An efficient fusion algorithm combining feature extraction and variational optimization for CT and MR images

Qinxia Wang¹ | Xiaoping Yang²

¹School of Science, Nanjing University of Science & Technology, Nanjing, P.R.China
²Department of Mathematics, Nanjing University, Nanjing, P.R.China

Abstract
In medical image processing, image fusion is the process of combining complementary information from different or multimodality images to obtain an informative fused image in order to improve clinical diagnostic accuracy. In this paper, we propose a two-stage fusion framework for computed tomography (CT) and magnetic resonance (MR) images. First, the intensity and geometric structure features in both CT and MR images are extracted by the saliency detection method and structure tensor, respectively, and an initial fused image is obtained. Then, the initial fused image is optimized by a variational model which contains a fidelity term and a regularization term. The fidelity term is to retain the intensity of the initial fused image, and the regularization term is to constrain the gradient information of the fused image to approximate the MR image. The primal-dual algorithm is proposed to solve the variational problem. The proposed method is applied on five pairs of clinical medical CT and MR-T1\MR-T2 images, and the comparison metrics SF, MI, Q₁₀₀/F, Q_w, and VIFF are calculated for assessment. Compared with seven state-of-the-art methods, the proposed method shows a comprehensive advantage in preserving the salient intensity features, as well as texture structure information, not only in visual effects but also in objective assessments.

KEY WORDS
image fusion, primal-dual algorithm, saliency detection, structure tensor, variational model

1 | INTRODUCTION

The rapid development of computer technology promotes the progress of medical imaging, clinic practices, and medical science. Different imaging methods are complementary and provide more information than any individual among those. For example, computed tomography (CT) can capture the high density bone structures, while magnetic resonance imaging (MRI) can provide clear soft tissue structures in organs, such as brain, heart, and liver. Medical image fusion is the process of combining various information from multimodality medical images to obtain a fused image in order to increase clinical applicability for medical diagnosis and treatment of medical problems. Medical image fusion methods have been applied in human organs, including brain, breast, liver, and other organs.¹²

Researchers have paid more attention to fuse multimodal medical images in order to obtain a single image with more information for physicians. Medical image fusion can be performed through the multiscale transform (MST) methods.³⁴ The laplace pyramids (LP)⁵⁶ and wavelet transforms (WT)⁷⁻⁹ are two kinds of classical MST methods. The multiscale geometric analysis (MGA) tools have been employed in image fusion, including curvelet transform,¹⁰ non-subsampled contourlet transform (NSCT),¹¹¹² and non-subsampled shearlet transform (NSST).¹³¹⁴ Moreover, the MST methods can extract difference in resolution information, which can be used to combine
with other image processing methods, such as sparse representation (SR)\textsuperscript{15} and pulse coupled neural network (PCNN).\textsuperscript{16}

Two-scale image fusion methods can decompose each input image into a base layer and a detail layer through a filter method. The guided filter-based fusion method (GFF)\textsuperscript{17} first extracts the average contrast to build the weight maps, then the guided filter is used to refine the weight maps of the two layers, and the final fused image is generated by the weighted average method. In Ref. \[18\], the authors established a gradient domain guided filter-based fusion model, and the weight maps are obtained by combing multiple visual features including contrast, sharpness, and structure saliency measures.

Medical image fusion can also be regarded as optimization problems. The variational image fusion methods always contain two terms. The first term is a fidelity norm used to constrain the gray approximation between the fused image and input images, and the second term is a regularization constraint. In Ref. \[19–21\], the authors utilized the TV norm as the regularization term, while the fidelity term is chosen to be the quadratic norm or $L_1$ norm. As a higher-order regularization constraint, the total generalized variation (TGV) is applied for image fusion as it can suppress the staircasing effect.\textsuperscript{22,23}

As two of most commonly used medical imaging modalities, CT and MR images are frequently combined for medical image analysis. CT images are density imaging, and can capture bone structures with high intensity. MR images can capture clear soft tissue structure information, especially the tumors which have high intensity in MR-T2 images. For CT and MR image fusion, the fusion strategies of most existing methods usually extract the same features from source images and use them to construct weight maps which may lead to the “feature average” problem. In order to extract different and complementary features from source CT and MR images, we propose a novel fusion method which can keep these extracted features enhancement in the fused images. For example, the proposed method can extract not only the high intensity features (such as bone structures and tumor) from CT and MR brain images, but also the soft tissue texture structures from MR images. As a result, we can obtain fused images which preserve the bone structures, tumor features, as well as the soft tissue texture information in CT and MR images.

