

Multi-modality image fusion via generalized Riesz-wavelet transformation

Bo Jin¹, Zhongliang Jing¹ and Han Pan¹

¹ School of Aeronautics and Astronautics, Shanghai Jiao Tong University
Shanghai, 200240, China

[e-mail: d.kampok@gmail.com, zljing@sjtu.edu.cn, hanpan@sjtu.edu.cn]

*Corresponding author: Bo Jin and Zhongliang Jing

Received July 7, 2014; revised September 4, 2014; accepted October 1, 2014; published November 30, 2014

Abstract

To preserve the spatial consistency of low-level features, generalized Riesz-wavelet transform (GRWT) is adopted for fusing multi-modality images. The proposed method can capture the directional image structure arbitrarily by exploiting a suitable parameterization fusion model and additional structural information. Its fusion patterns are controlled by a heuristic fusion model based on image phase and coherence features. It can explore and keep the structural information efficiently and consistently. A performance analysis of the proposed method applied to real-world images demonstrates that it is competitive with the state-of-art fusion methods, especially in combining structural information.

Keywords: Image processing, image fusion, Riesz-wavelet transform, heuristic fusion model

1. Introduction

Image fusion [1] is a unified process to provide a global physical scene from multi-source images. It commonly contains spatial or temporal representation. This technology has been applied to different application domains, such as medical imaging, night vision and environment monitoring [1]. Image fusion usually follows into three categories: pixel-level, feature-level and decision-level. It can take advantage of different sensors' basic features and improve visual perception for various applications. In this paper, we focus on developing a multi-modality image fusion algorithm, including visual, infrared, computed tomography (CT) and magnetic resonance imaging (MRI).

There are various image fusion methods for pixel-level fusion. First, wavelet-like image fusion methods or multi-resolution analysis [1] (MRA) based transformations, were proposed. Discrete wavelet frame [2] (DWF) transform is presented for fusing different directional information. A low redundancy extension of DWF, named by low redundant discrete wavelet frame transform (LRDWF), was proposed by Bo Yang *et. al.* [3]. Meanwhile, non-separable wavelet frame [4] based fusion method was also proven to be an effective method for the application of remote sensing. Moreover, lifting stationary wavelet [5] combined with pulse coupled neural network (PCNN) was applied to multi-focus image fusion. Several composite image fusion methods [6], based on non-subsampled contourlet transform [22] (NSCT) and shearlet [7], were proposed for different fusion problems. However, it should be noted that NSCT based fusion methods have low space and time efficiency. These methods were denoted as multi-resolution geometric analysis (MGA) based fusion methods. Second, sparse representation (SR) based image fusion methods were developed for handling related fusion problems, such as remote sensing image fusion [8]. Third, compressed sensing (CS) was popular for image fusion [9]. In summary, MGA based image fusion methods commonly lead to contrast reduction or high memory requirement. The SR based fusion methods exist some problems, for example time-space complexity and the blocking-effect [1]. The CS based image fusion may suffer from the reconstruction error.

Recently, dynamic image fusion is assumed to be a general challenging problem and has gained lots of attention. This problem was converted into an object detection based dynamic image fusion [10]. Lately, some researchers focused on the scheme of fusing spatial-temporal information [12, 13]. First, Kalman filtered compressed sensing [11] (KFCS) was presented for dealing with this problem. This method can capture spatial-temporal changes of multi-source videos by separable space-time fusion method. Second, Surfacelet transform was applied to video fusion by utilizing its tree-structure based directional filter banks [12]. Third, temporal information [13] was considered for providing more presented formation for the interest regions of the observation of the satellite in the field of remote sensing. Despite these works, the fusion pattern, considering spatial and temporal direction or other information, has not been explored fully. Beyond these problems, the spatial or temporal consistency remains as a challenging problem.

To preserve the spatial consistency and fuse structural information effectively, a novel image fusion method, based on generalized Riesz-wavelet transform [14], is proposed. Its main idea is to develop a heuristic fusion model, based on the capability of GRWT, to combine structural information adaptively and consistently. Its main feature lies in providing a generalization representation of low-level features based fusion pattern, which can be extended to other fusion problems easily. Meanwhile, the integration of high order Riesz

transformation and the proposed heuristic fusion model can keep and implement the fusion of structural information, such as gradient, contour or texture. This fusion pattern can detect and pick up low-level features by utilizing high-order steerability and its excellent angular selectivity [14]. Different from other MGA based image fusion methods, the GRWT based fusion method can improve the ability to keep the spatial consistency in high dimensions. The real-world experiments demonstrated that GRWT based fusion method achieves a fusion performance improvement, especially on the consistency of structural information.

The rest of the paper is organized as follows. A summary of generalized Riesz-wavelet transform is described in Section 2. Section 3 presents the details of the proposed heuristic fusion model. Section 4 presents experimental results based on multi-modality images. At last, discussions and conclusions are presented in Section 5 and 6, respectively.

2. Generalized Riesz-wavelet transform

2.1. Riesz transform and its high order extension

Riesz transform [14] can be viewed as a natural extension of Hilbert transform. It is a scalar-to-vector signal operation. For the Hilbert transformation, it performs as an all-pass filter, whose transfer function can be defined as follows

$$\hat{T}(w) = -j \text{sign}(w) = -jw / \|w\|, \quad (1)$$

where $\hat{T}(w)$ stands for the transfer function in frequency-domain, $T(w)$ for the space-domain version, and w for the frequency variable. Based on the definition of Hilbert transform, Riesz transform can be defined in the frequency-domain as follows

$$Ri[s(w)] = -j \frac{w}{\|w\|} \hat{s}(w), \quad (2)$$

where $\hat{s}(w)$ is the Fourier transform of the input signal $s(w)$ and $Ri[s(w)]$ denotes Riesz operator on the signal $s(w)$. The representation of this transformation in space-domain can be reformatted as follows

$$Ri[s(x)] = \begin{pmatrix} Ri_1 s(x) \\ \dots \\ Ri_d s(x) \end{pmatrix} = \begin{pmatrix} (h_1 * s)(x) \\ \dots \\ (h_d * s)(x) \end{pmatrix}, \quad (3)$$

where d denotes the dimension of $Ri[s(x)]$, and the filters $(h_n)_{n=1}^d$ are denoted as the resulting frequency responses $T_n(w) = -jw_n / \|w\|$. The space-domain based equation can be expressed directly as follows

$$y(x) = F^{-1}\{\|w\|^{-1}\}(x), \quad (4)$$

where $F^{-1}(\bullet)$ denotes the inverse fast Fourier transformation. Eq. (4) can be viewed as the impulse response of isotropic integral operator $(-\Delta)^{-1/2}$. More exactly, these operations, performed by $Ri_i s(x)$ s, can be viewed as partial derivatives of $y(x)$. Table 1 demonstrates the connection between the differential operators and the Riesz transformation.

Remark: Riesz transformation has a natural connection to part derivative or gradient operator. The details about this method can refer to the paper [14].

