Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-018-0873-5

ORIGINAL RESEARCH

An effective image fusion algorithm based on grey relation of similarity and morphology

Caiping Liu[1] · Yahui Long[1] · Jianxu Mao[2] · Hui Zhang[3] · Ruizhi Huang[1] · Yang Dai[2]

Received: 12 January 2018 / Accepted: 19 May 2018

Abstract
Considering the problems of the limited energy in wireless multi-media sensor networks (WMSNs) and the focused regions discontinuity of the fused image obtained using traditional multi-scale analysis tools (MST)-based methods, an effective multi-focus image fusion algorithm is proposed in this paper. In this method, the original fused image is obtained based on wavelet transform where the low-frequency coefficients are fused by average scheme, whereas the high-frequency coefficients are fused by the proposed merging rule consisting of the grey relation analysis of similarity and local area energy. Then, grey absolute relation analysis is again utilized as measurement indicator to estimate the similarities between the initial fused image and source images, during which the initial map is acquired and then corrected by the mathematical morphological opening and closing. Finally, the fused image is obtained with the guidance of the corrected map, namely the decision map. Experiment results demonstrate that the fused image using the proposed algorithm is more continuous in focused region and more similar to source images in brightness compared with state-of-art multi-focus image fusion algorithms, such as Curvelet transform, lifting stationary wavelet transform (LSWT), non-subsampled contourlet transform (NSCT) and non-subsampled shearlet transform (NSST). Meanwhile, the proposed method shows better superiority in term of the computational complexity.

Keywords Image fusion · Multi-focus image · Grey relational analysis · Mathematical morphology · Computational complexity

1 Introduction

Due to the limited depth of focus in optical lenses, it is difficult to describe the complex situation with a single image accurately (Shreyamsha et al. 2013). In wireless multi-media sensor networks (WMSNs), image fusion techniques are important and have been widely applied to obtain a full-focus image. Image fusion can integrate multiple images from several sensors or the same sensor at different time into a single image which combines complementary, multi-temporal or multi-view information from source images (Iovane et al. 2011; Yang et al. 2016). The fused image is more suitable for human visual perception and computer-processing tasks such as segmentation, feature extraction and target recognition (Nirmala et al. 2013; Cao et al. 2016; Wang et al. 2015; Fan et al. 2015).

Recently, image fusion methods are basically divided into two categories, one is based on spatial domain and the other based on transform domain (Miti-anoudis and Stathaki 2008). The former one directly processes the pixels in spatial