In this study, a two-stage fusion framework is proposed for CT and MR images. First, the contrast-based method and structure tensor are used to extract salient intensity information and geometry structures of medical images, respectively. The extracted intensity and structure features are used to construct weight maps and obtain an initial fused image. Then, a variational model is proposed to optimize the initial fused image. The variational framework consists of two terms: the fidelity term which is used to retain the saliency intensity information, and the regularization term which is used to constrain the approximation of gradient information between the fused image and MR image, and keep the fused image smooth at the same time. In the numerical experiments, the proposed method is compared with seven state-of-the-art medical fusion methods on several pairs of CT and MR images. The experimental results show that the proposed method can well preserve the clear bone structures, tumor features, as well as the soft tissue texture information from the source images, and get a good visual quality.

The rest of the paper is organized as follows. In Section 2, we introduce some related works about saliency detection and structure tensor. In Section 3, an initial fusion image is generated with the weighted fusion method, and then it is optimized by the proposed variational fusion model. Numerical experiments and discussion are shown in Section 4, and the fusion performance is evaluated by visual observation and quantitative evaluation. The conclusion is given in Section 5.

\section{RELATED WORKS}

\subsection{Saliency detection}

Human vision perception system is highly sensitive to contrast in visual signals, such as color, intensity, and texture. Based on this observation, the saliency values of image pixels are defined by using color statistics of the input images.\textsuperscript{24,25} The saliency value of a pixel $i_k$ in an image $l$ is

$$S(l_k) = \sum_{i \in l} D(l_k, l_i)$$

where $D(l_k, l_i)$ denotes the color distance metric, and $l_i$ calculated by the absolute value of the color difference between pixels $l_k$ and $l_i$. The saliency of a pixel is defined by using its color contrast to all other pixels in the image.

For gray input images, the color distance metric $D(l_k, l_i)$ is calculated by the intensity difference. When using the intensity values to rearrange the pixels of the image $l$, the pixels with the same intensity value $c_i$ are grouped together, and they have the same saliency value $S(c_i)$.

Let the intensity value of the pixel $l_k$ be equal to $c_i$, then the Eq. (1) can be rewritten as

$$S(l_k) = S(c_i) = \sum_{-0}^{n} f_i D(c_i, c_j)$$

where $n$ is the number of intensities and $f_i$ is the probability of the pixel intensity value $c_i$ appearing in the image $l$. The saliency value $S(c_i)$ is calculated by the gray difference, and the pixels with high intensity have high contrast. So, the high intensity regions can be detected globally in the saliency map $S(l)$. In Fig. 1, we compare the saliency maps obtained by the different methods. Fig. 1(a) is a medical MR-T2 image $l$, (b) is the saliency map generating by the intensity distance metric, (c) is the distance between $l$ and its mean value, (d) and (e) are their binary map after a threshold method. By comparing (d) and (e) in Fig. 1, we can observe that there are many unexpected blocks in the map $C_w(l)$ for detecting the high-intensity regions of the source MR-T2 image.