The higher-order Riesz transform with respect to the input signal is defined as follows

$$Ri_{i_1} Ri_{i_2} \dots Ri_{i_N} s(x), \quad (5)$$

where $i_1, i_2, \dots, i_N \in \{1, \dots, d\}$ denotes N -th individual signal components with different order Riesz transformation. There exist d^N ways to construct N -th order terms. The directional behavior of generalized Riesz-wavelet transform can be obtained as follows

$$Y_u[s(w)] = \sum_{i=1}^d u_i \frac{-jw_i}{\|w\|} \hat{s}(w), \quad (6)$$

where Y denotes Fourier transform version of $y(x)$, and $u = [u_1, \dots, u_i, \dots, u_d]$ is the unit vector for angular selectivity [14]. Iterating N times, the previous equation can be expressed as follows

$$\hat{T}_u^N s(w) = \sum_{|n|=N} \frac{N!}{n!} u^n \left(\frac{-jw}{\|w\|} \right)^n \hat{s}(w). \quad (7)$$

The formation of space-domain, corresponding to previous Eq. (7), can be expressed as follows

$$\hat{T}_u^N s(w) = \sum_{|n|=N} c_{n_1, \dots, n_d}(u) s^{n_1, \dots, n_d}(x). \quad (8)$$

where $c_{n_1, \dots, n_d}(u)$ denotes the steering coefficients by N -th order Riesz transform. Moreover, $n_1, \dots, n_d$ denotes multi-index vector for representing the N -th order Riesz transform. $c_{n_1, \dots, n_d}(u)$ can be obtained by

$$c_{n_1, \dots, n_d}(u) = \sqrt{\frac{N!}{n_1! \dots n_d!}} (u_1^{n_1} \dots u_d^{n_d}). \quad (9)$$

Table 1. Summary of Riesz transform and other differential operators [14] with $d = 1, 2, 3, 4$.

Operator	$(-\Delta)^{-\frac{1}{2}}$	$Ri^n = -\frac{\delta}{\delta x_n} (-\Delta)^{-\frac{1}{2}}$	$(-\Delta)^{\frac{1}{2}}$
Frequency response	$\frac{1}{\ w\ }$	$\frac{-jw_n}{\ w\ }$	$\ w\ $
Impulse response			
$d = 1$	$\frac{\log x }{\pi}$	$\frac{1}{\pi x}$	$\frac{-1}{\pi x ^2}$
$d = 2$	$\frac{1}{2\pi\ x\ }$	$\frac{x_n}{2\pi\ x\ ^3}$	$\frac{-1}{2\pi x ^3}$
$d = 3$	$\frac{1}{2\pi^2\ x\ ^2}$	$\frac{x_n}{\pi^2\ x\ ^4}$	$\frac{-1}{\pi^2 x ^4}$
$d = 4$	$\frac{1}{4\pi^2\ x\ ^2}$	$\frac{3x_n}{4\pi^2\ x\ ^5}$	$\frac{-3}{4\pi^2 x ^5}$

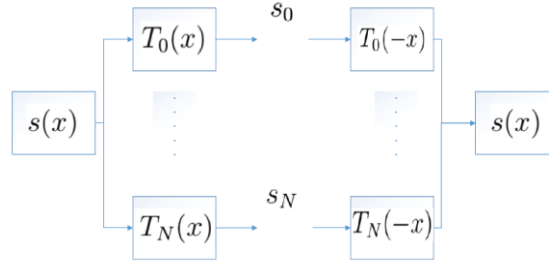


Fig. 1. Steerable filter banks of GRWT

2.2. 2D Generalized Riesz-wavelet transform

The 2D-version generalized Riesz-wavelet transform [14] can be given directly based on the definition of Riesz transform in Subsection 2.1. For N th-order Riesz transform with $d = 2$, there are $N + 1$ individual components. This equation leads to the subspace $\hat{V}_{N,2} = \text{span}\{\hat{T}_{n_1, N-n_1}(w)\}_{n_1=0}^N$. The explicit formation of $\hat{T}_{n_1, N-n_1}(w)$ can be given as follows

$$\hat{T}_{n_1, N-n_1}(w) = \sqrt{\binom{N}{n_1}} \left(\frac{-jw_1}{\|w\|} \right)^{n_1} \left(\frac{-jw_2}{\|w\|} \right)^{N-n_1}, \quad (10)$$

where we can compute $\cos(\theta) = w_1 / \sqrt{w_1^2 + w_2^2}$, and $\sin(\theta) = w_2 / \sqrt{w_1^2 + w_2^2}$ in the system of polar coordinate. These basic functions have a more simple formation (2π -periodic radial profile functions [14]), which can be expressed as follows

$$T_{m,n}(z) = (-j)^{m+n} \sqrt{\binom{m+n}{n}} \left(\frac{z+z^{-1}}{2} \right)^m \left(\frac{z-z^{-1}}{2j} \right)^n, \quad (11)$$

where $z = e^{j\theta}$, m, n are the orders of the frequency responds of different Riesz components, and the sum of these orders is equal to N . The correspond wavelet basic function or directional analysis based Wavelet transformation can be obtained as follows

$$\phi^{n_1, \dots, n_d}(x) = \sqrt{\frac{N!}{n_1! \dots n_d!}} R_{i_1}^{n_1} \dots R_{i_d}^{n_d} (-\Delta)^{\psi/2} \theta^{2\psi}(2x). \quad (12)$$

where $\psi > d/2$ is the order of the wavelet. The function $\theta^{2\psi}(2x)$ is the B-spline of order 2ψ , which is a smoothing kernel that converges to a Gaussian as ψ increases [14]. **Fig. 1** presents the flow chart of GRWT. Based on Eq. (12), the Riesz-wavelet coefficients $C_i^{(n_1, \dots, n_d)}[k]$, can be obtained by

$$C_i^{(n_1, \dots, n_d)}[k] = \langle s, R^{(n_1, \dots, n_d)} \phi_{i,k} \rangle, \quad (13)$$

where k stands for the location of multi-resolution transformation. Similarly, i stands for the current decomposition level.

In summary, the mapping pattern provided by generalized Riesz-wavelet transform [14] preserves image structure with the L_2 -stability of the representation. This property can keep the multi-scale decomposition process from the blocking-effect and introducing artificial

artifacts. Moreover, this operation doesn't amplify the high frequency components. This decomposition is fast with a moderate redundant representation [14] compared to NSCT based signal decomposition method.

3. Heuristic fusion model and framework

3.1 Heuristic fusion model

In this subsection, the proposed fusion model is proposed. It is named with heuristic fusion model, which can be expressed generally as follows

$$C(\sigma^2, u) = \sum_{i=1}^N \sigma_i^2 u_i, \quad (14)$$

where σ_i^2 is the degradation factor for the feature space u_i . σ_i^2 is not only able to project u_i into a suitable scale, but it also can weigh the u_i discriminatingly. u_i determines the importance of the correlation of the feature space $f_i, i = 1, \dots, N$. The lesser values of σ^2 represent less importance of u_i , while larger values correspond to more importance. The feature spaces are assumed to be leaded into a tensor based on the distance of the input low-level features, i.e., $u_i = \text{Dist}(f_1, f_2)$. Feature spaces f_1 and f_2 are established from each subband of GRWT from multi-modularity images. In this parameterization, the selection of u_i is denoted as the strength of the feature spaces' interaction.

