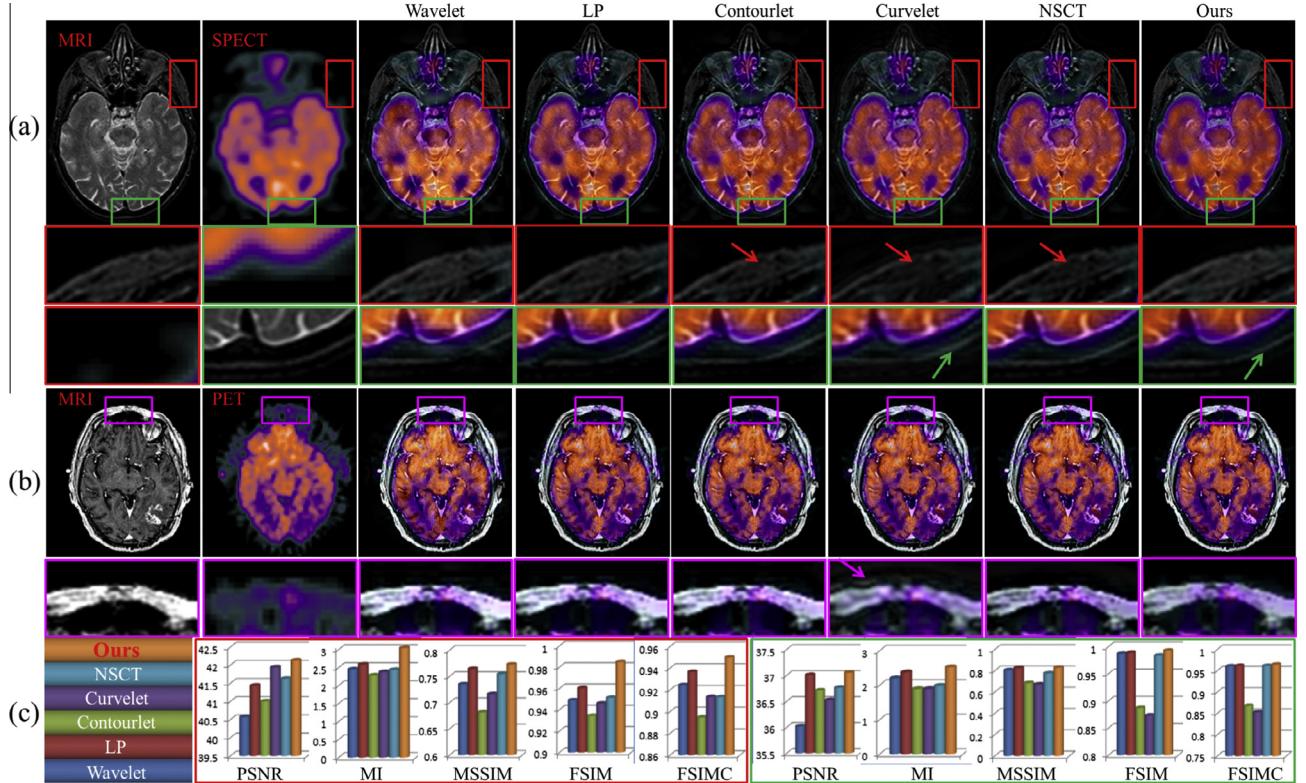


Fig. 13. The fusion results and their comparison/evaluation over noise-free images. (a) MRI/SPECT image fusion. (b) MRI/PET image fusion, $\beta = 1300, \theta = 0.2$. (c) The quantitative evaluation, the left column is for (a) and the right column is for (b).

method can re-integrate much more anatomical and physiological information into the fused image to guarantee better visual perception.

8.3. Noise-perturbed image fusion experiments and evaluation

We perform three more sets of experiments on noise-perturbed images to verify the robustness of our method. Fig. 14 lists the Gaussian noise-perturbed fusion results from different methods.

Table 3

Quantitative comparison of average fusion quality over all 254-pair noise-free images. Bold values indicates the maximum value of each column.

Method	PSNR	MI	MSSIM	FSIM	FSIMC
Wavelet	38.47	1.66	0.76	0.981	0.922
LP	38.48	1.70	0.74	0.984	0.921
Contourlet	41.84	1.94	0.73	0.947	0.881
Curvelet	38.48	1.61	0.68	0.923	0.875
NSCT	41.82	2.18	0.80	0.988	0.922
Ours	45.06	2.58	0.82	0.990	0.956

In Fig. 14(a)–(b), the source images are respectively perturbed with 5 dB and 25 dB white Gaussian noise. By comparison, it is obvious that other five methods are more sensitive to noise, which seriously decrease the visual quality. In sharp contrast, our method can effectively decouple the noise while simultaneously extracting most of the structural and functional details from the source images. Meanwhile, we also document the quantitative evaluation results in Fig. 14(c), which indicates that our method performs far better than most of the state-of-the-art methods in the overall fusion quality. Besides, Fig. 15 shows other eight fusion results from our method for the images with multiple noise levels. The fusion results have proper contrast while preserving the intrinsic structures when noise level is below 30 dB, which demonstrates that our method can effectively suppress the noise blow 30 dB. In fact, the noise compression of our method mainly benefits from low-rank analysis. Unlike the traditional methods, before constructing a scale space, we first conduct low-rank analysis on multi-scale image groups constructed by the smoothed images and source image, and use the low-rank component to indicate the salient information commonly occurring in the image group. So that, we can simultaneously decouple the noise from the