\subsection{Structure tensor}

Given a gray image $l: \Omega \rightarrow [0, 1]$, for every pixel $x = (x_1, x_2)$, and $l(x)$ is the intensity of the input image $l$ at $x$. The gradient $\nabla l$ is defined as follows to describe the contrast change
\( I(x, y) = \nabla I^2 = \left( \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)^T \),
\( I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y} \)
\( \lambda_1, \lambda_2 \) are eigenvalues, \( \theta_1 \) and \( \theta_2 \) are the corresponding eigenvectors. The eigenvalues \( \lambda_1 \) and \( \lambda_2 \), respectively represent maximum and minimum contrast change rate of the gray image \( I \). The gradient at pixel \( p \) can be calculated by
\( ST(p) = \theta_1(p)\sqrt{(s_1 + s_2)^2 + (s_1 - s_2)^2}, \)
where \( s_1 = \sqrt{\lambda_1(p)}, s_2 = \sqrt{\lambda_2(p)} \) and \( \alpha \) is a positive parameter. In Fig. 2, we calculate the gradient values in different directions.

3 | PROPOSED FUSION METHOD

In this section, we describe the implementation of the proposed medical image fusion method in details. The goal of the proposed method is to preserve the salient intensity information and clear soft tissue structures from CT and MR images. First, the contrast based method and structure tensor are used to capture the multiple intensity and structure features, and the weighted multiple features fusion is performed. Then the generated image is used as an initial input of the variational model, and the final fused image can be obtained by solving the variational model using the optimization algorithm.

3.A | Weighted fusion based on multiple features

In image processing, the contrast feature can be used to detect the saliency intensity information, and the structure tensor is an effective tool to describe image geometry structures. In this section, the contrast and structure features are detected to generate the weight maps for weighted image fusion. Given two input images \( I_{ct} \) and \( I_{mr} \), the weighted fusion process is shown in Fig. 3, and the main steps to get the initial fused image \( F_0 \) are described in the following.
First, we use Eq. (2) to detect the high-intensity regions of the input images, and the saliency maps $S(I_{ct})$ and $S(I_{mr})$ are segmented to generate two binary weight maps $S_w^1$ and $S_w^2$. Meanwhile, the structure tensor is used to calculate the structure feature maps $ST(I_{ct})$ and $ST(I_{mr})$, and the two structure weight maps $STw^1$ and $STw^2$ can be obtained by normalization, that is,

$$STw^1 = \frac{ST(I_{ct})}{ST(I_{ct}) + ST(I_{mr})}, \quad STw^2 = \frac{ST(I_{mr})}{ST(I_{ct}) + ST(I_{mr})}, \quad (9)$$

The goal of the weighted fusion is to extract the high-intensity regions from the input medical images, and preserve the structures of other regions, so we calculate the final weight maps as follows:

$$w_1 = \begin{cases} 1 & \text{if } S_w^1 \neq 0 \text{ and } S_w^2 \neq 0 \\ 0 & \text{if } S_w^1 = 0 \text{ and } S_w^2 \neq 0, \quad w_2 = \begin{cases} 1 & \text{if } S_w^2 \neq 0 \text{ and } S_w^1 = 0 \\ 0 & \text{if } S_w^2 = 0 \text{ and } S_w^1 \neq 0 \\ STw_2 & \text{otherwise} \end{cases} \end{cases} \quad (10)$$

Finally, the weight maps $w_1, w_2$ are used to fuse the input medical images $I_{ct}$ and $I_{mr}$, and the initial fused image $F_0$ is generated by the weighted average method, that is,

$$F_0 = w_1 \cdot I_{ct} + w_2 \cdot I_{mr}. \quad (11)$$

### 3.B | Variation-based fusion

Given the image domain $\Omega \subset \mathbb{R}^2$ and a pair of input images $u_0, u_1$, we propose the variational fusion method through a minimization problem:

$$\arg \min_u E(u) = \frac{\alpha}{2} \int_\Omega (u - u_0)^2 dx + \int_\Omega |\nabla u - \nabla u_1| dx. \quad (12)$$

Here, the final fused image and two initial input images are denoted to be $u, u_0,$ and $u_1$, respectively. The first term is a square norm, and it is used to constrain the intensity of the fused image $u$ to approximate the input image $u_0$. The second term is a regularization term to constrain the fused image $u$ smooth, and preserve the gradient information of $u_1$ at the same time.