For example, the fusion coefficient C_v , obtained by the visual images, can be presented explicitly by

$$C_v(\sigma_1^2, \sigma_2^2, u_1, u_2) = \sigma_1^2 u_1 + \sigma_2^2 u_2, \quad (15)$$

where u_1 and u_2 denote the strength of the feature spaces of image phase and coherence respectively, which can be extracted from the input multi-modality images, which located in each subband of GRWT. C_v denotes the fusion coefficient, which is calculated for the visual image. The determination of σ_1^2 and σ_2^2 is corresponded to the selection of the feature space, especially based on the volume.

Remark: It can be noted that the proposed heuristic fusion model is a generalized representation of the feature-based image fusion pattern. The types of low-level features may include image phase, coherence, orientation and regions, etc. In this paper, two feature spaces are generated via image phase denoted by u_p , and image coherence by u_c . These features can span a tensor-based feature spaces, which may provide a potential research direction to develop a heuristic or additive feature based fusion model.

At last, according to a common assumption in the context of image fusion [1], the sum of all the fusion coefficient C_i in the fusion process is equal to 1. For example, the fusion process of the visual and infrared images can be expressed mathematically

$$C_v + C_I = 1, \quad (16)$$

where the symbols V and I are abbreviated with visual and infrared images respectively. These fusion coefficients are determined by Eq.(15). In other words, the weighting process of the fusion coefficient C_i can be viewed as a convex combination.

3.2. Proposed fusion method

In this section, the fusion process is proposed through GRWT. The flow chart of the proposed fusion framework is presented in Fig. 2. We assume that the input image signals are spatially registered. The main workflow is summed up as follows

1. Input two images, denoted by image A and B .
2. Registering the input two images.
3. Selecting the order of 2D-Riesz transformation. Its definition can be referred to Eq.(5).
4. Choosing the number of decomposition levels of GRWT, which determines the number of subbands.
5. Analyzing multi-modality images A and B by GRWT based on Eq.(12) and Eq.(13). This process produces various transformation coefficients in different scales.

$$DC_{Ai} = GRWT(A), DC_{Bi} = GRWT(B), \quad (17)$$

where DC_{Ai} and DC_{Bi} denotes the decomposition coefficients of image A and B respectively, based on GRWT, in scale i .

6. Extracting the corresponding feature space based on the decomposition coefficients in each decomposition scale, which can be expressed as follows

$$\begin{aligned} PC_{Ai} &= Phase(DC_{Ai}), & PC_{Bi} &= Phase(DC_{Bi}), \\ CC_{Ai} &= Coherence(DC_{Ai}), & CC_{Bi} &= Coherence(DC_{Bi}). \end{aligned} \quad (18)$$

The function $Phase(\bullet)$ is aimed at extracting the image phase based on decomposition coefficients of image A and B . Similarly, the function $Coherence(\bullet)$ is a function for calculating the image coherence. PC_{Ai} and PC_{Bi} denotes the phase coefficients of image A and B respectively. CC_{Ai} and CC_{Bi} are the image coherence coefficients of image A and B respectively. The details steps to obtain the information of image phase and coherence can refer to the reference [21].

7. Based on these definitions, the features u_1 and u_2 can be obtained by

$$u_1 \leftarrow Dist(PC_{Ai}, PC_{Bi}), \quad u_2 \leftarrow Dist(CC_{Ai}, CC_{Bi}). \quad (19)$$

where the function $Dist(\bullet, \bullet)$ denotes the distance between two coefficients matrix. In this paper, this distance is chosen to be chessboard distance [23] because of the larger value indicates the better stability of the fusion images.

8. Reconstructing the fusion image by the fused coefficients FC_i in scale, the fusion procedure can be summed as follows

$$C_A(\sigma_1^2, \sigma_2^2, u_1, u_2) = \sigma_1^2 u_1 + \sigma_2^2 u_2. \quad (20)$$

$$FC_i(x, y) = C_A(\sigma_1^2, \sigma_2^2, u_1, u_2) \circ DC_{Ai} + (1 - C_A(\sigma_1^2, \sigma_2^2, u_1, u_2)) \circ DC_{Bi}. \quad (21)$$

where the symbol $\circ$ denotes the Hadamard product.

9. Based on the fused coefficients FC_i , the resulting fusion image F can be generated by

$$F = \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} FC_i(x, y) \phi_{i,k}^n, \quad (22)$$

where i denotes the scale or the decomposition level, $k \in \mathbb{Z}^2$ stands for the location, i.e., (x, y) , and the symbol $\phi_{i,k}^n(x)$ denotes the n -th order Riesz-wavelet basis function [14].

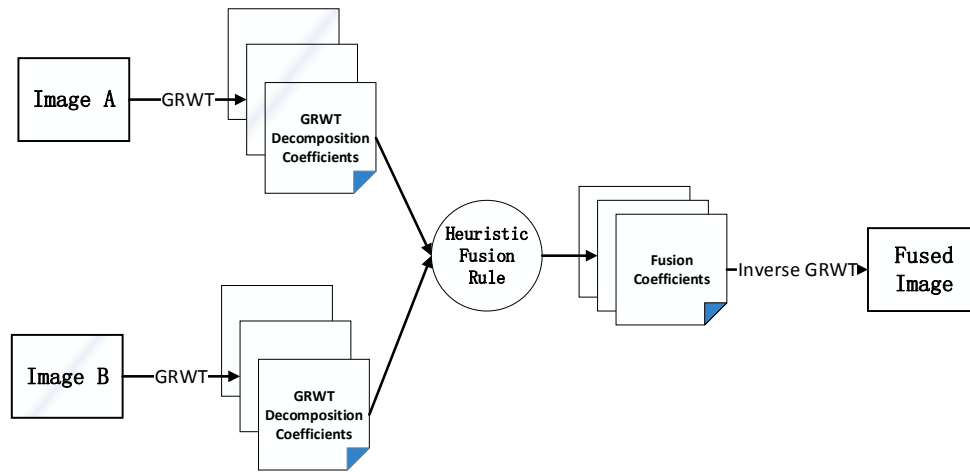


Fig. 2. The flow chart of GRWT based fusion method

4. Experiments

4.1 Quantitative evaluations

To assess the effectiveness of the proposed fusion, empirical experiments are performed on multi-modality images. Five objective indexes are taken for the evaluation of fusion performance, which consist of entropy [1] (EN), mutual information [1] (MI), structural similarity index [15] (SSIM), feature similarity index [16] (FSIM) and edge information preservation index [17] denoted by $Q^{ab/f}$. The definitions of FSIM, EN and MI can be referred in Appendix A. It should be noted that the larger values of these evaluation indexes indicate that the better fusion performance performed by the fusion methods. For the FSIM and SSIM, they would produce two numerical results based on two referred input images. In this paper, the larger ones are selected for representing the real fusion capability of the referred fusion methods. It should be noted that the completion and fidelity of structural information of the fusion image is important to the success of image fusion. To evaluate the computational cost of these fusion methods, the execution time, denoted by Time(s), is also taken for assessing the time complexity of all referred fusion methods in second(s). The experiments are performed on a computer shipped with Intel Core Quad CPU Q6700, and equipped with 3G RAM. All algorithms are implemented in Matlab 2010.