✉ Jianxu Mao
maojianxu@hnu.edu.cn

Caiping Liu
liucaiping@hnu.edu.cn

Yahui Long
yahuilong@hnu.edu.cn

Hui Zhang
zhanghuihby@126.com

Ruizhi Huang
huangruizhi@hnu.edu.cn

Yang Dai
daiyang@hnu.edu.cn

[1] College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China

[2] College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

[3] College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410012, China

Published online: 14 June 2018

domain according to merging rules among which one of the simplest methods is to take average of image pixel by pixel. This kind of method can effectively extract the useful information from source images to the fused image and enhance the quality of images. However, side effects such as low contrast are inevitable in the fused image because of the simplicity of the methods. Recently, with the deepening of fusion technique theories, many researchers have proposed plenty of multi-scale analysis tools (MST), such as Pyramid transform (Kakerda et al. 2015), Wavelet transform (WT) (Ellmauthaler et al. 2013; Iovane et al. 2011; Xie et al. 2015), Curvelet transform (Ma et al. 2012), Contourlet transform (CT) (Zhao et al. 2017), non-sample contourlet transform (NSCT) (Bhateja et al. 2015; Chen et al. 2014; Heshmati et al, 2016; Kong et al. 2015; Wang et al. 2015; Yang et al. 2014; Yu et al. 2015), Shearlet transform, non-sampled shearlet transform (NSST) (Gao et al. (2013); Yin et al. 2014) and so on.The Pyramid transform based methods can accurately identify the feature information in focused regions, and, thus, have a better fusion performance than the spatial domain based ones. However, the relevance among different coefficients can easily lead this kind of method to be instable. As the most commonly used MST, the wavelet transform can decompose the details on different scales and directions, which makes it more accurate to estimate the sharpness of pixels. As a result, the fused result can be obviously improved. However, the wavelet transform cannot effectively describe the line and plane singularities of source images. Moreover, it can capture only limited directional information due to the limitation in the direction of the 2D separable wavelets. It thus cannot represent the directions of the edges of source images accurately (Yang et al. 2014). In order to overcome the limitation of decomposition direction of wavelet, many MTSs are gradually proposed and become more and more popular in image fusion domain, such as Curvelet transform, CT, NSCT, NSST and so on. These MTSs have better time-frequency characteristics than wavelet transform. In addition, they can decompose source images into more directions at the same scale and describe straight line and curve line very well. Therefore, fused images obtained by the methods based on these MTSs have a better performance and are much closer to the standard image. Unfortunately, the decomposition of more direction makes it so complex and time-consuming that it is not suitable for the application in the high real-time situation, such as WMSNs. In WMSNs, besides the high real-time requirement, the network requires that the algorithms should be quite energy-efficient, as the nodes can only store limited energy. Hence, it is significantly easy for the nodes to fail due to the high energy consumption when the image fusion algorithm with high computational complexity is applied to the network.

In addition, the MST-based methods probably result in the situation that plenty of useful information of source images will have different degrees of loss in the fused image. This problem principally results from the improper selection of merging rule in the process of multi-scales decomposition and reconstruction. For example, the high frequency coefficients describe the detailed components of the source images, such as the edges, textures, boundaries and so on. Generally, used fusion rule is to take average of coefficients, which probably leads to the change of the grey values of coefficients. In addition, most of the MST-based image fusion schemes assume that input images are ones without noise. In fact, images captured from the scene are usually corrupted by noise. Therefore, there might be inevitable phenomenon of distortion in the fused images obtained based on the MST-based methods. Also, the conventional MST-based methods show poor robustness, as they can only process input images with registration very well and cannot obtain desired fusion results for input images without registration.

To address problems mentioned above, an effective multi-focus image fusion algorithm based on grey relation of similarity and mathematical morphology technique is proposed in this paper. The novel method effectively integrates grey relational analysis with mathematical morphology technique. Firstly, the initial fused image is obtained based on wavelet transform after which a novel fusion rule composed of grey absolute relation with local region energy is proposed to merge the high frequency coefficients. And then, grey absolute relation is selected as measurement indicator to estimate the similarities between the initial fused image and source images, which, as a result, will produce initial decision map. And the mathematical morphological opening and closing are employed for post-processing. Finally, the fused image is obtained with the guidance of the result, namely the decision map.

The rest of the paper is organized as follows. Section 2 introduces Grey Relation Analysis. The proposed method is described in Sect. 3 in detail. Section 4 shows experiment results and analysis. The conclusion is given in Sect. 5.

2 Grey relation analysis

Grey System Theory (GST), proposed by Professor Deng Julong from China in Deng (1982), is a new discipline which refers to system engineer based on mathematical theory. It is considered as a new way to investigate few data, poor information and uncertain problems. As an important part in GST, grey relational analysis (GRA) is a kind of impact evaluation model evaluating the degree of similarity or difference between two sequences based on their relation Goyal and Grover (2012). The basic principle is to compare and evaluate geometry shape between two curve lines which consist

of sequences. The more similar the geometry shape is, the stronger the relevance between sequences is Ma et al. (2012). Recently, many researchers have recognized the absolute advantage of grey relational analysis in processing small sample data and, thus, apply it to image fusion domain. Ma et al. (2012) selected Euclid relation as a metric to analyse the relevance among the high frequency coefficients in Curvelet domain. He et al. (2007) utilized Deng relation to distinguish if pixels come from edged region or not in the synthetic aperture radar (SAR) image.