In Section 3.A, the high-intensity information is well preserved through the weighted-fusion, while the tissue structures are partially missing. In Figs. 4(a) and 4(b) are two input CT and MR-T2 images, and Fig. 4(c) are the weighted-fusion result $F_0$. Compared with the input MR image, the gray matter and white matter of the fused image $F_0$ are blurred, and some noise is introduced into $F_0$ because of different gray distributions of CT and MR images. When the TV norm is used as the regularization term, we get a smoothed result shown in Fig. 4(d). Compared with CT images, the MR images can capture clearer soft tissue texture structures. The differences between soft tissues are more distinct, such as the intensity difference between the gray matter and white matter of human brain. The contrast of soft tissues in MR images is salient. So, the structure features of the MR image $I_{mr}$ are preserved in the final fused image, and the variational model (12) can be rewritten as

![Fig. 3. The weighted fusion process.](image-url)
\[ \arg \min_u E(u) = \frac{\alpha}{2} \int (u - F_0)^2 dx + \int_\Omega |\nabla u - \nabla I_{\text{mr}}|^2 dx \]  

(13)

The solution of problem (13) is obtained as the final fused result.

3.C | Algorithm

The proposed variational model (13) is convex, and can be solved by the primal-dual algorithm. The minimization problem (13) can be converted to a saddle-point problem which is derived by duality principles:

\[
\min_u \max_p \left\{ \frac{\alpha}{2} \|u - F_0\|^2_2 + \langle \nabla u - \nabla I_{\text{mr}}, p \rangle - l_{|p|_{\text{L}_1}} - \frac{1}{2\sigma} \|p - \tilde{p}\|^2_2 \right\}.
\]  

(14)

Let \((\tilde{p}^{t+1}, u^{t+1})\) be approximating solutions in the iteration process, then we have

\[
p^{t+1} = \arg \max_p \left\{ \langle \nabla u - \nabla I_{\text{mr}}, p \rangle - l_{|p|_{\text{L}_1}} - \frac{1}{2\sigma} \|p - \tilde{p}^t\|^2_2 \right\}.
\]  

(15)

\[
u^{t+1} = \arg \min_u \left\{ \frac{\alpha}{2} \|u - F_0\|^2_2 + \langle \nabla u - \nabla I_{\text{mr}}, p \rangle + \frac{1}{2\tau} \|u - u^t\|^2_2 \right\}.
\]  

(16)

By applying the proximal point algorithm, the proposed model can be easily solved as follows:

Algorithm 1: Numerical Algorithm:

- **Initialization**: Choose \(\sigma > 0, \tau > 0, \bar{u}, \tilde{p}, p^0\) and set \(\bar{u}^0 = 0, \bar{p}^0 = 0\).
- **Iteration** (\(t \geq 0\)): Update \(u^{t+1}, \tilde{u}^{t+1}, \bar{u}^{t+1}, \bar{p}^{t+1}\) as follows

\[
\bar{p}^{t+1} = p^t + \sigma (\nabla u^t - \nabla I_{\text{mr}}).
\]

\[
p^{t+1} = \max\{1, \bar{p}^{t+1}\}.
\]

\[
\tilde{u}^{t+1} = \frac{u^t + \sigma \bar{u}^{t+1} + \alpha F_0}{1 - \sigma},
\]

\[
u^{t+1} = 2\tilde{u}^{t+1} - u^t.
\]  

(17)

- **Output**: \(u^*\).

4 | NUMERICAL EXPERIMENTS AND DISCUSSION

Experiments are implemented on the clinical medical CT and MR images from the whole brain atlas website (http://www.med.harvard.edu/anilb). Five pairs of CT and MR-T1\_MR-T2 images of human brain are experimented to evaluate fusion performance. Each pair of images is sampled from the same slice plane and coregistered. The size of the images is 256 by 256. Besides, all the experiments are performed in MATLAB R2015b running on an Intel Core i7-5500U CPU 2.40 GHz with 4 GB of RAM.

4.A | Visual observation

To evaluate the fusion performance, the fusion result of the proposed method is compared with seven state-of-the-art fusion methods which are TGV, GTF, MGGFF, GFF, PAPCNN, MFDF-NSS, and LP-SR based methods, respectively.