The proposed method is compared to five fusion methods, including wavelet, dual tree-complex wavelet transform (DT-CWT), low redundant discrete wavelet frame transform [3] (LRDWF) and discrete wavelet frame transform [2] (DWF) and Shearlet [19]. For the referred multi-scale decomposition methods, the decomposition levels are chosen to be four, the fusion rule is the selection of the largest decomposition coefficients, and the basic function is chosen to be 'db4'. For the proposed GRWT based fusion method, when the order of Riesz transform is 1, and the decomposition level is 4, it gains a better fusion performance in our experiments.

4.2 Fusion results by navigation images

To assess the proposed method in different imaging situations, we test the fusion methods on navigation image. These images were captured in visual and infrared image. The samples of

the navigation image are displayed in Fig. 3. The size of navigation images is 512×512 . The resulting visual fusion results are presented in Fig. 4. The numerical results are presented in Table 2. It can be seen that GRWT based method can construct a more complete representation of the perceived scene than other fusion methods. Although the visual result of Shearlet based method is similar to the fine detail of GRWT based method, the numerical results indicated that Shearlet's fusion process, based on $Q^{ab/f}$, SSIM and FSIM, may damage the perception of the local image content. Clearly, the proposed method outperforms the other fusion methods because of the proposed fusion model and its ability to select and reconstruct structure information. At last, it can be concluded that the presented fusion performance validated the effectiveness of the proposed fusion method again.

GRWT based fusion method can make a balance between the fusion performance and computation requirement. Compared to Wavelet, LRDWF, DT-CWT and DWF based fusion methods, Shearlet based fusion method completed higher fusion performance in term of EN, MI, SSIM and FSIM. In a numerical view of these fusion processes, Shearlet [19] based fusion method required 9.6650 second, but GRWT based fusion only required 0.9135. Compared to the other fusion methods, the proposed fusion algorithm just slightly higher than them. It can be seen that Wavelet based fusion method required the least time about 0.1480 second, but its overall fusion performance is the lowest among compared fusion methods.

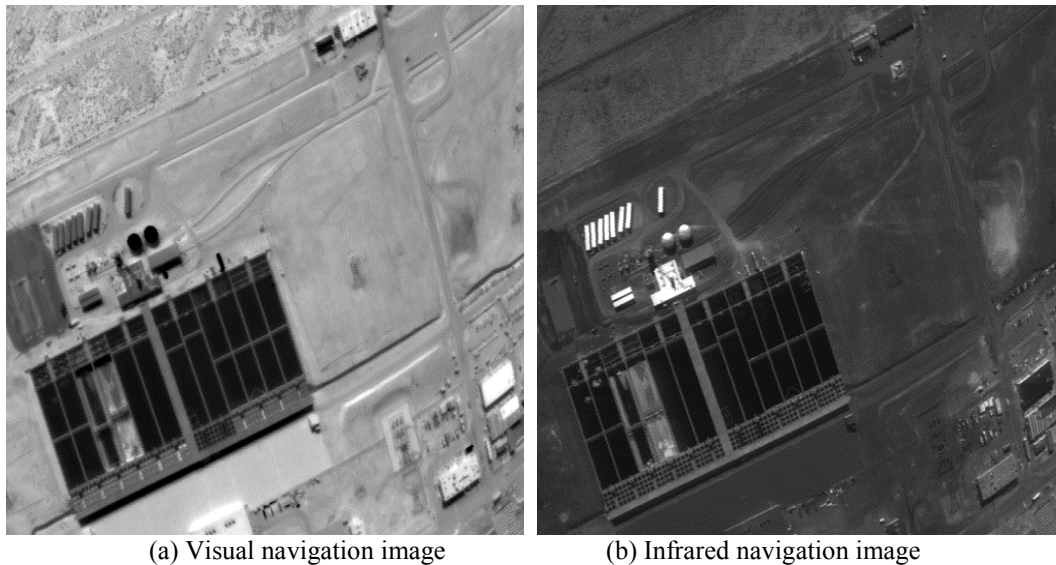


Fig. 3. Samples of navigation images

Table 2. Fusion performance evaluation by navigation images

Algorithm	EN	MI	SSIM	FSIM	$Q^{ab/f}$	Time (s)
Wavelet	6.7131	2.1534	0.7975	0.7417	0.5072	0.1480
LRDWF [3]	6.6397	2.3263	0.8314	0.7590	0.5684	0.2726
DT-CWT	6.6510	2.3065	0.8294	0.7551	0.5725	0.5173
DWF [2]	6.6344	2.3441	0.8343	0.7594	0.5778	0.5013
Shearlet [19]	6.8683	3.1231	0.8580	0.8905	0.4304	9.6650
GRWT	6.9874	3.9650	0.9785	0.9832	0.6175	0.9135

