Image Fusion Using Quaternion Wavelet Transform and Multiple Features

Pengfei Chai, Xiaoqing Luo and Zhancheng Zhang

Abstract—**Multi-scale based Image fusion is one of main fusion methods, in which multi-scale decomposition tool (MSD) and feature extraction play very important roles. The quaternion wavelet transform (QWT) is one of effective multi-scale decomposition tools. Therefore, this study proposes a novel multimodal image fusion method using QWT and multiple features. First, we perform QWT on each source images to obtain low frequency coefficients and high frequency coefficients. Second, a weighted average fusion rule based on the phase and magnitude of low frequency subband and spatial variance is proposed to fuse the low frequency subbands. Next, a choose-max fusion rule based on the contrast and energy of coefficient is proposed to integrate the high frequency subbands. Finally, the final fused image is constructed by inverse QWT. The proposed method is conducted on multi-focus images, medical images, infrared-visible images and remote sensing images, respectively. Experimental results demonstrate the effectiveness of the proposed method.**

Index Terms—**image fusion, quaternion wavelet transform, phase, magnitude, feature**

I. INTRODUCTION

IN recent years, image fusion technology becomes more and more important, which has been widely used in many fields such as multi-focus image[1-3], medical image[4, 5], infrared-visible image[6] and remote sensing image[7]. The purpose of image fusion is to combine images containing the same scene from different sensors to generate a more comprehensive and accurate image including all useful information of these source images.

The current image fusion methods are divided into two categories. One is to directly fuse source images in spatial domain. However, this kind of methods is not good in dealing with edge. The other one is to integrate source images in transform domain. This type of approaches could remove the block effect and get more consistent fusion result. Image fusion methods based on MSD draw researcher's attention in recent years. For example, discrete wavelet transform based method[8, 9], stationary wavelet transform based method[10], double-tree complex wavelet transform based method[11], curvelet transform based method[7], contourlet transform based method[12], non-subsampled contourlet transform based method[2] etc.

QWT, as a popular MSD tool, is first proposed in 2004[13]. It provides a richer scale space analysis for image compared to other MSD tools because it can decompose image into magnitude and phase information. The magnitude of QWT is near shift invariant so it have better texture representation than wavelet and complex wavelet, and the phase of QWT contains richer geometric information. Based on the above merits, QWT has been used in image denoising [14, 15] and image classification [16]. Therefore, we use QWT as MSD tool in our work.

The performance of feature extraction directly affects the quality of fusion result in MSD based image fusion method. In traditional MSD based fusion methods, often using a single feature to describe the attribute of image, such as the contrast of image for fusion[9], the phase congruency and directional contrast for the fusion of low frequency subbands and high frequency subbands[17], the sparse and energy for the fusion of low frequency subbands and high frequency subbands[18]. However, a single feature is often a partial description of image. To overcome this drawback, multiple features of image are extracted and integrated in our work to get more accurate representation of image. Based on this idea, three features are extracted for the fusion of low frequency subbands. The first one is based on the phase information of low frequency coefficient, which can be used to measure its clarity. The second one is the magnitude information of low frequency coefficient, which presents its energy information. Especially in medical or infrared images, the magnitudes of target and bone area are outstanding. The third one is the variance of spatial region corresponding to the low frequency coefficient, which is used to represent the salience of coefficient. As for the high frequency subbands, two features are extracted. Considering the fact that the human eyes are more sensitive to the contrast than the intensity of image, we take the contrast of image as the first feature. In addition, the details of image often present large coefficient values. Therefore, the energy of high frequency coefficient is employed as the second feature. After feature extraction, we integrate multiple features of low frequency coefficient and high frequency coefficient into two comprehensive features, respectively. The fusion result can be obtained based on them.

The rest of the paper is organized as follows: The quaternion

P. Chai and X. Luo are with School of Internet of Things, Jiangnan University, Wuxi, Jiangsu, 214122, China (e-mail: chai_pengfei@163.com, xqluo@jiangnan.edu.cn).

Z. Zhang is with the College of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou, Jiangsu, 215009, China (e-mail: cimszhang@163.com)

and quaternion wavelet transform is introduced in Section 2; The framework of the proposed method is shown in Section 3; The fusion rules for low frequency subbands and high frequency subbands are presented in Section 4 and 5, respectively; The experimental results and analyses are given in Section 6; Finally, we conclude this paper in Section 7.

II. QUATERNION WAVELET TRANSFORM

A. Quaternion algebra

The quaternion algebra $\mathcal{H}$ is a generalization of the complex algebra[19], and the mathematical representation is:

$$\mathcal{H} = q = a + bi + ci + dk \quad a,b,c,d \in R \tag{1}$$

where the orthogonal imaginary numbers i, j and k satisfy the following rules:

$$i^2 = j^2 = k^2 = -1, \; ij = k, \; jk = i, \; ki = j \tag{2}$$

The quaternion could also represent as

$$Q = |q| e^{i\phi} e^{k\psi} e^{j\theta} \tag{3}$$

where $(\phi,\theta,\psi) \in [-\pi,\pi) \times [-\pi/2,\pi/2) \times [-\pi/4,\pi/4]$. $|q|$ is magnitude and ϕ, θ, ψ are phases. The computational formulas of them is written as

$$\begin{cases} \phi = \arctan\left(\dfrac{2(ac+bd)}{a^2+b^2-c^2-d^2}\right) \\ \theta = \arctan\left(\dfrac{2(ab+cd)}{a^2-b^2+c^2-d^2}\right) \\ \psi = -\dfrac{1}{2}\arcsin\left(2(ad-db)\right) \end{cases} \tag{4}$$

B. 2-D Quaternion wavelet construction

The QWT of image f(x, y) can be defined as:

$$f(x,y) = A_n^q f(x,y)$$
$$+ \sum_{s=1}^{n} \left[D_{s,1}^q f(x,y) + D_{s,2}^q f(x,y) + D_{s,3}^q f(x,y) \right] \tag{5}$$

where $A_s^q f(x,y)$ and $D_{s,p}^q f(x,y)$ $(p=1,2,3)$ are approximation and difference directional components respectively, which also be called as the low frequency subbands and high frequency subbands of image. The analytic extension is constructed by real wavelet and its 2-D Hilbert transform:
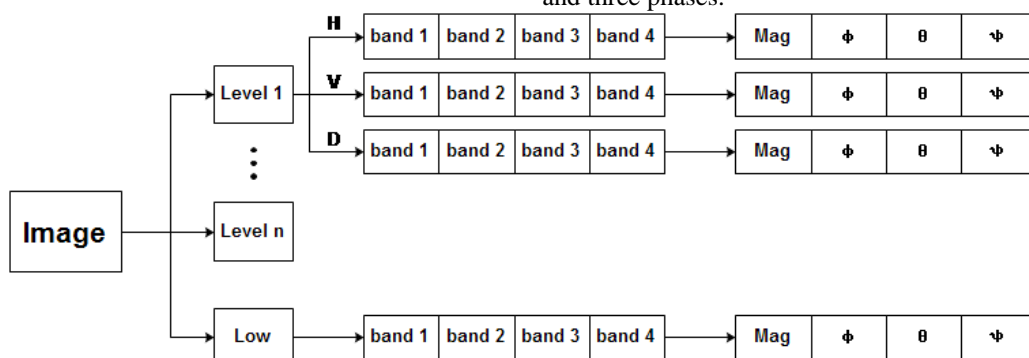
$$\begin{cases} \psi^D(x,y) = \psi_h(x)\psi_h(y) \\ \quad \rightarrow \psi^D + iH_{i1}\psi^D + jH_{i2}\psi^D + kH_i\psi^D \\ \psi^V(x,y) = \phi_h(x)\psi_h(y) \\ \quad \rightarrow \psi^V + iH_{i1}\psi^V + jH_{i2}\psi^V + kH_i\psi^V \\ \psi^H(x,y) = \psi_h(x)\phi_h(y) \\ \quad \rightarrow \psi^H + iH_{i1}\psi^H + jH_{i2}\psi^H + kH_i\psi^H \\ \phi(x,y) = \phi_h(x)\phi_h(y) \\ \quad \rightarrow \phi + iH_{i1}\phi + jH_{i2}\phi + kH_i\phi \end{cases} \tag{6}$$