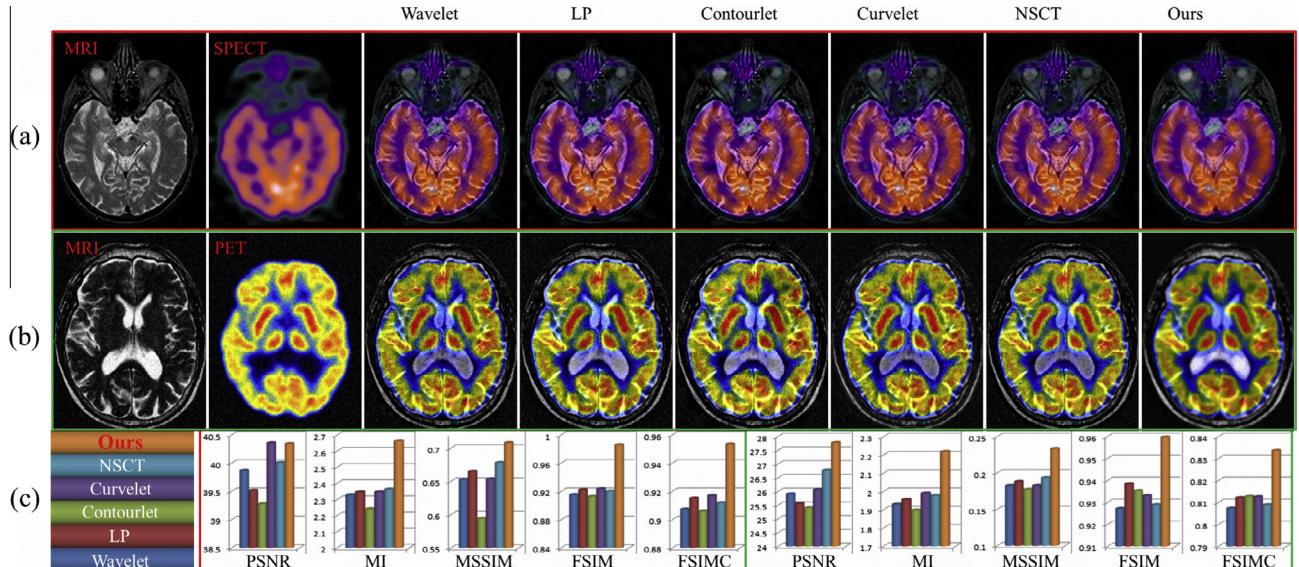


Fig. 14. The image fusion results and quantitative evaluation on white Gaussian noise-perturbed images (5 dB (a) and 25 dB (b)). (c) The quantitative evaluation, the left column is for (a) and the right column is for (b).

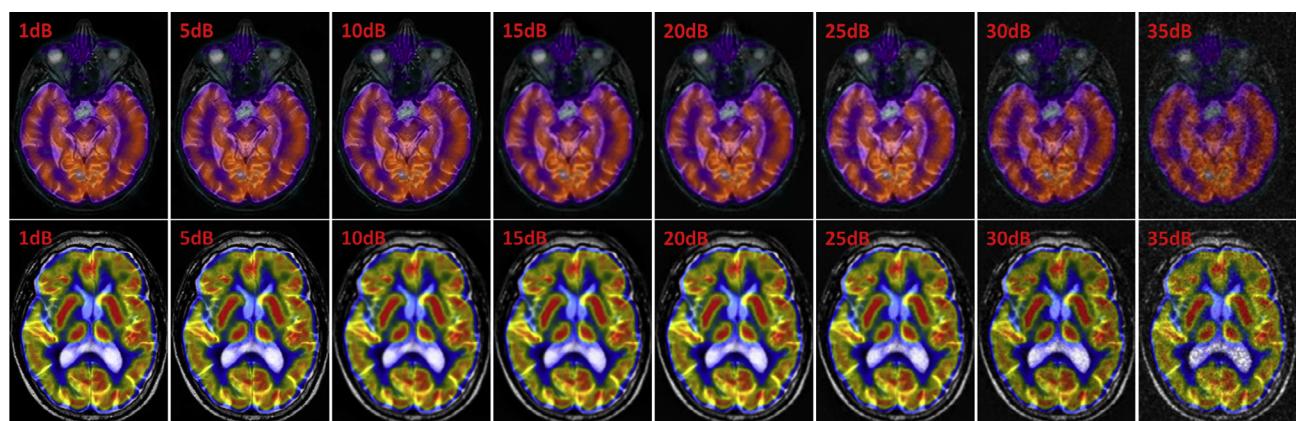


Fig. 15. The fusion results of our method on multi-level white Gaussian noise-perturbed images.

source image. Therefore, our method performs better on the noise-perturbed cases.

Moreover, to more objectively evaluate the robustness of the six methods, we measure their average performances according to five indicators (PSNR, MI, MSSIM, FSIM, and FSIMC) by taking into account the fusion results under different noise levels, wherein all the methods involved in the comparison are implemented by the authors themselves. Let us examine the statistic curves corresponding to the five quality indicators shown in Fig. 16, when increasing the noise level, the performances of other methods deteriorate rapidly, but the increasing noise has little negative influence on the performance of our method. In particular, Fig. 16(d) and (e) indicate that our method can preserve the intrin-

sic features even better when the noise level is up to 30 dB, which demonstrates the superiority of our method in robustness. Thus, our method fully outperforms such state-of-the-art methods at all the testing noise levels.

As for salt and pepper noise-perturbed images, Fig. 17 lists the fusion results when 6% and 10% pixels are perturbed by noise. The quantitative evaluation is documented in Fig. 17(c). We can notice that the competing methods are sensitive to noise and the performance degrades drastically for the noisy cases. In contrast, our method not only suppresses the salt and pepper noise but also fuses the complementary information well. The quantitative evaluation (Fig. 17(c)) also demonstrates that our method provides the best fusion quality. Similar to Fig. 16, Fig. 18 depicts the statistical

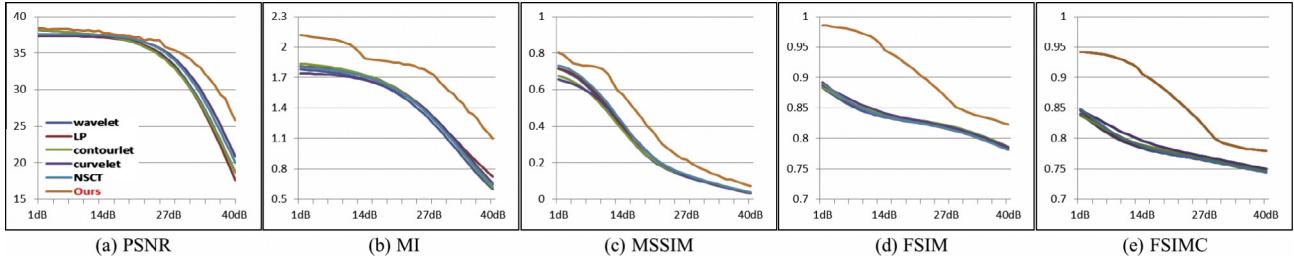


Fig. 16. Quantitative evaluation of average fusion quality on all the white Gaussian noise-perturbed images, wherein the horizontal coordinates represent adopted noise levels {1 dB, 2 dB, ..., 40 dB}.