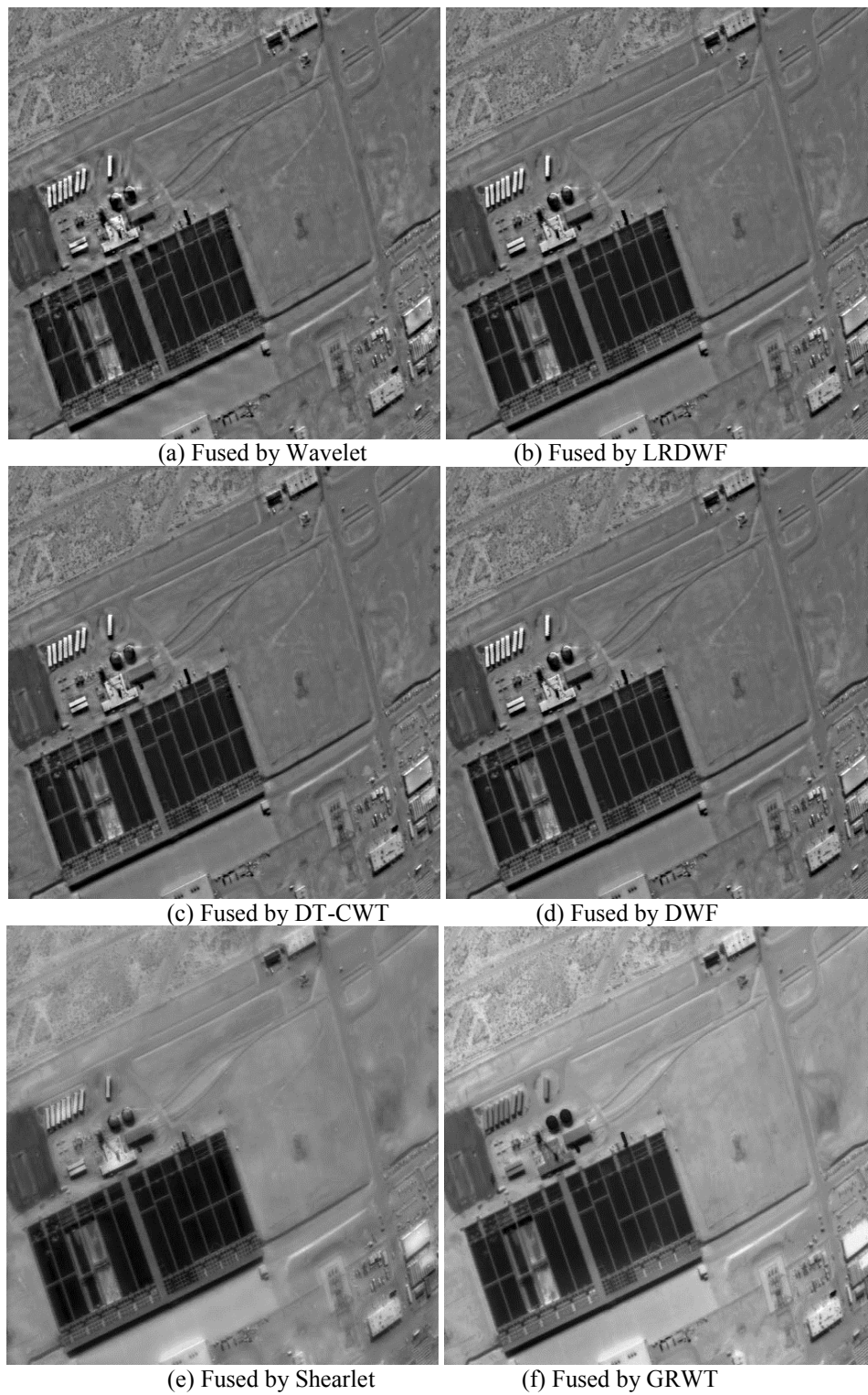


Fig. 4. Visual comparison of six fusion methods by navigation images

4.3 Fusion results by visual and near-infrared images

In this subsection, some experiments are performed for assessing the fusion performance on visual and near-infrared images. The dimension of these two images is 256×256 . The samples of these two images are presented in Fig.5. The numerical outcomes are illustrated in Table 3. Although Shearlet based method's visual contrast is higher than GRWT's, its numerical evaluation results indicate that the fusion process by Shearlet based method may destroy local structures of the scene, which results in low fusion performance in term of EN, $Q^{ab/f}$, SSIM and FSIM. In summary, the fusion performance, obtained by GRWT based method, demonstrated that its fusion pattern can capture and select natural structural information or low-level features, and complete a better fusion performance improvement. Meanwhile, it can be seen that the computation time of the proposed fusion method behaves as the section 4.2. The visual results are displayed in Fig. 6. It should be noted the fusion result of GRWT based method contains more contrast details than other fusion methods. The visual results of Wavelet, LRDWF, DT-CWT and DWF lost some contrast information originated in the original scenes. The reason for these outcomes is that the smoothing effect of these multi-scale decomposition methods would discard some contrast details in some degree. In other words, these methods suffer from the loss of local contrast or the damage of structural information. The numerical evaluation results indicate the proposed fusion method preserves more high-order structural information, such as gradient and texture. It is clear that the fusion image created by GRWT based method is superior to other fusion methods.

Table 3. Fusion performance evaluation by visual and near-infrared images

Algorithm	EN	MI	SSIM	FSIM	$Q^{ab/f}$	Time(s)
Wavelet	6.7737	2.6208	0.9099	0.6530	0.7069	0.0653
LRDWF [3]	6.7687	2.7431	0.9145	0.6721	0.7399	0.1180
DT-CWT	6.7641	2.7788	0.9135	0.6712	0.7515	0.2020
DWF [2]	6.7653	2.7698	0.9179	0.6761	0.7459	0.1906
Shearlet [19]	6.7965	3.1925	0.8503	0.8704	0.4028	3.1208
GRWT	6.9275	4.3387	0.9859	0.9810	0.7695	0.2120



(a) Visual image

(b) Near-infrared image

Fig. 5. Samples of visual and near-infrared images

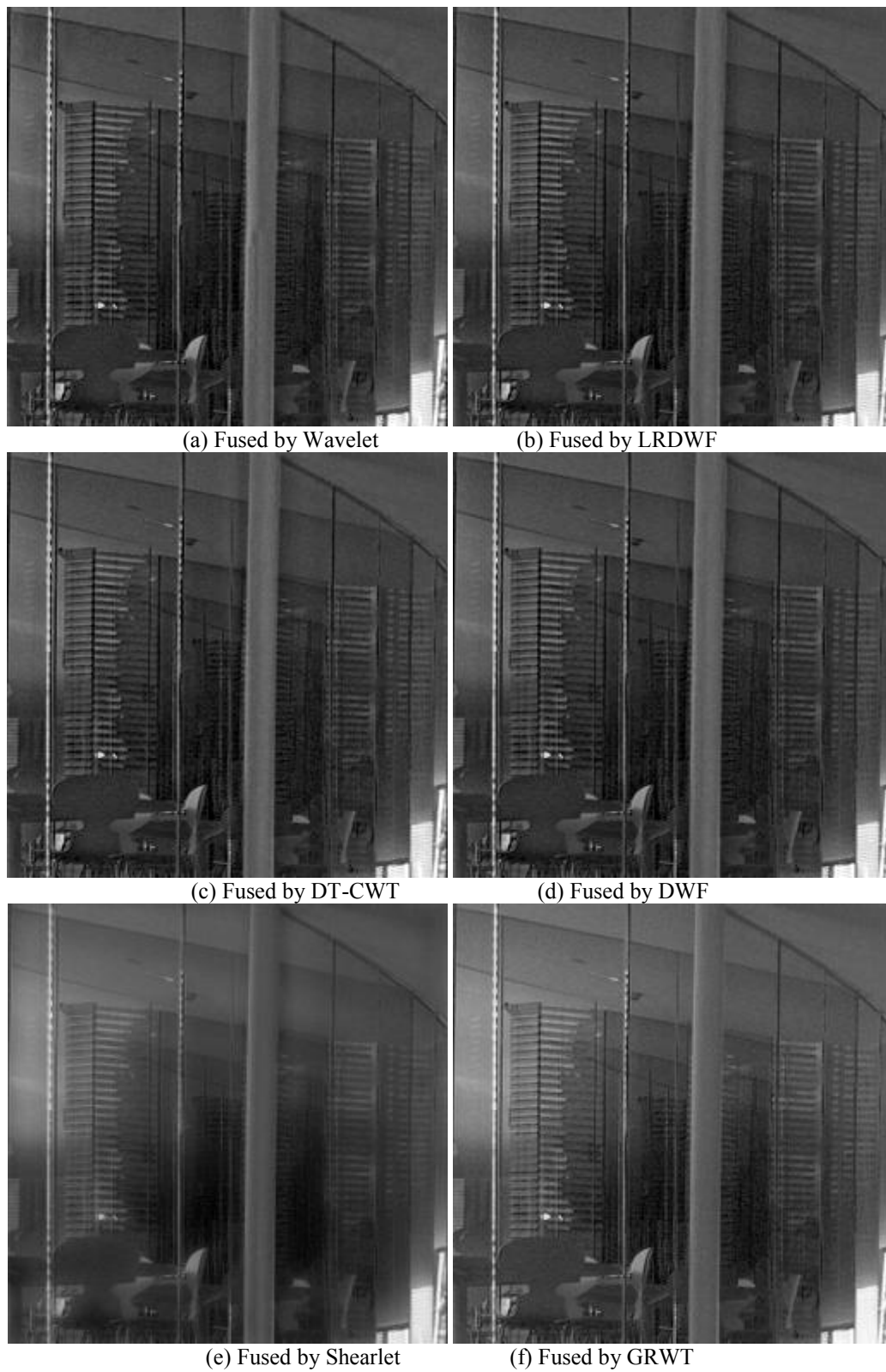


Fig. 6. Visual comparison of six fusion methods by visual and near-infrared images