where ϕ is scaling function, ψ^D, ψ^V, ψ^H is wavelet functions which are oriented at diagonal, vertical and horizontal respectively. 2-D Hilbert Transform can get by 1-D Hilbert Transform along x and y axis respectively:

$$\begin{cases} (\psi_h, \psi_g = H\psi_h) \\ (\phi_h, \phi_g = H\phi_h) \end{cases} \tag{7}$$

The 2-D QWT is defined as:

$$\begin{cases} \psi^D(x,y) = \psi_h(x)\psi_h(y) + i\psi_g(x)\psi_h(y) \\ \quad + j\psi_h(x)\psi_g(y) + k\psi_g(x)\psi_g(y) \\ \psi^V(x,y) = \phi_h(x)\psi_h(y) + i\phi_g(x)\psi_h(y) \\ \quad + j\phi_h(x)\psi_g(y) + k\phi_g(x)\psi_g(y) \\ \psi^H(x,y) = \psi_h(x)\phi_h(y) + i\psi_g(x)\phi_h(y) \\ \quad + j\psi_h(x)\phi_g(y) + k\psi_g(x)\phi_g(y) \\ \phi(x,y) = \phi_h(x)\phi_h(y) + i\phi_g(x)\phi_h(y) \\ \quad + j\phi_h(x)\phi_g(y) + k\phi_g(x)\phi_g(y) \end{cases} \tag{8}$$

where the first three rows are the computational formulas of QWT high coefficients of image in diagonal, vertical and horizontal directions, respectively. The last row presents the computational formula of QWT low coefficient of image. The 2-D QWT can be represented by magnitude and phases by means of substituting Eq.(8) into Eq.(4).

C. An example of Quaternion wavelet transform for image

After image is decomposed by QWT, we can obtain a low frequency part and n groups of high frequency parts. The QWT decomposition structure of image is shown in Fig. 1. "Low" represents the low frequency part which composed by four low frequency subbands, that is, band 1 to band 4. At each level, the high frequency information is presented in 3 directions (horizontal (H), vertical (V) and diagonal (D)), and there are four subbands (band 1, band 2, band 3, band 4) in each direction. These four subbands can be transformed into one magnitude and three phases.

Fig. 1. The QWT decomposition structure of image.

To intuitively illustrate the QWT decomposition, the QWT is conducted on MRI medical image (Fig.2) and the decomposition results are shown in Fig.3 and Fig.4. Fig. 3 and Fig. 4 are the low frequency part and high frequency part of Fig.2 obtained by using one level quaternion wavelet decomposition. In the first row of Fig.3, Fig. 3(a) to 3(d) are four low frequency subbands. In the second row of Fig.3, the low frequency part of image is transformed into magnitude and phases. Fig. 3(e) is the magnitude which provides overview information of source image and Fig. 3(f) to 3(h) are three phases. The edge and contour information of image can be seen from these three phases. Considering the spatial limitation, Fig.

4 shows the high frequency part of Fig.2 in the vertical direction. Similar to Fig.3, the first row is four high frequency subbands, and the second row is the magnitude and phase representation of high frequency part. The QWT magnitude (Fig. 4(e)) represents structural. The two phases(ϕ and θ) (Fig.4(f) and Fig.4(g)) represent the image local shift information and the other phase ψ (Fig.4(h)) represents the texture features of the image. Through this example, we can observe that QWT is an effective MSD tool because it can reveal more internal information of image.
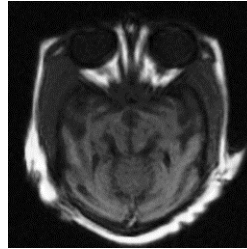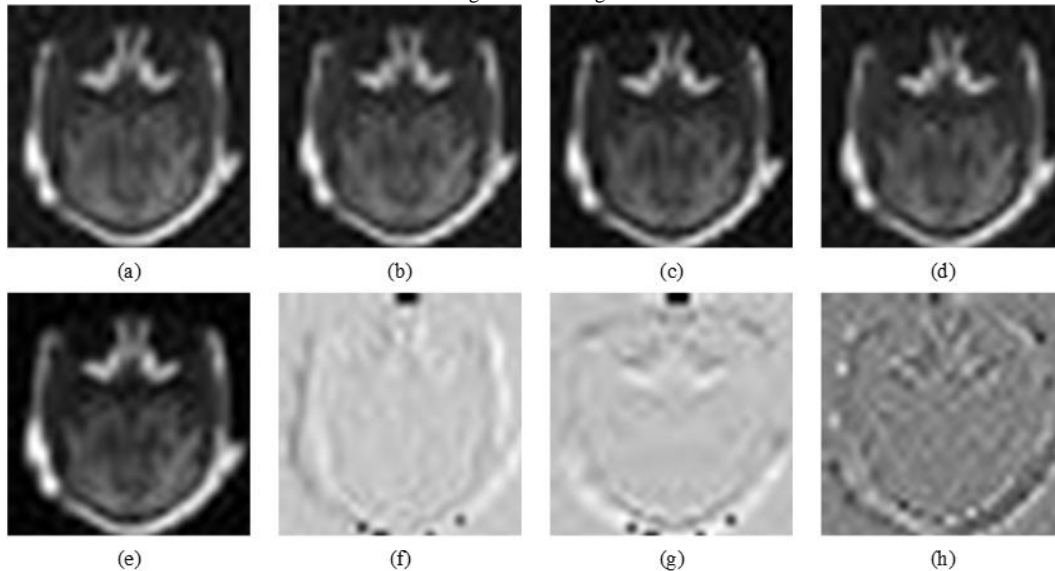

Fig. 2. MRI image.