As shown in Figs. 5(a) and 5(b) are a pair of input CT and MR-T2 images, and Figs. 5(c)–5(j) are the fusion results of the proposed method, TGV, GTF, MGGFF, GFF, PA-PCCNN, MFDF-NSS, and LP-SR based methods, respectively. The fusion results of TGV and GTF based methods lose some intensity information, which has led to low visual quality [see Figs. 5(d) and 5(e)]. Some details are blurred in the fused images of MGGFF and MFDF-NSS based methods, and the corresponding regions are marked with red arrows as shown in Figs. 5(f) and 5(h). In Fig. 6, the gradients of the fusion results are calculated, and the proposed method can extract smooth and clear structure information from the input images [see Fig. 6(c)].

As shown in Figs. 7(a) and 7(b) are a pair of input CT and MR-2 images, and there exists a lesion which is reflected as one high-signal intensity region in the MR-T2 image (the red rectangle). As shown in Figs. 7(c), 7(f), and 7(i), the fused images obtained by the proposed method, MGGFF and PAPCNN based methods preserve the lesion region and bone structures well without reducing the intensity contrast. The lesion regions of the fused images in Figs. 7(c)–7(j) are enlarged, and the edges of the lesion region are clearer in the fused image obtained by the proposed method [see Fig. 8(a)].

In Figs. 9(a) and 9(b) are the input CT and MR-T2 images. The proposed method can extract clear detail information from the source images. The fusion result of GFF based method preserves the soft tissues with reduced contrast. As shown in Figs. 9(h)–9(j), the detail textures are blurred in the fused images of MFDF-NSS, PAPCNN, and LP-SR based methods. Another set of brain CT and MR-T2 images is experimented, and the performance comparison results are shown in Fig. 10.
Figure 11 shows the fusion results of CT and MR-T1 images, and the proposed method can get better visual result. The experimental results in this section demonstrate advantages of the proposed method in preserving bone structures and clear soft tissues from source CT and MR images.

4.B | Quality assessment

For objective quality assessment, we use five metrics to evaluate the fusion quality, which are spatial frequency (SF),\(^{28}\) mutual information (MI),\(^{29}\) gradient based fusion metric \(Q^{G} \),\(^{30}\) the weighted fusion quality metric \(Q^{W} \),\(^{31}\) and visual information fidelity fusion (VIFF).\(^{32}\)

The five image quality metrics are introduced as follows:

- **Spatial Frequency (SF)**

  \[
  SF = \sqrt{RF^2 + CF^2},
  \]

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

- **Mutual Information (MI)**

  MI measures the degree of dependency between two variables X and Y, and it is defined as follows:
Fig. 7. Performance comparison in the "Med-2" experiment.

Fig. 8. The zoomed region of fusion result images in Fig. 7.

Fig. 9. Performance comparison in the 'Med-3' experiment.
where $x$ and $y$ are sampling variables, $p_{XY}$ is the joint probability distribution for $X$ and $Y$, $p_X$ and $p_Y$ are the probability distribution for $X$ and $Y$, respectively. In addition, $H(X) = I(X,X)$ is the entropy of $X$. For the fused image $F$ and input images $A$ and $B$, we calculate the two mutual information $I(F,A)$ and $I(F,B)$, then the normalized mutual information is calculated by

$$M_{AB} = 2 \left( \frac{I(F,A)}{H(F)+H(A)} + \frac{I(F,B)}{H(F)+H(B)} \right).$$  

The metric $Q^{AB}/F$ measures the relative amount of edge information that is transferred from the input images $A$ and $B$ into the fused image $F$.
where $Q_{AF}^{i,j}$ and $Q_{BF}^{i,j}$ are the edge preservation values at the pixel location $(i,j)$. $W_{AF}(i,j)$ and $W_{BF}(i,j)$ are the weights, which indicate the importance of $Q_{AF}^{i,j}$ and $Q_{BF}^{i,j}$, respectively.