In this paper, we employ a new grey relational analysis model named grey absolute relation (GAR) to determine the relevance among pixels in wavelet transform domain. Li and Yang (2010) presented grey absolute relational model. However, Liu and He (2011) found that the model had significantly high computational complexity and the results showed improper phenomenon. Therefore, Liu proposed an improved grey absolute relational model. The new model describes the similarities between sequences by the summation of the absolute values of subtractions between two sequences. Experiment results demonstrate that the improved grey absolute relational model is such easier to be calculated and even has better performance of evaluation than common grey relational analysis models. The important steps to calculate grey absolute relation can be summarized as follows:

Step 1 Pick out the referential sequence $H_0 = \{h_0(k)|k = 1, 2, ..., N, N \in \text{int}\}$ from the observation system and the comparative sequences $H_i = \{h_i(k)|i = 1, 2, ..., t, k = 1, 2, ..., N, t \in \text{int}\}$, respectively, where N, t represents the length of sequences and the number of the comparative sequences, respectively.

Step 2 Normalize both the referential sequence and the comparative sequences so that it becomes free from any units. The results are expressed as $H_i^0 = H_iD = \left\{h_i^0(k)|i = 0, 1, 2, ..., t, k = 1, 2, ..., N\right\}$, $H_iD = (h_i(1)d, h_i(2)d, ..., h_i(N)d)$ where $h_i(k)d = h_i^0(k) = h_i(k) - h_i(1)$, and D, H_iD are separately considered as normalized operator and normalized map of H_i.

Step 3 Calculate the grey absolute relation values between the referential sequence H_0 and the comparative sequences H_0 as follow:

$$\xi_{0i} = \frac{1 + |s_0| + |s_i|}{1 + |s_0| + |s_i| + |s_0 - s_i|} \tag{1}$$

where $|s_0| = \left|\sum_2^n (h_0^0(k))\right|$, $|s_i| = \left|\sum_2^n (h_i^0(k))\right|$, $|s_0 - s_i| = \left|\sum_2^n (h_0^0(k) - h_i^0(k))\right|$.

3 The proposed image fusion scheme

As shown in Fig. 1, the main process of the proposed image fusion scheme is accomplished by the following steps:

Step 1 Obtain the initial fused image based on the wavelet transform according to the new proposed merging rules.

Step 2 Detect focus region by comparing and evaluating the similarities between the initial fused image and source images pixel by pixel based on the improved grey absolute relation, and, thus, produce the initial decision map.

Step 3 Modify the initial decision map using mathematical morphology technique to obtain the decision map.

Step 4 Reconstruct the fused image with the guidance of the decision map.

3.1 Initial fusion image based on wavelet transform

In this paper, the initial fused image plays an important role in the proposed image fusion scheme. On one hand, the process of followed focus region detection is on the foundation of the initial fused image. The better the performance of the initial fused image is, the more accurate the detection becomes. On the other hand, the pixels located at the