Fig. 3. The low frequency part of Fig.2 at level=1; (a) to (d) are four low frequency subbands of QWT; (e) to (h) are the low frequency subbands express in terms of magnitude and phases,;(e) the magnitude of low frequency part; (f) the phase ϕ of low frequency part; (g) the phase θ of low frequency part; (h) the phase ψ of low frequency part.
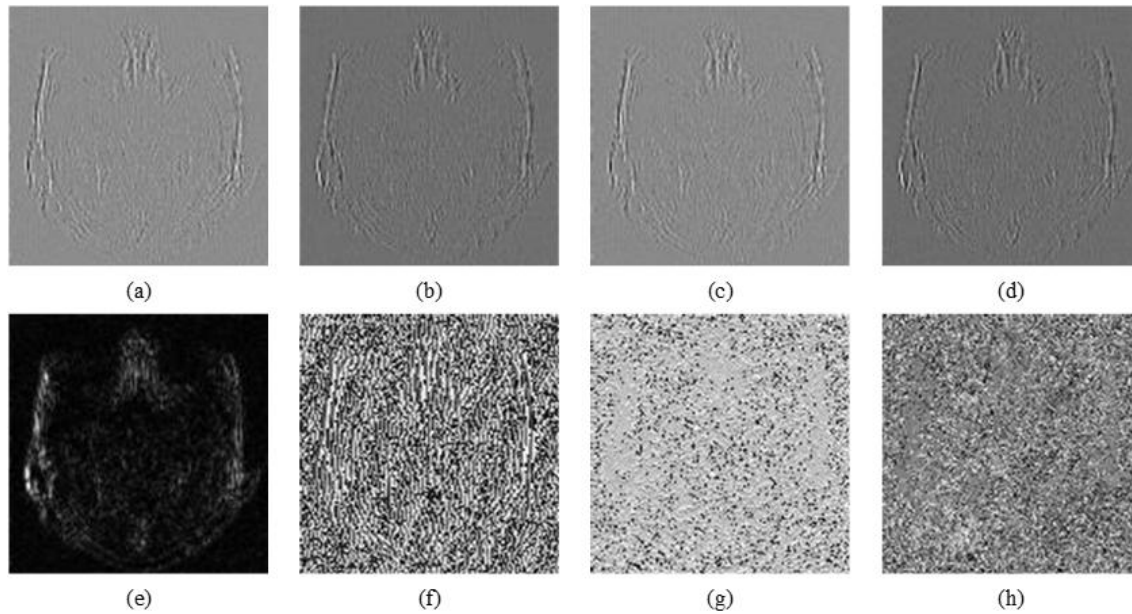
Fig. 4. The high frequency part of Fig.2 at the first level in the vertical direction; (a) to (d) are four high frequency subbands of QWT ; (e) to (h) are the high frequency subbands express in terms of magnitude and phases; (e) the magnitude of high frequency part in the vertical direction; (f) the phase ϕ of high frequency part in the vertical direction; (g) the phase θ of high frequency part in the vertical direction; (h) the phase ψ of high frequency part in the vertical direction.

III. THE FRAMEWORK OF THE PROPOSED METHOD

In this paper, a novel multimodal image fusion using quaternion wavelet transform and multiple features is proposed. The framework of the proposed method is given in Fig.5. As can be seen from Fig.5, the fusion processing can be divided into three parts. In the first part, the source image A and B are decomposed by QWT into low frequency part and high frequency part. In the second part, the weighted average fusion rule and the choose-max fusion rule are proposed to fuse the low and high frequency coefficients respectively. Finally, the fused image is obtained by using inverse QWT. The detail fusion procedure is described as follows.

1) Assuming that the source images have been registered. QWT is performed on the two source images A and B to obtain low frequency part L_A^n, L_B^n and high frequency part $H_A^{l,d,n}$, $H_B^{l,d,n}$. L_*^n is the n-th low frequency subband, $H_*^{l,d,n}$ represent the n-th high frequency subband at level l in d direction.

2) The weighted average fusion rule based on multi-feature is proposed for low frequency parts, in which the multi-feature is proposed by the combination of the phase information of coefficient, the magnitude information of

coefficient and the regional variance of image.

$$L_F^n(x,y) = L_A^n(x,y) \times \omega(x,y) + L_B^n(x,y) \times (1 - \omega(x,y)) \quad (9)$$

where (x,y) is the location of coefficient and $\omega(x,y)$ is the weight of coefficient. The calculation of weight $\omega(x,y)$ is based on multi-feature of low frequency coefficient, which will be discussed in Section 4.

3) The choose-max fusion rule based on multi-feature is designed to get the fused high frequency parts, in which the multi-feature $M_{*,l,d,n}^H(x,y)$ is obtained by combining the energy and local contrast of coefficient:

$$H_F^{l,d,n}(x,y) = \begin{cases} H_A^{l,d,n}(x,y), & if\ M_{A,l,d,n}^H > M_{B,l,d,n}^H(x,y) \\ H_B^{l,d,n}(x,y), & if\ M_{A,l,d,n}^H \leqslant M_{B,l,d,n}^H(x,y) \end{cases} \quad (10)$$

where $M_{*,l,d,n}^H(x,y)$ are defined as the activity level measure of high frequency coefficient, which will be described in Section 5.

4) The final fused image F can be obtained by inverse QWT over the fused low frequency coefficients L_F^n and the fused high frequency coefficients $H_F^{l,d,n}$.
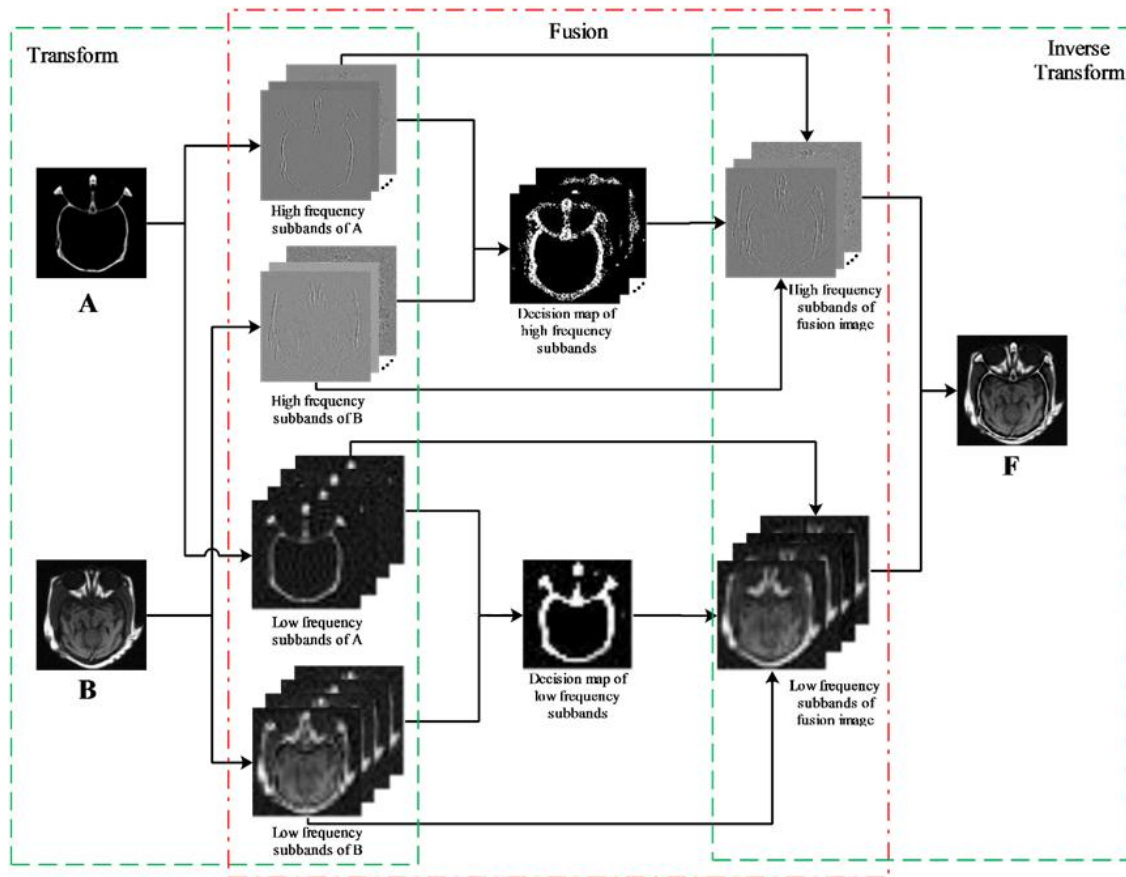
Fig. 5. The framework of the proposed method.