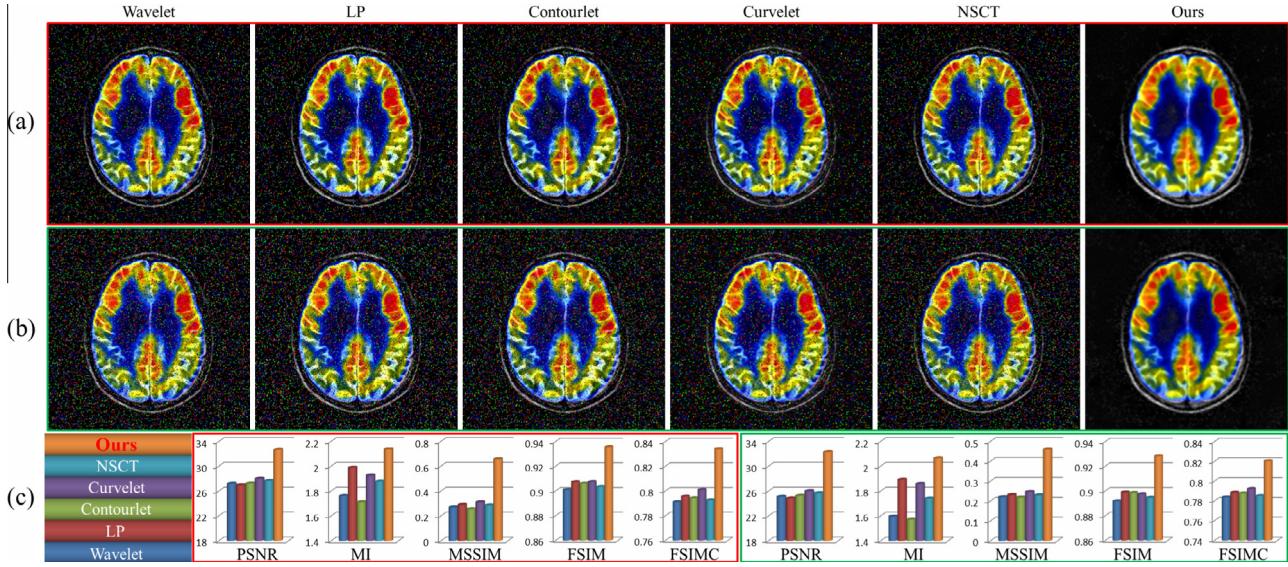


Fig. 17. The fusion results and quantitative evaluation on salt and pepper noise-perturbed images, wherein 6% (a) and 10% (b) pixels are respectively perturbed. (c) The quantitative evaluation, the left column is for (a) and the right column is for (b).

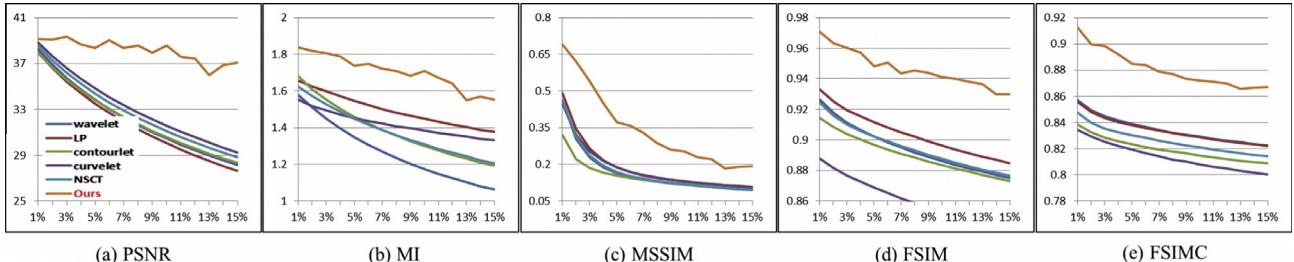


Fig. 18. Quantitative evaluation of average fusion quality on salt and pepper noise-perturbed images, wherein the percentages of noise-perturbed pixel are represented along the horizontal axis.