4.4 Fusion results by medical images

In this section, the proposed method's fusion performance is assessed numerically and inspected visually on CT and MRI images. The sample images are displayed in Fig. 7. The size of these images is 256×256 . It should be noted that our method is better than other fusion methods in term of EN, SSIM, FSIM and $Q^{ab/f}$ based on Table 4. These numerical results indicated that the proposed fusion method can transfer more information than other methods. Moreover, the computational requirement of the proposed fusion method almostly reaches to the outcomes of the DT-CWT and DWF based fusion methods. For a visual examination of the fused image by GRWT, displayed in Fig. 8, it can be seen that GRWT based fusion method contains more salient or details than other fusion methods. In other words, different features from these two images are combined into Fig. 8(f). This phenomenon, generated by GRWT based methods, can be supported by the real fusion performance measured by SSIM, FSIM and $Q^{ab/f}$. The definitions of these indexes indicate that structural information, such as the information of gradient, texture and edge, is integrated into the final fused result efficiently. In other words, the proposed method can complete a comprehensive fusion performance improvement with regard to objective evaluation indexes and visual effects.

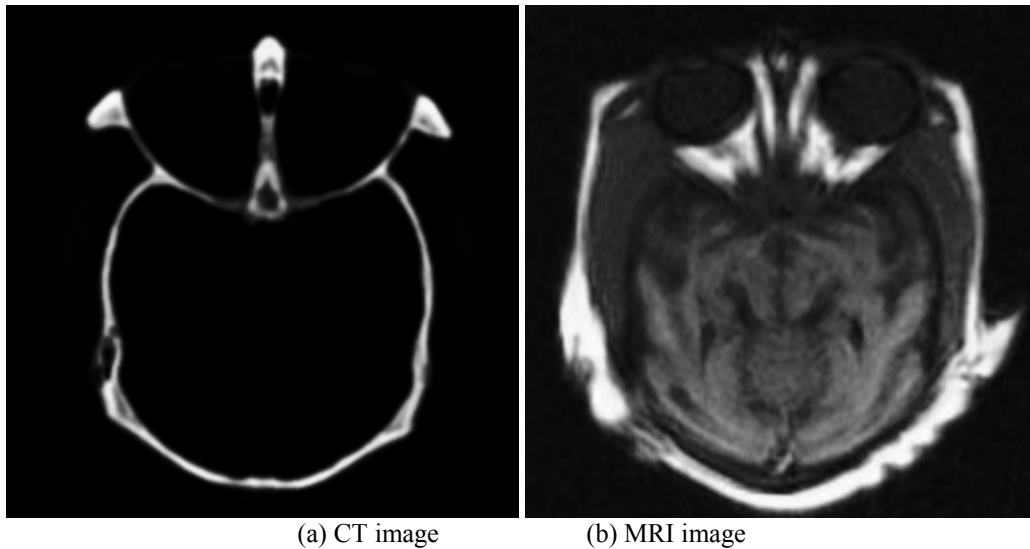


Fig. 7. Samples of CT and MRI images

Table 4. Fusion performance evaluation by medical image

Algorithm	EN	MI	SSIM	FSIM	$Q^{ab/f}$	Time(s)
Wavelet	6.2166	1.6202	0.6146	0.7440	0.4045	0.0777
LRDWF [3]	6.1896	1.9580	0.6883	0.7736	0.4861	0.1110
DT-CWT	6.1768	1.9280	0.6887	0.7707	0.4832	0.2103
DWF [2]	6.6344	2.0002	0.6953	0.7777	0.5017	0.2001
Shearlet [19]	6.9238	2.3176	0.7244	0.8476	0.6308	2.9344
GRWT	6.9714	2.5253	0.7581	0.9152	0.6879	0.2024