IV. THE FUSION RULE FOR LOW FREQUENCY SUBBANDS

There is most of energy in the low frequency part of the source image. The fused low frequency part is the approximation of fused image. Therefore, the fused low frequency information affect the whole perception of fused image. In this paper, a weighted average fusion rule based on multi-feature is proposed for low frequency subbands to get better fused low frequency part. The fusion weight is calculated by a comprehensive feature of low frequency coefficient consisting of the phase and magnitude of coefficient and the spatial variance of coefficient. The detail description of the comprehensive feature is given below. The QWT low-frequency part can be transformed into one magnitude matrix and three phase matrices. The first two phases ϕ and θ represent local image shifts, and the last one denotes the texture information of image. As can be seen from Fig. 3, phase ϕ and θ contain rich useful information. Therefore, the first salient feature of coefficient can be defined based on phase ϕ and θ, which reflects the clarity of low frequency coefficient[17, 20].

$$P(x,y) = \sqrt{\sigma_1^2(x,y) + \sigma_2^2(x,y)} \qquad (11)$$

where, $\sigma_1^2(x,y), \sigma_2^2(x,y)$ is the variance of the phase ϕ, θ in a window centered on (x, y). Here, the windows size is 5×5.

Besides the phases, the magnitude of coefficient is an effective feature. Especially in medical image and infrared image, the useful information such as the bone information in CT image and the infrared object present more energy and less detail. Therefore, we should measure the salience of them from the energy point of view. Fortunately, the magnitude of coefficient $mag(x,y)$ is a good description of energy information, so it is employed as the second feature of coefficient and is named as G.

Considering the fact that the low frequency subband is the approximation of source image, there are strong correlative between the source image and its low frequency subband. Moreover, the source image is downsampled in the decomposition processing. Therefore, we take the variance of region corresponding to the low frequency coefficient in the source image as the third feature of low frequency coefficient. Specifically, this feature is

$$S(x,y) = var\{I(i,j) \mid (i,j) \in \Omega(x,y)\} \qquad (12)$$

where $I(i,j)$ is the intensity value of pixel, $var(*)$ represents the variance operation and $\Omega(x,y)$ is the corresponding spatial window of the low frequency coefficient $L_*^n(x,y)$. For example, if the original image size is 512×512 and the decomposition level of QWT is 3, then the size of $\Omega(x,y)$ is 8×8. That is to say, when the decomposition level is defined as n, the size of $\Omega(x,y)$ is $2^n \times 2^n$ and the starting position is $(x \cdot 2^n, y \cdot 2^n)$.

Next, we integrate these three features into a comprehensive feature employed as the activity level measure to accurately

measure the salience of low frequency coefficient.

$$M^L(x,y) = |P(x,y)|^{\alpha_1} \cdot |G(x,y)|^{\alpha_2} \cdot |S(x,y)|^{\alpha_3} \quad (13)$$

where α_1, α_2, α_3 are the weigh factor. The contributions of features to M can be adjusted by them.

Finally, the weight of low frequency fusion rule is calculated by formula (14). The fused low-frequency subbands can be obtained by using formula (9):

$$w(x,y) = \frac{M_A^L(x,y)}{M_A^L(x,y) + M_B^L(x,y)} \quad (14)$$

To verify the effectiveness of M^L compared to individual feature P, G and S, experiments are conducted on multi-focus images (Fig.6(a1) and Fig.6(b1)), medical images (Fig.6(a2) and Fig.6(b2)), infrared-visible images (Fig.6(a3) and Fig.6(b3)) and remote sensing images (Fig.6(a4) and Fig.6(b4)). Fig.6(c(*)), Fig.6(d(*)) and Fig.6(e(*)) are the weight maps w obtained by separately using three features

such as P, G and S to replace M^L in the formula (14). Fig. 6(f(*)) are the weight map calculating by using the comprehensive feature M^L. In the weight map, the intensity correspond to the size of weight. The white points mean that the weight is equal to 1, that is to say, the fused low frequency coefficient is fully from low frequency coefficient of image A, and vice versa. As can be seen from these weight maps, the performances of features P, G, S rely on the type of image. Feature P is good at remote sensing images, feature G have a better performance in multi-focus and medical images, feature S is better in infrared-visible images. However, the performance of feature M^L is efficient and robust in all kinds of images because M^L integrates the advantages of three feature P, G and S.
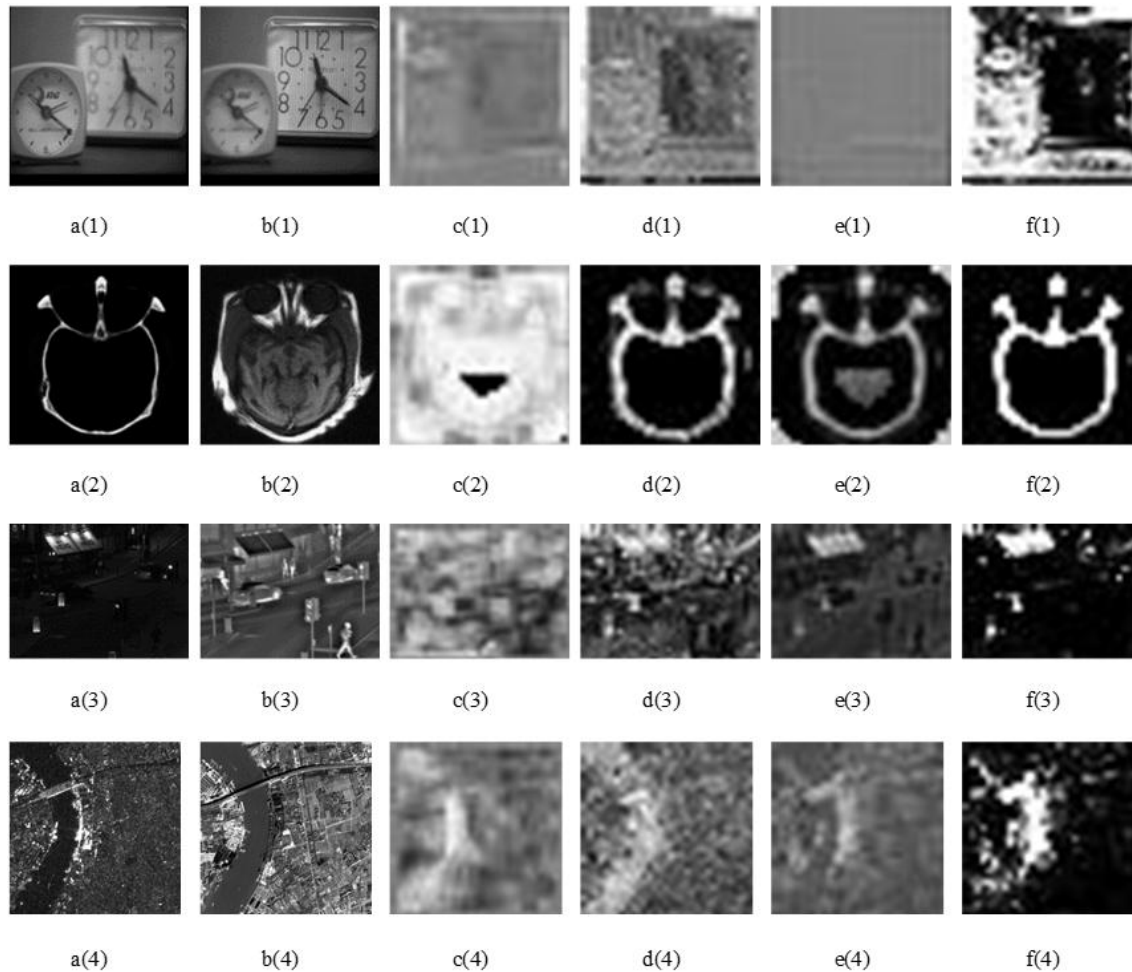

Fig. 6. The source images (a(*) and b(*)) and the weight maps obtained by using P (c(*)), G(d(*)), S(e(*)) and M^L (f(*)). (*=1, 2, 3, 4), α_1, α_2, α_3 are defined as 5, 5, 0.1; 2, 1, 4; 1, 1.4, 2 and 1, 1, 5 in four experiments, respectively).