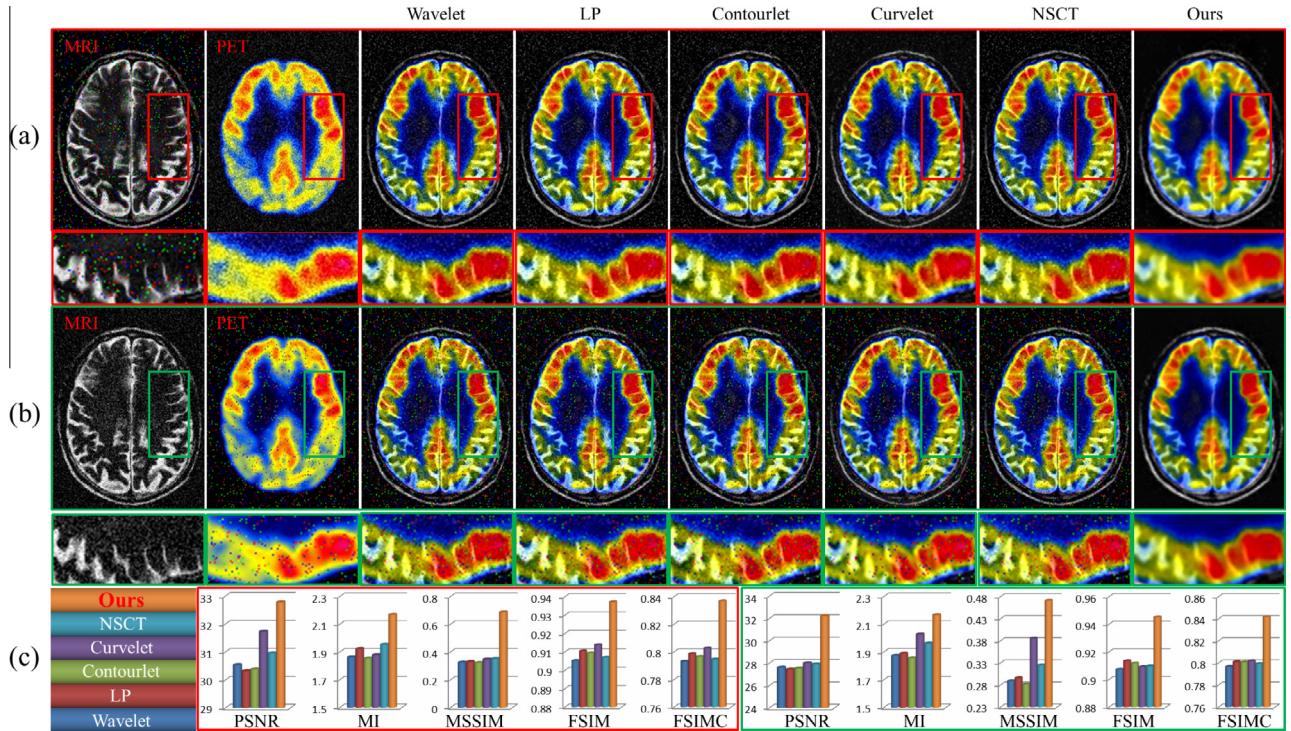


Fig. 19. The fusion results and quantitative evaluation for the noise-perturbed images of mixed-types (30 dB white Gaussian noise, 6% salt and pepper noise). (a) The fusion result of salt and pepper noise-perturbed MRI image and white Gaussian noise-perturbed PET image. (b) The fusion result of white Gaussian noise-perturbed MRI image and salt and pepper noise-perturbed PET image. (c) The quantitative evaluation, the left column is for (a) and the right column is for (b).

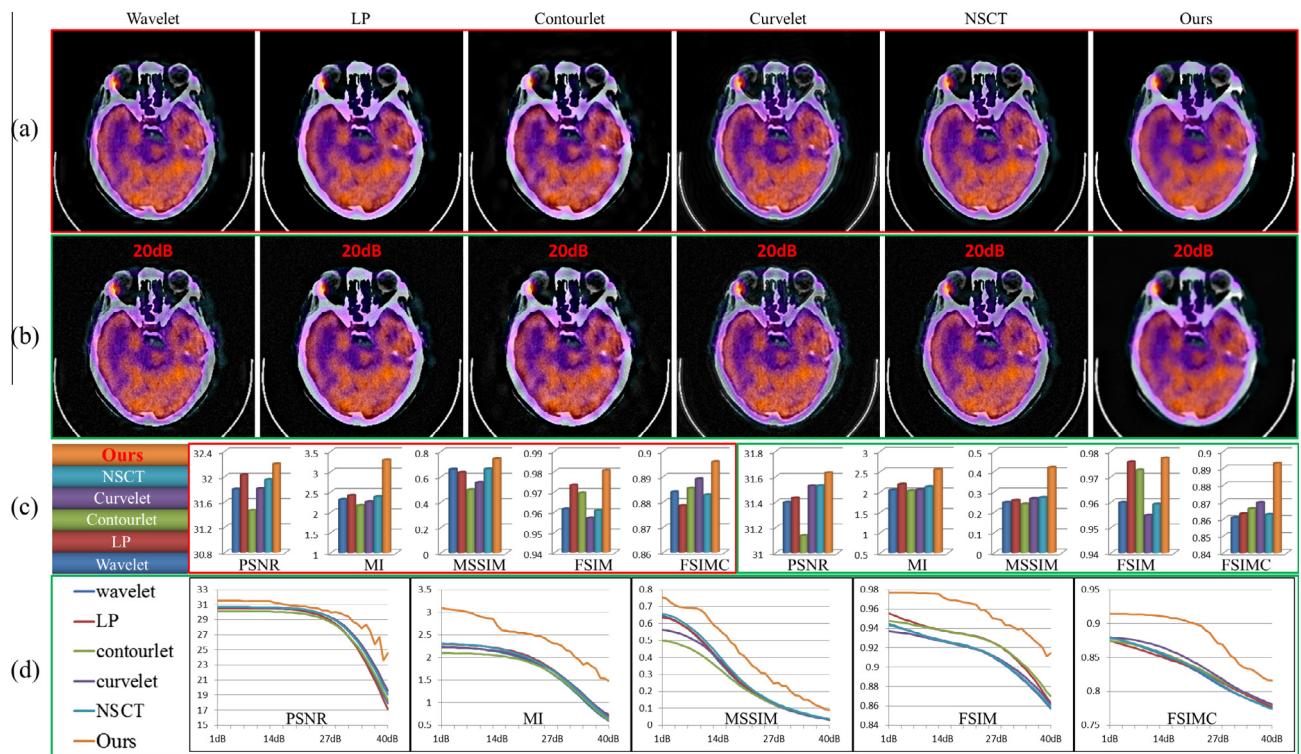


Fig. 20. CT/SPECT image fusion results and quantitative evaluation. (a) The fusion results of different methods on noise-free images. (b) The fusion results of different methods on 20 dB white Gaussian noise-perturbed images. (c) The quantitative evaluation for (a) and (b). (d) The quantitative evaluation for different methods on multi-level Gaussian noise-perturbed images.

curve of the average quality for fifteen noise levels which all indicate the robustness of our method in concert.