The metric $Q_{W}$ is calculated to measure how well the structural information is preserved from input images. The metric is first calculated in a local region.

$$Q_{W}(A, B, F) = \sum_{w \in W_C(w)} \lambda(w) Q_0(A, F|w) + (1 - \lambda(w)) Q_0(B, F|w),$$

where $Q_0$ is the universal image quality index, $\lambda(w)$ is a local weight indicating the relative importance of image $A$ compared to image $B$.

The metric VIFF is a multiresolution image fusion metric using visual information fidelity. The metric measures the effective visual information of the fusion in all blocks in each sub-band. The metric value is calculated by integrating all the information in each sub-band.
For SF, MI, $Q^{AB/F}$, $Q_W$, and VIFF, they reflect different information including image spatial frequency, the correlation between the fused image and input images in gray level, and gradient level, the structure similarity and the visual information fidelity. For the five metrics, the values are bigger, the fusion results are better. As shown in Figs. 12 and 13, the proposed method has the best MI and $Q^{AB/F}$ values, which indicates the fused image of the proposed method preserve better gray and gradient features from source medical CT and MR images.

For CT and MR images fusion, there may be artifacts in the fused image because of different gray distributions. The proposed method uses the optimization model to reduce the artifacts, so the fused results turn to be a bit smooth, and the SF values of the proposed method are lower compared with MFDF-NSST based method (see Fig. 14). As shown in Fig. 15, the values of the metric $Q_W$ are too close and thus it is difficult to estimate image quality difference. The comparison of the metric VIFF is shown in Fig. 16; the proposed method has higher values compared with the seven fusion methods.
In Table 1, the average running times of different methods are calculated. For LP-SR based method, the running time is the smallest as the sparse dictionary untrained. The GFF based method costs less time because it only extracts a contrast feature for image fusion.

### 4.C | Discussion

For the visual observation, the proposed method can extract the salient features from source CT and MR images, such as the high-intensity bone structures, tumor features, and soft tissue texture information, and these salient features are well preserved in the fused images, while the seven compared methods tend to generate average or blurred features in the process of fusion. For quantitative quality assessment, the proposed method has the highest MI and $Q^{AF}$ values. The SF and VIFF values are higher because of the smoothing effect in the fusion process. In summary, the proposed method outperforms in intensity contrast preservation and texture structure extraction. In addition, our experiments are performed on brain images, and other images will be processed in the future work.

### 5 | CONCLUSION

In this paper, we propose a two-stage fusion method for CT and MR images. The intensity and structure features are extracted from source images for initial weighted fusion, and then a variational model is applied to optimize the initial fusion result. The comparison between the proposed method and seven state-of-the-art fusion methods are performed in the numerical experiments, and five comparison metrics are used for objective assessment. The experimental results and discussion show that the proposed method can well preserve the intensity and texture structure information from source images, and get a good visual effect.

### ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (No. 11531005, No. 91630311).

### CONFLICT OF INTEREST

No conflicts of interest.

### REFERENCES

1. James AP, Dasarathy BV. Medical image fusion: a survey of the state of the art. Inf Fusion. 2014;19:4–19.
2. Du J, Li W, Lu K, Xiao B. An overview of multi-modal medical image fusion. Neurocomputing. 2016;215:3–20.
3. Li S, Yang B, Hu J. Performance comparison of different multi-resolution transforms for image fusion. Inf Fusion. 2011;12:74–84.
4. Liu Y, Liu S, Wang Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf Fusion*. 2015;24:147–164.

5. Burt P, Adelson E. The Laplacian pyramid as a compact image code. *IEEE Trans Commun*. 1983;31:532–540.

6. Toet A. Multiscale contrast enhancement with applications to image fusion. *Opt. Eng.*. 1992;31(5):1026–1032.

7. Pajares G, De La Cruz JM. A wavelet-based image fusion tutorial. *Pattern Recognit.* 2004;37:1855–1872.

8. Lewis JJ, O'Callaghan RJ, Nikolov SG, Bull DR, Canagarajah N. Pixel and region-based image fusion with complex wavelets. *Inf Fusion*. 2007;8:119–130.