V. THE FUSION RULE FOR HIGH FREQUENCY SUBBANDS

The high frequency subbands represent the details of image such as edge, line, and corner. In general, the detail of image presents a large absolute value of the high frequency coefficient. However, sometimes noise has the same phenomenon. Fortunately, the local contrast of image can address this problem[5]. The contrast is represented as follows.

$$D(i,j) = \begin{cases} \left(\dfrac{1}{L(i,j)}\right)^{\alpha} \dfrac{H(i,j)}{L(i,j)}, & if\ L(i,j) \neq 0 \\ H(i,j), & if\ L(i,j) \neq 0 \end{cases} \quad (15)$$

where α is a visual constant and is determined by a physiological visual test in a range of 0.6 to 0.7, L(i,j) and H(i,j) are low frequency coefficient and high frequency coefficient, respectively. The relationship between contrast and background intensity is non-linear, which makes the human visual system highly sensitive to contrast variations.

It is worth noting that the contrast of high frequency coefficient can effective distinguish noise and detail, however, it weaken the energy information of that. Therefore, to accurately represent the salience of high frequency coefficient, a combination of contrast and energy of coefficient is proposed as the comprehensive feature of coefficient $M^{H}_{*,l,d,n}$ in this paper.

$$M^{H}_{*,l,d,n}(i,j) = |D_{*,l,d,n}(i,j)|^{\beta_1} \cdot |E_{*,l,d,n}(i,j)|^{\beta_2} \quad (16)$$

where, $E_{*,l,d,n}(i,j) = |H^{l,d,n}_{*}(i,j)|^2$ denotes the energy of high-frequency coefficient. By changing the value of β_1, β_2 to adjust the proportion of these two indicators in $M^{H}_{*,l,d,n}$. Then, $M^{H}_{*,l,d,n}(i,j)$ is defined as the activity level measure of

high frequency subband. The fused high frequency subbands can be obtained by formula (10).

A series of experimental results are shown in Fig.7 to test the effectiveness of $M^{H}_{*,l,d,n}$. Fig.7(a) and Fig.7(b) are the source images. Fig.7(c(*)) , Fig.7(d(*)) and Fig.7(e(*)) are the selection maps obtained by $D_{*,l,d,n}$, $E_{*,l,d,n}$ and $M^{H}_{*,l,d,n}$. if the activity level measure of $H^{l,n,d}_{A}(i,j)$ is larger than that of $H^{l,n,d}_{B}(i,j)$, then the pixel of selection map is white. Otherwise, the pixel of selection map is black. As can be seen from Fig. 7, the selection map using $D_{*,l,d,n}$ feature is rough, which means that the $D_{*,l,d,n}$ feature can represent the large-scale characteristics, and as for the $E_{*,l,d,n}$ feature, thought the selection map is fine, there are noises in it. It is obvious that the multi-feature metric $M^{H}_{*,l,d,n}$ obtained by combining the feature $D_{*,l,d,n}$ and feature $E_{*,l,d,n}$ get more perfect selection results.
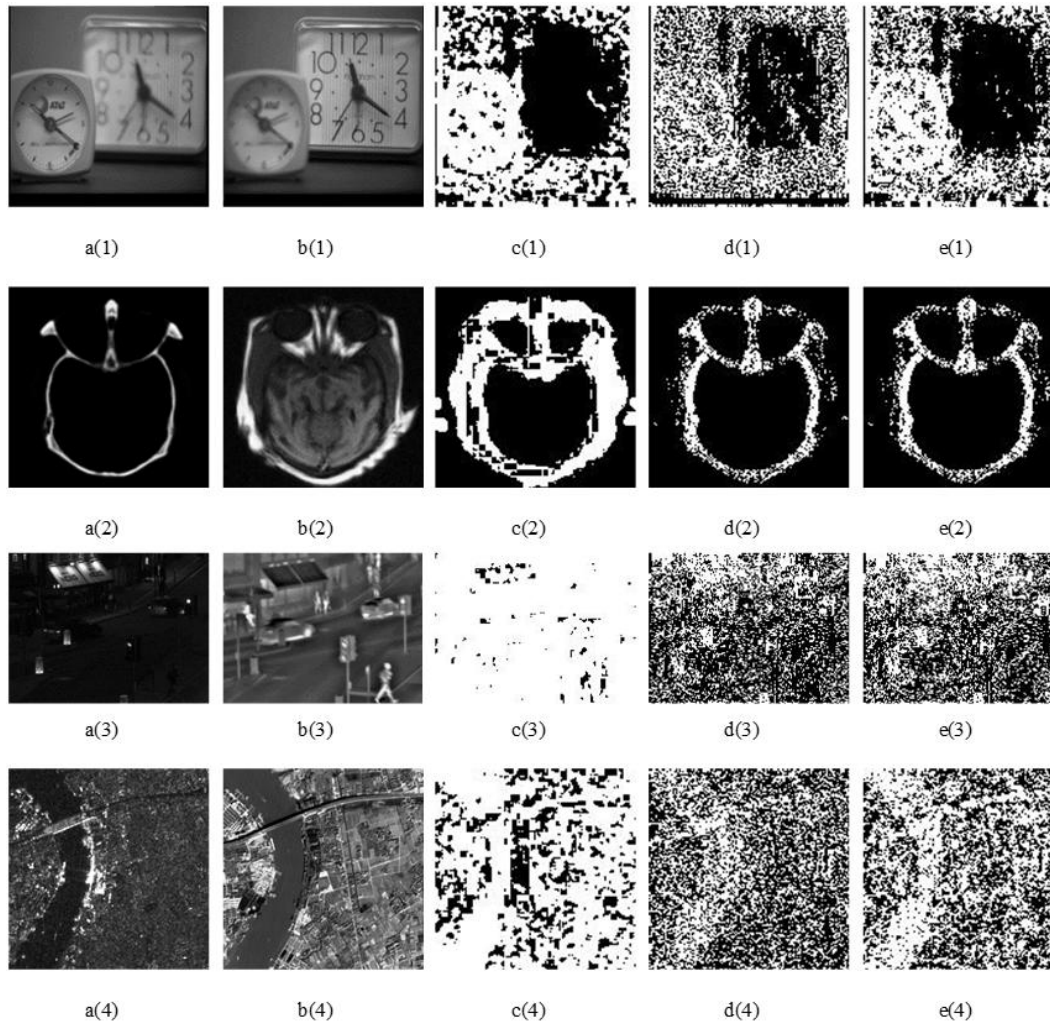

Fig. 7. The source images (a(*) and b(*)) and the selection maps obtained by using $D_{*,l,d,n}$ (c(*)), $E_{*,l,d,n}$ (d(*)) and $M^{H}_{*,l,d,n}$ (e(*)). (*=1,2,3,4), β_1 and β_2 are defined as 2, 1; 0.1, 2; 0.1, 1 and 1, 1 in four experiments, respectively, and l=1, d=1, n=1.