Meanwhile, to more comprehensively verify the robustness of our method, we test our method using different types of noise-perturbed images, wherein different-modality source images to be fused have different noise types. Fig. 19 presents the fusion results where one image has 30 dB white Gaussian noise and 6% pixels of another image is perturbed by salt and pepper noise. From the zoom-in effects shown in Fig. 19(a)–(b), it is obvious that the fusion results from other five methods have exhibited more noise,

which seriously decrease the visual quality. The quantitative evaluation (Fig. 19(c)) also indicates that our method gives rise to the better fusion quality, and thus our method is more robust to the noise mixture.

8.4. Other experiments and evaluation on versatility

To test the versatility of our method, we also conduct extensive fusion experiments on CT/SPECT and MRI-T1/MRI-T2 image pairs. Fig. 20(a) compares the CT/SPECT fusion results from different

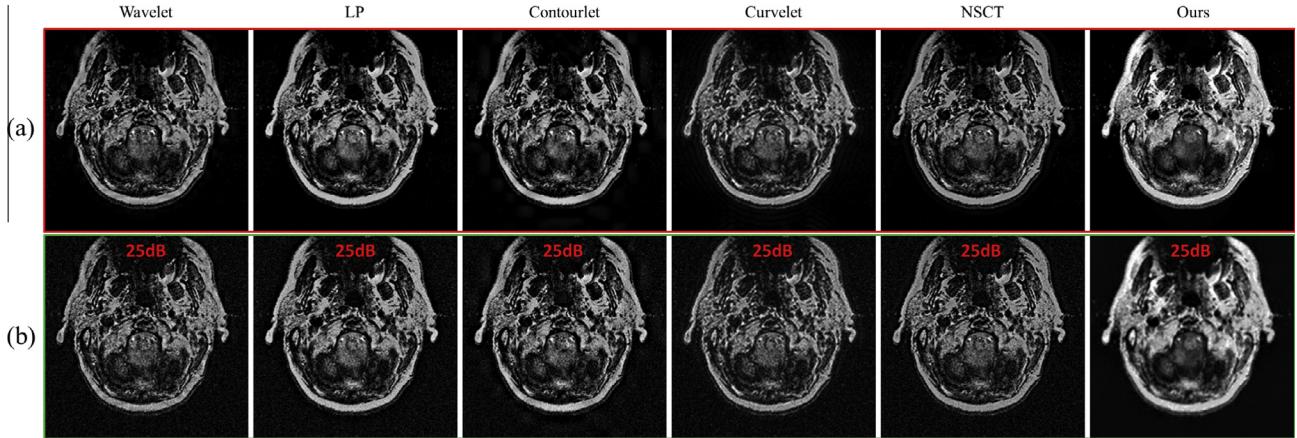


Fig. 21. MRI-T1 and MRI-T2 image fusion results. (a) The fusion results of different methods on noise-free images. (b) The fusion results of different methods on 25 dB white Gaussian noise-perturbed images.

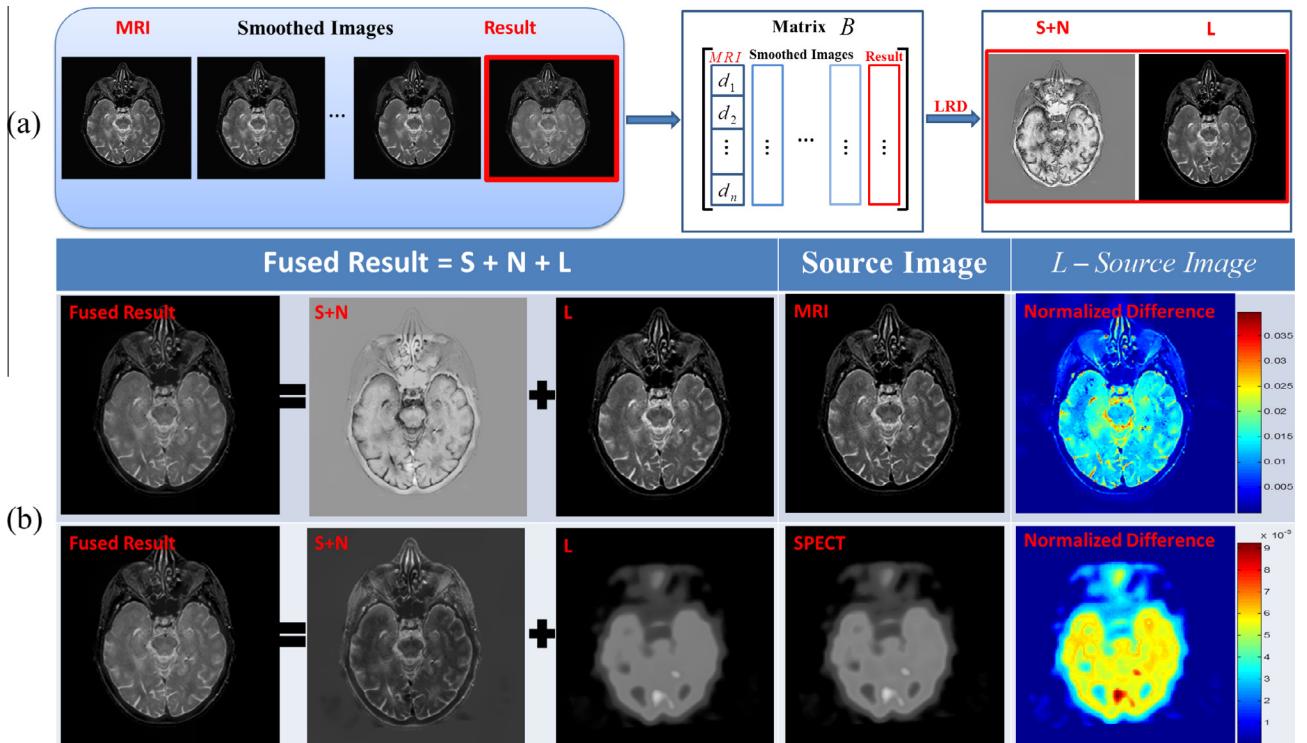


Fig. 22. Structure-preserving capability analysis of the fused result. (a) The principle of our analysis method. The fused result, the source image together with its multi-scale smoothed versions are reorganized to form a matrix B , and low-rank analysis is used to decompose the fused result into " $L + N + S$ ". L : low-rank component, $S + N$: sparse component and noise component. (b) In order to improve the visual effect, $S + N$ and difference images are normalized when performing visualization. Middle row: the low-rank component (L) preserves the salient structure in MRI source image, and " $S + N$ " retains the residuals of the extracted MRI salient information and the fused complementary information from SPECT source image. The amplitude difference, whose maximum value is 0.04, indicates that the proposed method can effectively preserve the salient information. Bottom row: the similar conclusions can be drawn for SPECT source image, wherein the maximum amplitude difference is 0.01.