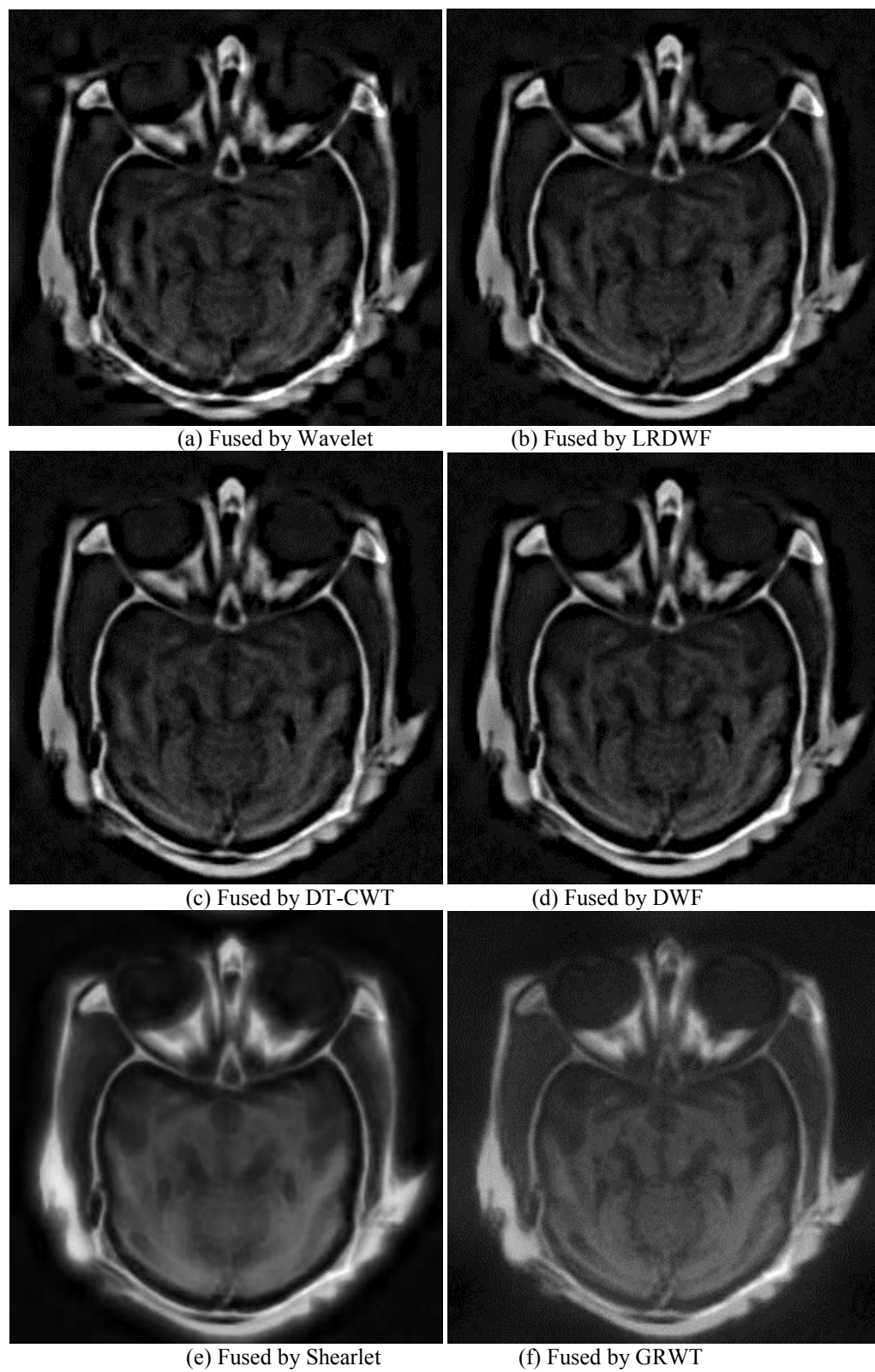


Fig. 8. Visual results of six fusion methods

4.5 Statistical fusion results by 20-pair visual and near-infrared images

To investigate the fusion performance fully, the proposed method was assessed on a classical public dataset, which was captured by a Canon camera with appreciating modifications. All fusion methods were applied to 20 pair images. The details of sampling the near-infrared image [20] can refer to the web site: <http://ivrg.epfl.ch/research/infrared/imaging>. The original size of these images is 1024×679 . To benefit the evaluation process of all fusion methods, the input images are cropped into a square size with 512×512 . These images are presented in Fig. 9. Meanwhile, this section would not display the resulting fusion images for the sake of the paper space. Table 5 presented the average numerical results based on 20-pair visual and near-infrared images. It can be seen again that GRWT based method outperforms other fusion methods in term of five objective fusion indexes. For the MI index, GRWT based method behaves much better than other fusion methods. Other four evaluation indexes, i.e., MI, SSIM, FSIM and $Q^{ab/f}$, indicate that GRWT based method can transfer more structural information into the fusion results. Similar to the execution times in Table 2, 3 and 4, the proposed fusion method can accomplish a balance between the fusion performance and computational cost again.

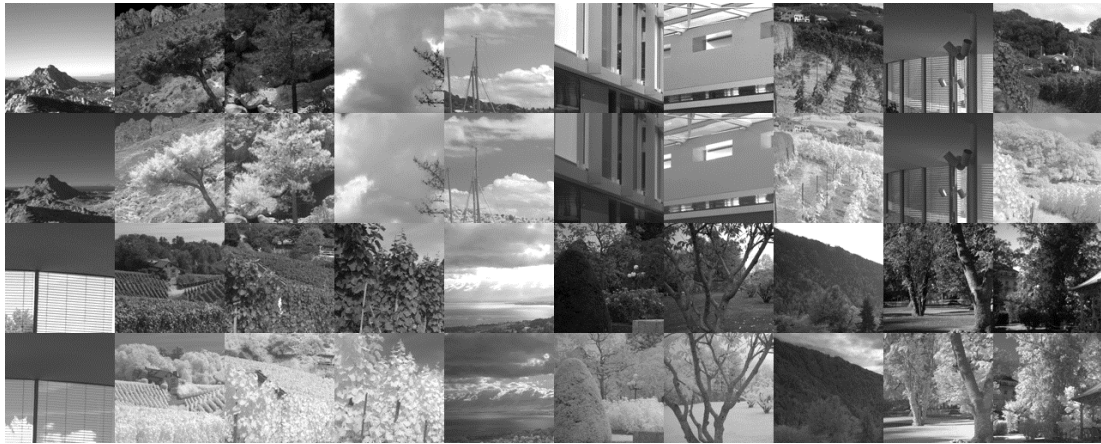


Fig. 9. Samples of 20 pair visual and near-infrared images

Table 5. Statistical numerical results by visual and near-infrared dataset

Algorithm	EN	MI	SSIM	FSIM	$Q^{ab/f}$	Time(s)
Wavelet	6.6911	4.4643	0.9405	0.8708	0.7010	0.1451
LRDWF [3]	6.6804	4.6162	0.9478	0.8815	0.7257	0.2706
DT-CWT	6.6749	4.6618	0.9499	0.8804	0.7383	0.4973
DWF [2]	6.6773	4.6626	0.9499	0.8779	0.7359	0.4816
Shearlet [19]	6.6613	4.8197	0.9355	0.9459	0.6394	9.4765
GRWT	6.7622	5.6641	0.9923	0.9919	0.7597	0.9196

5. Discussion

A general challenging problem in MRA based image fusion methods is the consistency of fusion result, which is owing to the representation accuracy of the MRA based method and the fusion rule. The proposed method can deal with this problem in some degree, which can be validated by visual results in Fig. 4(f), Fig. 6(f) and Fig. 8(f). Visual inspection of these visual results indicates that other MRA based methods may damage the local contrast features or

produce visual artifacts. The numerical results, based on [Table 2](#), [Table 3](#) and [Table 4](#), also specified that GRWT based method not only keeps the contrast of the input images, but it also preserves the spatial consistency of contour, gradient and texture. This promising process can be verified by numerical results provided by EN, MI, $Q^{ab/f}$, SSIM and FSIM.

Beyond these, the presented method's superiority lies in the preservation of image content coherence. The Shearlet based fusion methods may cost about eight times original image's computational cost, much more than GRWT. In real-time application, these behaviors may not be permitted. The proposed method can alleviate these situations in some degree and have a balance between the complexities of space and time.

6. Conclusion

A novel image fusion method, integrated by heuristic fusion model and generalized Riesz-wavelet transformation, is presented. Exploiting the proposed fusion model's excellent ability to investigate and select structure information, the proposed method can combine image content efficiently. A variety of experiments illustrated that the congruency of phase and gradient magnitude is important to the success of image fusion method. The numerical and visual results provided by five objective indexes and visual examination, have shown that the presented fusion method is suitable for multi-modality image fusion. Moreover, GRWT based fusion method can capture salient features with sharper intensity changes, and keep the consistency of directional edge and texture.

Appendix A

Mutual information (MI) [1]

MI is aimed at measuring the mutual similarity between the two input images. For the two digital image signals in pixels, MI can be expressed by

$$MI(x; y) = H(x) + H(y) - H(x, y),$$

And $H(x, y)$ is joint entropy, which is defined by

$$H(x, y) = -\sum_{i=0}^{255} P_i(x, y) \log P_i(x, y),$$

where $P_i(x)$ is the normalized histogram of the pixel value x . And the sum of the $P_i(x)$ is assumed to be 1, i.e., $\sum_{i=0}^{255} P_i(x) = 1$. And $P_i(x, y)$ is denoted by joint normalized distribution histogram about the pixel values x and y . Similarly, $P_i(x, y)$ is constraint to $\sum_{i=0}^{255} P_i(x, y) = 1$.