VI. EXPERIMENTAL RESULTS AND ANALYSES

To verify the effectiveness of the proposed method in various types of source images, the experiments are conducted on 4 pairs of source images. They are multi-focus images, medical images, infrared-visible images and remote images downloaded from the homepage of Yu Liu[18]. Several comparison fusion methods are employed including Laplacian pyramid based method (LP)[18], dual-tree complex wavelet transform based method (DTCWT)[18], curvelet transform based method (CVT)[18], nonsubsampled contourlet transform based method (NSCT)[18], directive contrast based method (DC)[5], Guided filter based method (GFF)[21] and Pulse Coupled Neural Networks based method (PCNN)[22]. The source images are decomposed into three levels in the above mentioned MSD based fusion methods. All parameters of comparison experiments are set according to the reference paper. All experiments are conducted with Matlab 2014b and run on a PC with a Pentium 3.5 GHz CPU and an 8GB RAM.

A. Subjective evaluations of experimental results

Fig. 8 is a set of multi-focus image fusion experimental results. Fig. 8(a) and 8(b) are the left focus image and the right focus image respectively. Fig. 8(c)-(j) are the fused images obtained by using various methods. Because these image fusion methods are implemented in transform domain, the block artifacts are overcome. To facilitate observation, the blue areas in fused results are zoomed. It can be seen that the amplified areas of Fig. 8(d), (f), (h) and (i) are unclear. There are some noises on both sides of "1" in Fig. 8(g). Subjectively, the rest fusion results get a good performance.

Next, medical image fusion experimental results are presented in Fig. 9. Fig. 9(a) is the CT image, which usually contains dense structures such as bone information. Fig. 9(b) is

the magnetic resonance imaging (MRI) image, which contains the details of soft tissue information. As we can see, Fig. 9(c), 9(e), 9(g), and 9(h) lose some CT image information, Fig. 9(d), (e), (f) and (i) lose some MRI image information. The MRI information is blurry in Fig. 9(g). In the zoomed region, it can be found that the amplified area of Fig. 9(g) loses the information of CT image, and the amplified areas of Fig. 9(c)-(f), (h) get low contrast. Though the global and local observation, it is easy to see that only Fig. 9(h) obtained by the proposed method contains both CT and MRI image information completely.

The third group experiment is the infrared-visible image fusion. Fig. 10(a) is a visible image containing the lights and billboard, Fig. 10(b) is an infrared image presenting the cars and pedestrians. It is easy to see that the fused images (Fig. 10(f), 10(g) and 10(i)) are coarse. The billboards of Fig. 10(e) are blurry.

Fig. 11 shows the fusion results of the remote sensing images Fig. 11(a) and Fig. 11(b). There are a number of white spots in Fig. 11(g). The boundaries of bridge are unclear in Fig. 11(c), (d), (f) and (h). Moreover, it can be found that in the right border of river of Fig. 11(i) and the land of right river of Fig. 11(e) also lose information. Generally, Fig. 11(j) is the best fusion result in subjective.

Based on the above four groups of experiments, it can be concluded that comparison fusion methods perform well in multi-focus image fusion while they get poor fusion quality in medical image fusion and infrared-visible image fusion because the comparison methods only use a signal feature to present the salience of source images which reduce the robustness of the fusion method. Owing to the usage of comprehensive features, the proposed method can well preserve the information of source images without producing artifacts and distortions in all types of images.
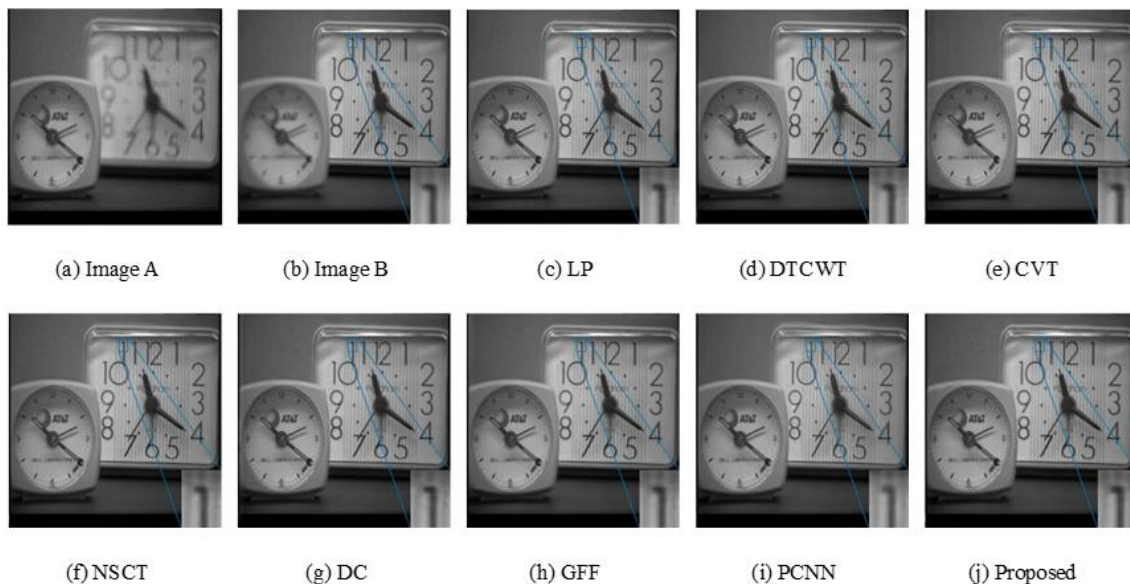

(a) Image A　　(b) Image B　　(c) LP　　(d) DTCWT　　(e) CVT

(f) NSCT　　(g) DC　　(h) GFF　　(i) PCNN　　(j) Proposed

Fig. 8. Multi-focus image fusion results.

(a) CT   (b) MRI   (c) LP   (d) DTCWT   (e) CVT

(f) NSCT   (g) DC   (h) GFF   (i) PCNN   (j) Proposed
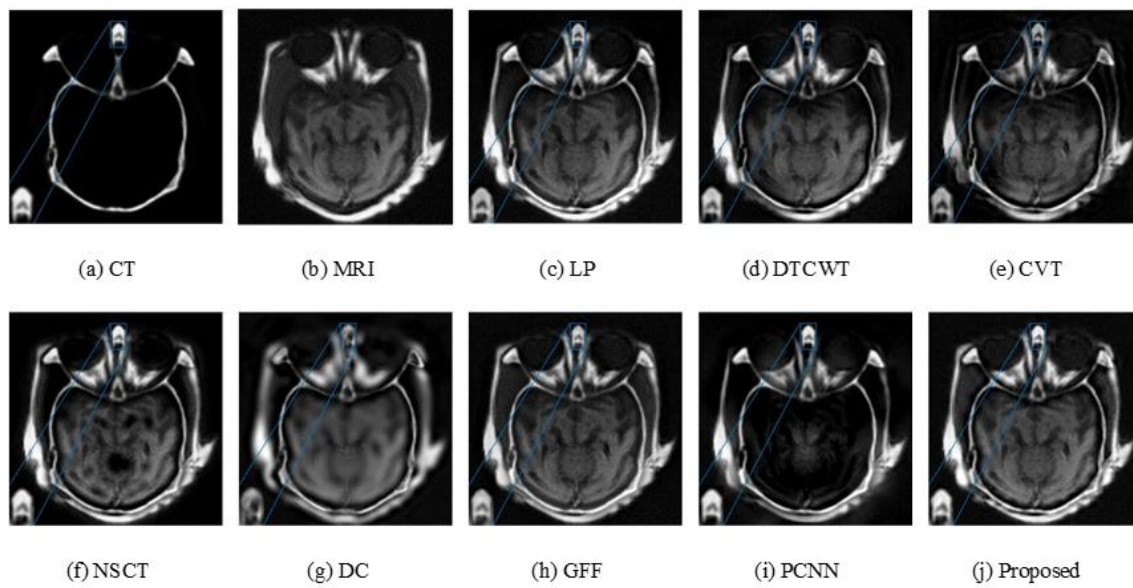
Fig. 9. Medical image fusion results.