methods. Fig. 20(b) shows the fusion results of 20 dB Gaussian noise-perturbed CT/SPECT image pairs. And the corresponding quantitative evaluation for Fig. 20(a)–(b) are respectively documented in Fig. 20(c), which indicates that our method outperforms the state-of-the-art methods in overall fusion quality. The average quantitative evaluation for 40 noise levels (Fig. 20(d)) demonstrates that our method possesses the better noise robustness than other methods. Similarly, Fig. 21(a) compares the MRI-T1/MRI-T2 fusion results from different methods, and Fig. 21(b) shows the fusion results of 25 dB Gaussian noise-perturbed MRI-T1/MRI-T2 image pairs. Comparing with other methods, our method can produce better fusion results, which has proper contrast and much more complementary information. Thus, our method can well accommodate multi-modal medical image fusion.

8.5. Evaluation for structure and salient information preservation

To verify the salient information-preserving capability of our method, we conduct sharp structure comparison between the fused image and source images in Fig. 22. Fig. 22(a) illustrates the principle of our low-rank salient structure abstraction based comparison analysis. Here, the fused result, one of the source images together with its multi-scale smoothed versions are reorganized to form a matrix B . Given the fused result, low-rank decomposition is employed to extract its involved common salient structure (L) with respect to current source image. The entire process will be executed twice, and the first one is for MRI source image and the second one is for SPECT source image. The first column of Fig. 22(b) shows the extracted three parts. It can be seen that the L component embedded in the fused image is very similar to its corresponding source image. And the $S + N$ component demonstrates the residual non-salient information from the current source image and the fused complementary information from the

other source image. Meanwhile, we also present the differences between the source image and the L component of the fused result in the right column of Fig. 22(b), wherein the maximum differences for MRI and SPECT are 0.04 and 0.01, respectively. It may be noted that, the illustrated difference is normalized to improve the visual effect. However, since the fusion operator necessarily impacts the sharpness of some salient edges where the complementary information is fused, it should also be noted that our method cannot guarantee to invariably preserve the amplitude of all the structures, and the difference in Fig. 22(b) clearly demonstrates this fact.

8.6. Evaluation for scale-specific fusion rule

To investigate the effects of our scale-specific rule, Fig. 23 respectively list the results obtained with the average–maximum rule and our rule, where we compare our result with NSCT-based method because it can achieve relatively better performance. Fig. 23(b) shows the results of our method and NSCT-based method when both of them adopt average–maximum rule. Comparing with the source images, we find that the results in Fig. 23(b) miss some anatomical information and appear too obscure. Fig. 23(c) shows the results of our method and NSCT-based method when using our scale-specific rule. Here it demonstrates that our fusion rule gives rise to larger MI both for our method and NSCT-based method, which means our rule can make an effective tradeoff when fusing the complementary information. Specially, by make comparison between the zoom-in effects in Fig. 23(b) and (c), it is obvious that our scale-specific fusion rule can maintain more proper contrast than average–maximum rule. Besides, even if using the same fusion rule, our results in Fig. 23(b) and (c) both have larger MI than those of NSCT-based method. Moreover, the NSCT-based method produces obvious errors (indicated with red