Entropy(EN) [1]

The EN can measure the amount of the information has been merged, which can be given by

$$EN = -\sum_{i=0}^{255} P_i \log_2 P_i,$$

where P_i denotes a normalized histogram of the fused image's all pixels.

Feature-similarity (FSIM) index [14, 18]

The definition of feature similarity index is defined by

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \bullet PC_m(x)}{\sum_{x \in \Omega} PC_m(x)},$$

where the symbol $S_L(x)$ is denoted by

$$S_L(x) = [S_{PC}(x)]^\alpha [S_G(x)]^\beta,$$

where Ω denotes the domain of the image plane, and $S_G(x)$ and $S_{PC}(x)$ are the distance functions for measuring the similarity of the gradient magnitude (GM) and phase congruency (PC).

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61175028) and the Ph.D. Programs Foundation of Ministry of Education of China (Grant Nos. 20090073110045).

References

- [1] Z. Jing, G. Xiao and Z. Li, "Image fusion: Theory and applications," *Higher Education Press*, Beijing, October, 2007.
- [2] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *Proc. of International Conference on Image Processing*, IEEE, Vol. 3, pp. 288-291, October, 1997.
[Article \(CrossRef Link\)](#)
- [3] B. Yang and Z. Jing, "Image fusion using a low-redundancy and nearly shift-invariant discrete wavelet frame," *Optical Engineering*, vol. 46, no. 10, 107002-107002-10, October, 2007.
[Article \(CrossRef Link\)](#)
- [4] H. Wang, Z. Jing, J. Li and H. Leung, "Image fusion using non-separable wavelet frame," *Intelligent Transportation Systems, 2003. Proceedings. IEEE*, vol. 2, pp. 988-992, October 12-15, 2003. [Article \(CrossRef Link\)](#)
- [5] Y. Chai, H. Li and M. Guo, "Multifocus image fusion scheme based on features of multiscale products and PCNN in lifting stationary wavelet domain," *Optics Communications*, vol. 284, no. 5, pp. 1146-1158, March 1, 2011. [Article \(CrossRef Link\)](#)
- [6] B. Guo, Q. Zhang and Y. Hou, "Region-based fusion of infrared and visible images using nonsubsampling contourlet transform," *Chinese Optics Letters*, vol. 6, no. 5, pp. 338-341, May 1, 2008. [Article \(CrossRef Link\)](#)
- [7] Q.G. Miao, C. Shi, P.F. Xu, M. Yang and Y.B. Shi, "A novel algorithm of image fusion using shearlets," *Optics Communications*, vol. 284, no. 6, pp. 1540-1547, March 15, 2011.
[Article \(CrossRef Link\)](#)
- [8] S. Li, H. Yin and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 9, pp. 4779-4789, September, 2013. [Article \(CrossRef Link\)](#)
- [9] J. Han, O. Loffeld, K. Hartmann and R. Wang, "Multi image fusion based on compressive sensing," in *Proc. of 2010 IEEE International Conference on Audio Language and Image Processing (ICALIP)*, pp. 1463-1469, November 23-25, 2010. [Article \(CrossRef Link\)](#)
- [10] C. Liu, Z. Jing, G. Xiao, B. Yang, "Feature-based fusion of infrared and visible dynamic images using target detection," *Chinese optics letters*, vol. 5, no. 5, pp. 274-277, May 10, 2007.
- [11] H. Pan, Z. Jing, R. Liu, B. Jin, "Simultaneous spatial-temporal image fusion using kalman filtered compressed sensing," *Optical Engineering*, vol. 51, no. 5, pp. 057005-1, May 22, 2012.

- [Article \(CrossRef Link\)](#)
- [12] Q. Zhang, L. Wang, Z. Ma and H. Li, "A novel video fusion framework using surfacelet transform," *Optics Communications*, vol. 285, no. 13-14, pp. 3032-3041, June 15, 2012.
[Article \(CrossRef Link\)](#)
 - [13] H. Song and B. Huang, "Spatiotemporal satellite image fusion through one-pair image learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 1883-1896, April, 2013.
[Article \(CrossRef Link\)](#)
 - [14] M. Unser and D. Van De Ville, "Wavelet steerability and the higher-order Riesz transform," *IEEE Transactions on Image Processing*, vol. 19, no. 3, pp. 636-652, March, 2010.
[Article \(CrossRef Link\)](#)
 - [15] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April, 2004. [Article \(CrossRef Link\)](#)
 - [16] L. Zhang, D. Zhang and X. Mou, "FSIM: a feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, August, 2011.
[Article \(CrossRef Link\)](#)
 - [17] C. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308-309, February 17, 2000. [Article \(CrossRef Link\)](#)
 - [18] B. Han, G. Kutyniok and Z. Shen, "Adaptive multiresolution analysis structures and shearlet systems," *SIAM Journal on Numerical Analysis*, vol. 49, no. 5, pp. 1921-1946, June 20, 2011.
[Article \(CrossRef Link\)](#)
 - [19] W.Q. Lim, "The discrete shearlet transform: A new directional transform and compactly supported shearlet frames," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1166-1180, May, 2010. [Article \(CrossRef Link\)](#)
 - [20] Y. M. Lu, C. Fredembach, M. Vetterli and S. Susstrunk, "Designing color filter arrays for the joint capture of visible and near-infrared images," in *Proc. of 16th IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 3797-3800, November 7-10, 2009. [Article \(CrossRef Link\)](#)
 - [21] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transaction on Signal Processing*, vol. 49, no. 12, pp. 3136-3144, December, 2001. [Article \(CrossRef Link\)](#)
 - [22] W.W. Kong, Y.J. Lei, Y. Lei, and S. Lu, "Image fusion technique based on non-subsampled contourlet transform and adaptive unit-fast-linking pulse-coupled neural network," *Image Processing, IET*, vol. 5, no. 2, pp. 113-121, March, 2011. [Article \(CrossRef Link\)](#)
 - [23] R. Plamondon and H.D. Cheng, "Pattern recognition: architectures, algorithms & applications," *World Scientific*, Vol. 29, May 14-18, 1991. [Article \(CrossRef Link\)](#)



Bo Jin received his B.S. degree in electronic information and electrical engineering from Shanghai Jiao Tong University, China, in 2004. He is currently pursuing the Ph.D. degree in electronic information and electrical engineering at Shanghai Jiao Tong University, Shanghai, China. His major research interests include machine learning, incremental learning, face recognition, and visual tracking, image fusion.



Zhongliang Jing received his B.S., M.S., and Ph.D. degrees from Northwestern Polytechnical University, Xi'an, China, in 1983, 1988, and 1994, respectively, all in Electronics and Information Technology. Currently, he is Cheung Kong professor, and executive dean at the School of Aeronautics and Astronautics, Shanghai Jiao Tong University, China. Prof. Jing is an editorial board member of the Science China: Information Sciences, Chinese Optics Letters as well as International Journal of Space Science and Engineering. His major research interests include multi-source information acquisition, processing and fusion, target tracking, and aerospace control.



Han Pan was born in GuangXi, PR. China in 1983. He received his Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2014. He is currently a postdoctoral fellow at Shanghai Jiao Tong University, Shanghai, China. His research interests include image restoration, information fusion, and convex optimization.