(a) Image A   (b) Image B   (c) LP   (d) DTCWT   (e) CVT

(f) NSCT   (g) DC   (h) GFF   (i) PCNN   (j) Proposed
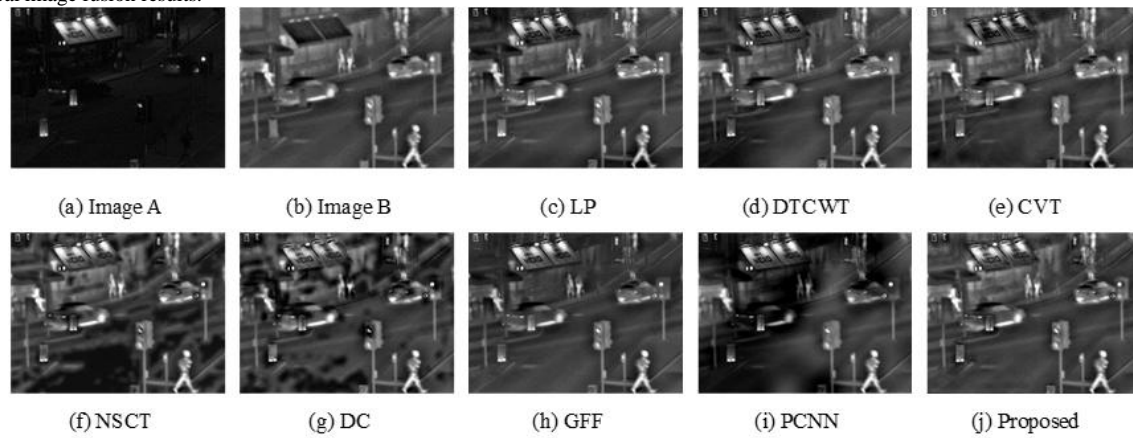
Fig. 10. Infrared-visible image fusion results.

(a) Image A   (b) Image B   (c) LP   (d) DTCWT   (e) CVT

(f) NSCT   (g) DC   (h) GFF   (i) PCNN   (j) Proposed
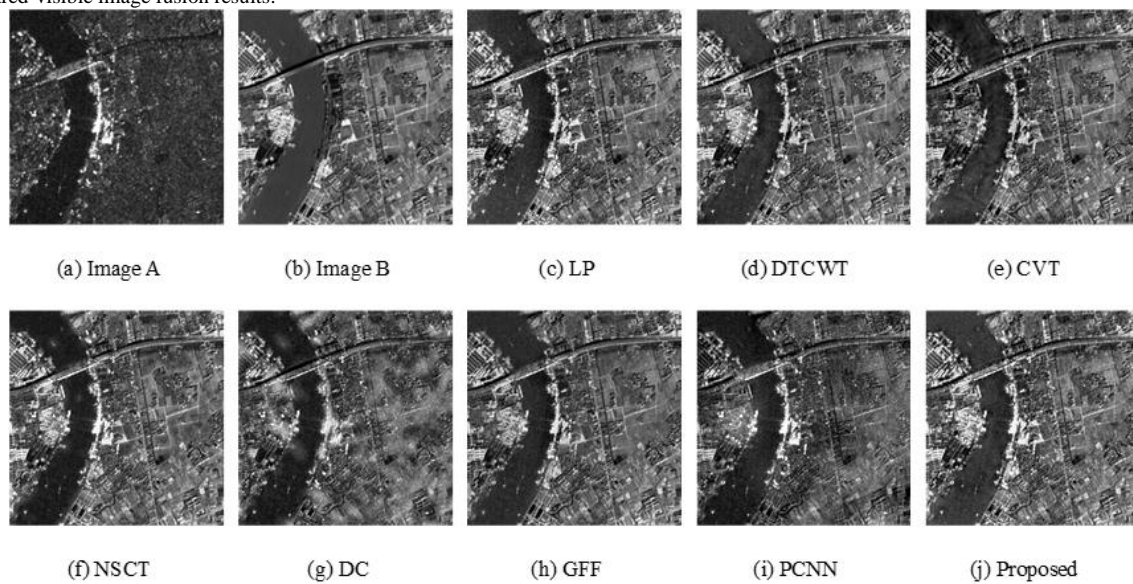
Fig. 11. Remote sensing image fusion results.

B. Objective evaluations of experimental results

To provide further quantitative comparison of various fusion methods, objective evaluations of fused results are given in this section. The average gradient (AG), edge intensity (EI), mutual information (MI), spatial frequency (SF) and $Q^{AB/F}$[23] are adopted as evaluation indexes. Large values indicate a better fusion result for all indexes.

The objective evaluation indexes of fusion results are shown in Table 1, in which the best results are labeled in bold. We did not consider the evaluation metrics of the NSCT method and DC method in infrared-visible fusion experiments because of bad subjective perception. Except for the infrared-visible fusion result obtained by NSCT based and DC based method, the proposed method gets the best values in terms of AG, EI and SF. We use the superscript (2) to mark them. It can be seen that the AG and EI indexes of the proposed method are the best in all experiments. This means that the proposed method can well preserve the useful information of source images. The SF shows that the proposed method has a good performance in all types of image fusion except for multi-focus image fusion. The MI and $Q^{AB/F}$ evaluate the amount of information and edge information from source image to fused image, respectively.

Though Fig. 9(h) get the best value in terms of $Q^{AB/F}$ index, it presents a bad work in the border, especially in zoomed region. In addition, the objective evaluation of Fig.11(c) is better than that of Fig. 11(j) in terms of MI index, but the border of bridge is blurry in Fig.11(c). The above analyses demonstrate that all metrics should be collectively considered in the objective evaluations of fused images and the performance evaluations of fused methods is determined by the subjective and objective evaluations together. In general, a conclusion can be drawn that the propose method can get better fusion performance in various types of image fusion whatever the subjective or objective evaluation because QWT can better represent the source image and the proposed comprehensive features are effective salience metrics.