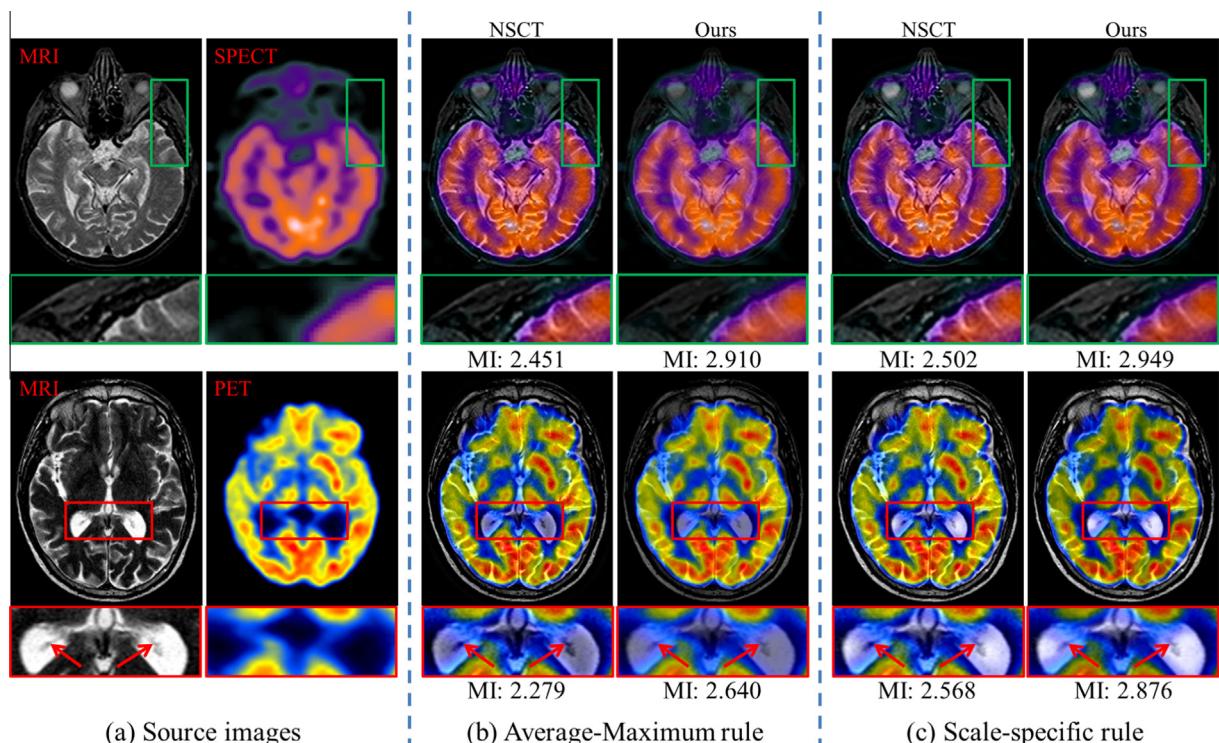


Fig. 23. Evaluation for scale-specific fusion rules. (a) Multi-modal source images. (b) Fusion results with Average–Maximum fusion rule. (c) Fusion results with our scale-specific fusion rule.

arrows) due to the limitations caused by its non-data-specific kernel, which also certifies the superiorities of our overall framework.

9. Conclusion and discussion

In this paper, we have developed a novel and comprehensive framework to address a suite of research challenges in multi-modal medical image fusion. The proposed technical solutions emphasize the physics-based data-specific multi-scale geometrical analysis to enable directional, structure-preserving image decomposition. Meanwhile, our method also focuses on the scale-aware salient information extraction from complementary multi-source images. In addition, our method offers an improved scale-specific fusion rule to adaptively optimize the contrast of fusion image, which can further improve the fusion quality. Extensive experiments on various medical images, together with their quantitative comparison with existing state-of-the-art fusion methods, have demonstrated the superior performance of our method. More importantly, the critical technical components of our method, including anisotropic heat diffusion over 2D manifold, multi-scale image decomposition, low-rank analysis model and reconstruction, divide-and-conquer algorithmic strategy, and CUDA parallel computation, can also contribute to other physics-based image processing applications (either individually or collectively).

However, our method still has some limitations that must be addressed. First, although our method is relatively robust, it may also fail to effectively handle the image containing massive noise. This is because the anisotropy of our method is solely determined by the local-structure encoded Laplacian matrix, which relies on the pixel intensity. Second, as documented in Table 1, although our method has shown high-quality performance, it is still time-consuming in principle. Specifically, when the dimension of image increases, the required computational load will increase rapidly. Third, our method cannot be directly used for sparse MR image fusion, because its valuable information may be presented as noise-like artifacts, while the noise tends to be compressed in our method. However, this problem may be solved by transforming the MR image from color space to certain high-dimensional feature space and then conducting fusion with our central framework. Thus, our ongoing research efforts are centered around these topics and their continuing improvement. Moreover, exploring new computer vision applications enabled by such techniques are also equally important and deserve more research endeavors.

Acknowledgments

This research is supported in part by National Natural Science Foundation of China (Nos. 61190120, 61190121, 61190125, and 61300067) and National Science Foundation of USA (Nos. IIS-0949467, IIS-1047715, and IIS-1049448). The authors would like to thank the anonymous reviewers for their helpful comments.