9. Na Y, Zhao L, Yang Y, Ren M. Guided filter-based images fusion algorithm for CT and MRI medical images. *IET Image Process*. 2017;12(1):138–148.

10. Nencini F, Garzelli A, Baronti S, Alparone L. Remote sensing image fusion using the curvelet transform. *Inf Fusion*. 2007;8:143–156.

11. Bhatnagar G, Wu QJ, Liu Z. A new contrast based multimodal medical image fusion framework. *Neurocomputing*. 2015;157:143–152.

12. Zhu Z, Zheng M, Qi G, Wang D, Xiang Y. A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain. *IEEE Access*. 2019;7:20811–20824.

13. Kong W. Technique for gray-scale visual light and infrared image fusion based on nonsubsampled shearlet transform. *Infrared Phys Technol*. 2014;63:110–118.

14. Liu X, Mei W, Du H. Multi-modality medical image fusion based on image decomposition framework and nonsubsampled shearlet transform. *Biomed Signal Process Control*. 2018;40:343–350.

15. Zhang Q, Liu Y, Blum RS, Han J, Tao D. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review. *Inf. Fusion*. 2018;40:57–75.

16. Yin M, Liu X, Liu Y, Chen X. Medical image fusion with parameter-adaptive pulse coupled neural network in nonsampled shearlet transform domain. *IEEE Trans Instrum Meas*. 2019;68:49–64.

17. Li S, Kang X, Hu J. Image fusion with guided filtering. *IEEE Trans Image Process*. 2013;22:2864–2875.

18. Yang Y, Que Y, Huang S, Lin P. Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion. *IEEE Trans Instrum Meas*. 2017;66:691–703.

19. Wang WW, Shui PL, Feng XC. Variational models for fusion and denoising of multifocus images. *IEEE Signal Process Lett*. 2008;15:65–68.

20. Yuan J, Miles B, Garvin G, Tai XC, Fenster A. Efficient convex optimization approaches to variational image fusion. *Num Math*. 2014;7:234–250.

21. Ma J, Chen C, Li C, Huang J. Infrared and visible image fusion via gradient transfer and total variation minimization. *Inf Fusion*. 2016;31:100–109.

22. Pock T, Zebedin L, Bischof H. TGV-fusion. In: *Rainbow of Computer Science*. Berlin, Heidelberg: Springer; 2011:245–258. https://doi.org/10.1007/978-3-642-19391-0_18

23. Wang Q, Yang X. Variational image fusion approach based on TGV and local information. *IET Comput Vis*. 2018;12:535–541.

24. Cheng MM, Mitra NJ, Huang X, Torr PH, Hu SM. Global contrast based salient region detection. *IEEE Trans Pattern Anal Mach Intell*. 2014;37:569–582.

25. Zhai Y, Shah M. Visual attention detection in video sequences using spatiotemporal cues. Paper presented at: the 14th ACM international conference on Multimedia. October; 2006. Santa Barbara CA, USA.

26. Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imaging Vis*. 2011;40:120–145.

27. Johnson KA, Becker JA. Whole brain website of the harvard medical school. http://www.med.harvard.edu/aanlib/. Accessed April 13, 2018.

28. Li S, Kwok JT, Wang Y. Combination of images with diverse focuses using the spatial frequency. *Inf Fusion*. 2001;2:169–176.

29. Hossny M, Nahavandi S, Grehton D. Comments on ‘Information measure for performance of image fusion’. *Electron Lett*. 2008;44:1066–1067.

30. Xydeas C, Petrovic V. Objective image fusion performance measure. *Electron. Lett.*. 2000;36:308–309.

31. Piella G, Heijmans H. A new quality metric for image fusion. Paper presented at: Proceedings 2003 International Conference on Image Processing; September 14-17, 2003; Barcelona, Spain.

32. Han Y, Cai Y, Cao Y, Xu X. A new image fusion performance metric based on visual information fidelity. *Inf Fusion*. 2013;14:127–135.