TABLE I
THE OBJECTIVE EVALUATION INDEXES OF FUSION RESULTS

Image	Indexes	LP	DTCWT	CVT	NSCT	DC	GFF	PCNN	Proposed
Multi-focus	AG	6.7302	6.6571	6.6859	6.6634	6.6727	6.5349	6.2978	**6.7654**
	EI	60.6542	60.5254	60.6250	60.3300	60.4671	59.5325	57.0380	**60.7847**
	MI	7.4504	7.7523	7.4564	7.5377	7.4997	**8.0518**	7.4383	7.3283
	SF	**16.3544**	16.2910	15.9972	16.2130	16.2535	16.0684	15.2428	16.1881
	$Q^{AB/F}$	0.7167	0.7164	0.7030	0.7158	0.7139	**0.7321**	0.6843	0.6959
Medical	AG	6.9644	6.6152	6.2812	7.1252	5.5517	6.5808	5.7960	**7.4087**
	EI	70.1917	66.3107	62.9066	72.9038	58.3791	66.4941	59.0610	**74.9809**
	MI	2.7430	2.3615	1.6664	2.4777	1.9229	3.4307	1.2819	**3.7158**
	SF	17.4651	16.3641	15.6952	17.5549	13.6848	16.0034	16.1484	**17.9141**
	$Q^{AB/F}$	0.7604	0.6593	0.5893	0.7046	0.4148	**0.7850**	0.5889	0.7338
Infrared-visible	AG	6.3119	6.1476	6.2225	**6.7664**	5.8896	5.6140	5.4547	6.5739[2]
	EI	57.9215	56.7403	57.7813	**64.5372**	57.0736	51.8637	51.2432	60.656[2]
	MI	2.8103	**3.3657**	1.9596	2.2367	1.3822	1.8676	1.2690	3.2112
	SF	16.2076	15.9539	15.9245	**16.5090**	14.8261	15.2460	15.4707	16.419[2]
	$Q^{AB/F}$	**0.6751**	0.6436	0.5691	0.5856	0.3227	0.6740	0.5212	0.6307
Remote sensing	AG	25.1626	23.4388	23.9729	24.4697	20.8794	21.7396	20.6969	**25.8159**
	EI	175.9880	168.5175	171.4893	175.4893	145.2283	157.0075	147.5072	**182.9052**
	MI	**3.2415**	2.9388	2.3715	2.7321	2.0729	3.0174	1.9628	2.5721
	SF	47.7289	44.9158	45.0617	46.3640	39.7994	42.1385	39.6685	**48.3638**
	$Q^{AB/F}$	0.5882	0.5366	0.4758	0.5231	0.3774	**0.6008**	0.3930	0.4901

VII. CONCLUSION

In this paper, a novel image fusion method Using QWT and multiple features is proposed. Compared to traditional MSD tools, the QWT can provide abundant magnitude and phase information, which meet approximate translation invariance and limited redundancy. Different from the traditional fusion methods using a single feature as the activity level measure, we combine the magnitude, phase and spatial variance of low frequency coefficient into a comprehensive feature as the activity level measure of low frequency coefficient and combine the contrast and energy of high frequency coefficient

into the other comprehensive feature as the activity level measure of high frequency coefficient. These two multi-features are reliable and robust, which are available for image fusion. Finally, the experimental results demonstrate the proposed method is effective in all kinds of image fusion.

REFERENCES

[1] H. Li, L. Li, and J. Zhang, "Multi-focus image fusion based on sparse feature matrix decomposition and morphological filtering," *Optics Communications,* vol. 342, pp. 1-11, 2015.
[2] Q. Zhang and B.-l. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Processing,* vol. 89, pp. 1334-1346, 2009.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2685178, IEEE Access

Access-2017-01106   11

[3]   J. Tian and L. Chen, "Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure," *Signal Processing,* vol. 92, pp. 2137-2146, 2012.

[4]   L. Yang, B. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing,* vol. 72, pp. 203-211, 2008.

[5]   G. Bhatnagar, Q. J. Wu, and Z. Liu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE transactions on multimedia,* vol. 15, pp. 1014-1024, 2013.

[6]   S. Singh, A. Gyaourova, G. Bebis, and I. Pavlidis, "Infrared and visible image fusion for face recognition," in *Proceedings of SPIE*, 2004, pp. 585-596.

[7]   F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, "Remote sensing image fusion using the curvelet transform," *Information Fusion,* vol. 8, pp. 143-156, 2007.

[8]   H. Li, B. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical models and image processing,* vol. 57, pp. 235-245, 1995.

[9]   T. Pu and G. Ni, "Contrast-based image fusion using the discrete wavelet transform," *Optical Engineering,* vol. 39, pp. 2075-2082, 2000.

[10] M. Beaulieu, S. Foucher, and L. Gagnon, "Multi-spectral image resolution refinement using stationary wavelet transform," in *International Geoscience and Remote Sensing Symposium*, 2003, pp. VI: 4032-4034.

[11] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel-and region-based image fusion with complex wavelets," *Information fusion,* vol. 8, pp. 119-130, 2007.

[12] M. Qiguang and W. Baoshu, "A novel image fusion method using contourlet transform," in *2006 International Conference on Communications, Circuits and Systems*, 2006, pp. 548-552.

[13] W. L. Chan, H. Choi, and R. Baraniuk, "Quaternion wavelets for image analysis and processing," in *Image Processing, 2004. ICIP'04. 2004 International Conference on*, 2004, pp. 3057-3060.

[14] M. Yin, W. Liu, J. Shui, and J. Wu, "Quaternion Wavelet Analysis and Application in Image Denoising," *Mathematical Problems in Engineering,* vol. 2012, pp. 587-612, 2012.

[15] S. Gai, L. Wang, G. Yang, and P. Yang, "Sparse representation based on vector extension of reduced quaternion matrix for multiscale image denoising," *IET Image Processing,* vol. 10, pp. 598-607, 2016.

[16] S. Gai, G. Yang, and S. Zhang, "Multiscale texture classification using reduced quaternion wavelet transform," *AEU - International Journal of Electronics and Communications,* vol. 67, pp. 233-241, 2013.

[17] Y. Liu, J. Jin, Q. Wang, Y. Shen, and X. Dong, "Region level based multi-focus image fusion using quaternion wavelet and normalized cut," *Signal Processing,* vol. 97, pp. 9-30, 2014.

[18] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion,* vol. 24, pp. 147-164, 2015.

[19] E. Bayro-Corrochano, "The Theory and Use of the Quaternion Wavelet Transform," *Journal of Mathematical Imaging and Vision,* vol. 24, pp. 19-35, 2006.

[20] Y. Liu, J. Jin, Q. Wang, and Y. Shen, "Phases measure of image sharpness based on quaternion wavelet," *Pattern Recognition Letters,* vol. 34, pp. 1063-1070, 2013.

[21] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing,* vol. 22, pp. 2864-2875, 2013.

[22] T. B. Ulow, "Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images," 1999.

[23] C. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics letters,* vol. 36, pp. 308-309, 2000.

First A. Pengfei Chai was born in Linfen, Shan Xi City, in 1991. He received the Bachelor's degree in Communication Engineering from Jiangnan University, Wuxi, China 2010. His current research interests include image fusion, pattern recognition and machine learning.

Second B. Xiaoqing Luo was born on 1980, received the BS degree in computer science and technology from Jingdezhen Ceramic Institute, Jiangxi, PR China, in 2001 and MS degree in computer science and technology in 2007, and PhD degree in Pattern Recognition and Intelligent System in 2010, both from Jiangnan University, Wuxi, PR China , respectively. She is currently taught in the School of Internet of things, an associate professor of Jiangnan University. Her current research interests are image fusion, pattern recognition, and other problems in image technologies.

Third C. Zhancheng Zhang was born in Pingyao, Shan Xi City, in 1977. He received the B.S. degree in material science from Jingdezhen Ceramic Institute, Jingdezhen, China, in 2000, the M.S. degree in computer science from Huazhong University of Science and Technology, Wuhan, China, in 2006, and the Ph.D. degrees from the School of Information Technology, Jiangnan University, in 2011. From 2012 to 2013, he was a post-doctor with the Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences. Since 2014, he has been an Assistant Professor with the College of Electronic and Information Engineering, Suzhou University of Science and Technology. He is the author of 15 articles, and holds five patents. His research interests include pattern recognition, AI in biomedical engineering and Health Informatics.