References

- [1] A. Goshtasby, S. Nikolov, Image fusion: advances in the state of the art, *Inf. Fusion* 8 (2) (2007) 114–118.
- [2] A. James, B. Dasarathy, Medical image fusion: a survey of the state of the art, *Inf. Fusion* 19 (2014) 4–19.
- [3] S. Paris, S.W. Hasinoff, J. Kautz, Local laplacian filters: Edge-aware image processing with a laplacian pyramid, *ACM Trans Graphics* 30 (2011) 68:1–68:12.
- [4] C.E. Heil, D.F. Walnut, Continuous and discrete wavelet transforms, *Soc. Ind. Appl. Math.* 31 (4) (1989) 628–666.
- [5] J. Hu, S. Li, The multiscale directional bilateral filter and its application to multisensor image fusion, *Inf. Fusion* 13 (3) (2012) 196–206.
- [6] J. Zhao, H. Feng, Z. Xu, Q. Li, T. Liu, Detail enhanced multi-source fusion using visual weight map extraction based on multi-scale edge preserving decomposition, *Opt. Commun.* 287 (2013) 45–52.
- [7] F. Nencini, A. Garzilli, S. Baronti, L. Alparone, Remote sensing image fusion using the curvelet transform, *Inf. Fusion* 8 (2) (2007) 143–156.
- [8] M. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, *IEEE Trans. Image Process.* 14 (12) (2005) 2091–2106.
- [9] A. Cunha, J. Zhou, M. Do, The nonsubsampled contourlet transform: theory, design, and application, *IEEE Trans. Image Process.* 15 (10) (2006) 3089–3101.
- [10] K. Guo, D. Labate, W. Lim, G. Weiss, E. Wilsonc, Wavelets with composite dilations and their mra properties, *Appl. Comput. Harmon. Anal.* 20 (2) (2006) 202–236.
- [11] A. James, S. Thiruvenkadam, J. Paul, M. Braun, Special issue on medical image computing and systems, *Inf. Fusion* 19 (2014) 2–3.
- [12] A. Wang, H. Sun, Y. Guan, The application of wavelet transform to multimodality medical image fusion, in: IEEE International Conference on Networking, Sensing and Control, 2006, pp. 270–274.
- [13] Q. Miao, C. Shi, P. Xu, M. Yang, Y. Shi, A novel algorithm of image fusion using shearlets, *Opt. Commun.* 284 (6) (2011) 1540–1547.
- [14] T. Tu, S. Su, H. Shyu, P. Huang, A new look at ihs-like image fusion methods, *Int. J. Inf. Fusion* 2 (3) (2001) 177–186.
- [15] S. Daneshvar, H. Ghassemian, Mri and pet image fusion by combining ihs and retina-inspired models, *Inf. Fusion* 11 (2) (2010) 114–123.
- [16] F. Ali, I. El-Dokany, A. Saad, F.A. ElSamie, A curvelet transform approach for the fusion of mr and ct images, *J. Mod. Opt.* 57 (4) (2010) 273–286.
- [17] L. Yang, B. Guo, W. Ni, Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform, *Neurocomputing* 72 (1–3) (2008) 203–211.
- [18] D. Labate, W. Lim, G. Kutyniok, G. Weiss, Sparse multidimensional representation using shearlets, *Wavelets XI* (San Diego, CA, 2005), vol. 5914, 2005, pp. 254–262.
- [19] X. Qu, J. Yan, H. Xiao, Z. Zhu, Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain, *Acta Automatica Sinica* 34 (12) (2008) 1508–1514.
- [20] Q. Zhang, B. Guo, Multifocus image fusion using the nonsubsampled contourlet transform, *Signal Process.* 89 (7) (2009) 1334–1346.
- [21] X. Yang, L. Jiao, Fusion algorithm for remote sensing images based on nonsubsampled contourlet transform, *Acta Automatica Sinica* 34 (3) (2008) 274–281.
- [22] T. Li, Y. Wang, Biological image fusion using a nsct based variable-weight method, *Inf. Fusion* 12 (2) (2011) 85–92.
- [23] L. Wang, B. Li, L. Tian, Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients, *Inf. Fusion* 19 (2014) 20–28.
- [24] P. Burt, E. Andelson, The Laplacian pyramid as a compact image code, *IEEE Trans. Commun.* 31 (4) (1983) 532–540.
- [25] W. Wang, F. Chang, A multi-focus image fusion method based on Laplacian pyramid, *Inf. Fusion* 6 (12) (2011) 2559–2566.
- [26] J. Scott, M.A. Pusateri, Laplacian based image fusion, in: IEEE 39th Applied Imagery Pattern Recognition Workshop, 2010, pp. 1–7.
- [27] T. Zhou, D. Tao, Godec: Randomized low-rank & sparse matrix decomposition in noisy case, in: International Conference on Machine Learning, 2011, pp. 33–40.
- [28] M. Yang, Background modeling from surveillance video using rank minimization, *Lect. Notes Comput. Sci.* 7350 (2012) 769–774.
- [29] E.J. Candès, X. Li, Y. Ma, J. Wright, Robust principal component analysis?, *J ACM* 58 (3) (2011) 11:1–11:37.
- [30] X. Zhou, W. Yu, Low-rank modeling and its applications in medical image analysis, in: SPIE Defense, Security, and Sensing: Signal, Image, and Neural Net Processing, vol. 8750, 2013, pp. 1–10.
- [31] H.M. Nguyen, X. Peng, M.N. Do, Z. Liang, Spatiotemporal denoising of mr spectroscopic imaging data by low-rank approximations, in: IEEE International Symposium on Biomedical Imaging, 2011, pp. 857–860.
- [32] E.J. Candès, C.A. Sing-Long, J.D. Trzasko, Unbiased risk estimates for singular value thresholding and spectral estimators, *IEEE Trans. Signal Process.* 61 (19) (2012) 4643–4657.
- [33] A.G. Christodoulou, S.D. Babacan, Z. Liang, Accelerating cardiovascular imaging by exploiting regional low-rank structure via group sparsity, in: IEEE International Symposium on Biomedical Imaging, 2012, pp. 330–333.
- [34] J. Cai, X. Jia, H. Gao, S.B. Jiang, Z. Shen, H. Zhao, Cine cone beam ct reconstruction using low-rank matrix factorization: algorithm and a proof-of-principle study, *IEEE Int. Med. Imaging* 33 (8) (2014) 1581–1591.
- [35] A. Rahmim, J. Tang, H. Zaidi, Four-dimensional (4d) image reconstruction strategies in dynamic pet: beyond conventional independent frame reconstruction, *Med. Phys.* 36 (8) (2009) 3654–3670.
- [36] T. Wan, C. Zhu, Z. Qin, Multifocus image fusion based on robust principal component analysis, *Pattern Recogn. Lett.* 34 (2013) 1001–1008.
- [37] D. Lee, B. Schachter, Two algorithms for constructing a delaunay triangulation, *Int. J. Comput. Inform. Sci.* 9 (3) (1980) 219–242.
- [38] J. Sun, M. Ovsjanikov, L. Guibas, A concise and provably informative multiscale signature based on heat diffusion, *Eurographics Symposium on Geometry Processing 2009*, vol. 28(5), 2009, pp. 1383–1392.
- [39] S. Li, Q. Zhao, S. Wang, T. Hou, A. Hao, H. Qin, Material-aware feature descriptor for volumetric image registration in diffusion tensor space, in: Proceedings of the 12th European Conference on Computer Vision, 2012, pp. 502–515.

- [40] S. Li, H. Qin, A. Hao, Multi-scale local features based on anisotropic heat diffusion and global eigen-structure, *Sci. China Inform. Sci.* 55 (10) (2013) 2234–2242.
- [41] F.M. Noguer, Deformation and illumination invariant feature point descriptor, in: IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 1593–1600.
- [42] U. Pinkall, K. Polthier, Computing discrete minimal surfaces and their conjugates, *Exp. Math.* 2 (1) (1993) 15–36.
- [43] M. Fazel, E. Candes, B. Recht, P. Parrilo, Compressed sensing and robust recovery of low rank matrices, in: Proc. 40th Asilomar Conference on Signals, Systems and Computers, 2008, pp. 1043–1047.
- [44] F. Wolfe, E. Liberty, V. Rokhlin, M. Tygert, A fast randomized algorithm for the approximation of matrices, *Appl. Comput. Harmon. Anal.* 25 (2008) 335–366.
- [45] S. Roweis, Em algorithms for pca and spca, in: Neural Information Processing Systems, 1998, pp. 626–632.
- [46] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (2004) 600–612.
- [47] L. Zhang, L. Zhang, X. Mou, Fsim: a feature similarity index for image quality assessment, *IEEE Trans. Image Process.* 20 (8) (2011) 2378–2386.