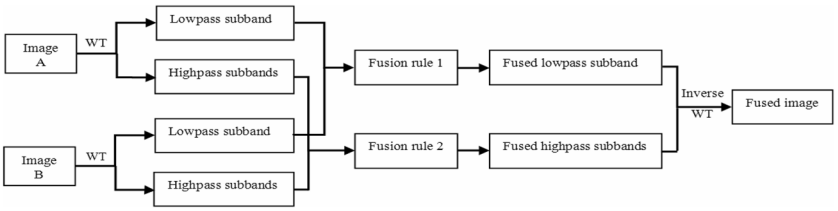

Fig. 1 Schematic diagram of the initial fusion image

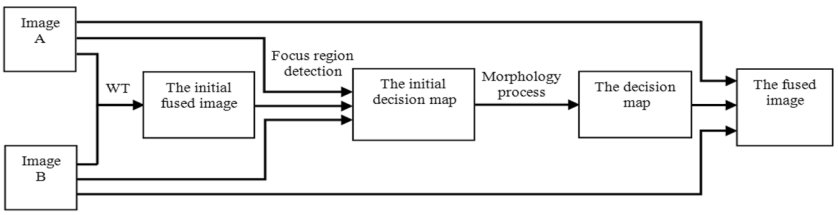

Fig. 2 Schematic diagram of the proposed image fusion scheme

boundary of focus regions of the fused image are obtained by directly selecting the corresponding pixels of the initial fused image. As displayed in Fig. 2, the fusion process of the initial fused image is conducted as follows:

Step 1 Decompose source images into lowpass subband and highpass subbands based on wavelet transform. The level of decomposition is determined based on the experiment results.

Step 2 Integrate lowpass subband and highpass subbands according to corresponding fusion rules, respectively.

Step 3 Apply the inverse wavelet transform to obtain the initial fused image.

3.1.1 Fusion of lowpass subband

In fact, lowpass subband obtained by wavelet decomposion represents the approximate information of source images and contain the most energy of source images. Due to the high similarity among the lowpass subbands of images focusing on different regions, it is reasonable to select simple merge rule with the consideration of the tradeoff between fusion performance and computational complexity. The most popular rule is to take average of pixels. In this paper, we assume two source images (image A and image B) as input images (more than two source images are also suitable for the proposed method). We choose average scheme to integrate the lowpass subband as follows:

$$C(x,y) = 0.5 \times (C_A(x,y) + C_B(x,y)) \tag{2}$$

where $C_A(x,y)$, $C_B(x,y)$, $C(x,y)$ represent the coefficients at the location (x, y) of image A, image B and the initial fused image, respectively.

3.1.2 Fusion of highpass subbands

The highpass subbands represent the detail information of source images, such as edges, textures, boundaries and so on. Generally, the coefficients with larger absolute values are considered as the ones which contain more detail information of source images or represent sharper tendency of brightness changes (Yang et al. 2014). The most common fusion rule is maximum-absolute-value scheme. However, it is noteworthy that source images are often mixed up with more or less noise, which probably results in the erroneous estimation of salient features, and finally weakens the quality of fused image. Local region energy is experimentally demonstrated to be a great metric to estimate sharp salient of pixels (Guo et al. 2012; Ma et al. 2012). The larger the value of the local region energy is, the sharper the feature of the coefficient is. The fusion performance of the highpass subbands determines the quality of the initial fused image to a great extent. Also, the accuracy of the followed focus region detection will be indirectly influenced by its' performance. Because focus region detection is based on the initial fused image. Thus, in order to effectively and accurately identify and extract more useful information, a novel fusion scheme is proposed. The proposed rule consists of the improved grey absolute relation mentioned in Sect. 2 and local area energy. The proposed scheme is described as:

Step 1 Assume that the level of wavelet decomposition is 3, and divide the highpass subbands into blocks with the same size of $s \times s$ (Generally the experiment result is the best when s is equal to 3). Denote N_A, N_B as the local region block around $F^h_{l,r}(x,y)$ (which represents the coefficient located at (x, y) in the highpass subbands of source image h at the l-th scale and r-th direction) of image A and image B, respectively. Transform N_A, N_B to one-dimension sequences $N_A(t)$ and $N_B(t)$ $(t = 1, 2, ..., 9)$, and then, consider N_A and N_B as the referential sequence and the comparative sequence, respectively.

Step 2 Calculate the value of grey absolute relation ξ_{AB} between N_A and N_B based on Eq. 1. Meanwhile, the local energy values of each blocks of image A and image B can be separately calculated as follows.

$$E^{l,r}_A(x,y) = \sum\nolimits^{1}_{m=-1} \sum\nolimits^{1}_{n=-1} \left(F^A_{l,r}(x+m, y+n)\right)^2 \tag{3}$$

$$E^{l,r}_B(x,y) = \sum\nolimits^{1}_{m=-1} \sum\nolimits^{1}_{n=-1} \left(F^B_{l,r}(x+m, y+n)\right)^2 \tag{4}$$

where $E^{l,r}_h(x,y)$ represents the local region energy of the block around pixel (x, y) in the highpass subbands of source image h at the l-th scale and r-th direction, and m, n denote the radius of block in the horizontal and vertical direction, respectively.

Step 3 Denote $T \in (0, 1]$ as the threshold of grey absolute relation. If ξ_{AB} is equal to T or greater than T, this shows that the structure features between local region blocks $N_A(t)$ and $N_B(t)$ look extraordinarily similar, which illustrates that both of the coefficients at the centre of each block are probably helpful to enhance the sharpness of corresponding fused coefficient. Thus, we can combine the highpass subband according to Eq. 5.

$$F_{l,r}(x,y) = \begin{cases} L^{\max}_{l,r}(x,y)F^A_{l,r}(x,y) + L^{\min}_{l,r}(x,y)F^B_{l,r}(x,y), \\ if\ E^{l,r}_A(x,y) \geq E^{l,r}_B(x,y) \\ \\ L^{\min}_{l,r}(x,y)F^A_{l,r}(x,y) + L^{\max}_{l,r}(x,y)F^B_{l,r}(x,y), \\ if\ E^{l,r}_A(x,y) < E^{l,r}_B(x,y) \end{cases} \tag{5}$$

where $L^{\max}_{l,r}(x,y) = \frac{1}{2} + \frac{1-\xi_{AB}}{2(1-T)}$, $L^{\min}_{l,r}(x,y) = 1 - L^{\max}_{l,r}(x,y)$. On the contrary, if ξ_{AB} is less than T, this shows that there exists great difference between local region blocks $N_A(t)$and $N_B(t)$ in term of structure features. Generally, the coefficient with sharper feature can represent the focus characteristic of

fused coefficient better than the another one. Therefore, the highpass subbands can be combined as follows.

$$F_{l,r}(x,y) = \begin{cases} F^{A}_{l,r}(x,y) & if\ E^{l,r}_{A}(x,y) \geq E^{l,r}_{B}(x,y) \\ F^{B}_{l,r}(x,y) & if\ E^{l,r}_{A}(x,y) < E^{l,r}_{B}(x,y) \end{cases} \tag{6}$$

Step 4 Reconstruct the initial fused image based on the combination lowpass subband and highpass subband by taking inverse wavelet transform. And then the initial fused image is obtained.

3.2 Focus region detection

Among the conventional multi-focus image fusion algorithms based on the multi-scales analysis tools, most of them are effective to distinguish and extract useful features and can also achieve great fusion performance. However, as the inherent limitation of multi-scales analysis tools or fusion rules, it is significantly difficult to avoid mistakenly selecting the pixels from defocus region as the ones from focus region into fused image. Generally, coupled with the inevitable existence of noise in the input images, the fused images become discontinuous in focus region in term of multi-focus images. What's more, the addition of weighting operator in the merge rules will probably distort the brightness of fused images.

Mathematical morphology technique is applied to address the common problems mentioned above in this paper. As introduced in the Sect. 2, grey absolute relation has plenty of advantages, such as easy to calculate, strong ability to estimate variance. Thus, it is particularly reasonable to evaluate the similarities between the initial fused image and source images with it. Practically, the proposed image fusion scheme requests decrease the computational complexity at the condition of great performance as soon as possible. Therefore, it needs to consider the cost of computation for every small step. Given that the block-based scheme can seriously accelerate the process of focus region detection instead of the conventional pixel-based scheme, which can provide smaller input image for next morphology process, we select the block-based scheme here.There are four main steps to detect focus region as follows.

Step 1 Divide image A, image B and the initial fused image F into blocks M_A, M_B and M_F with the same size of $s_1 \times s_1$ (According to the experiment analysis, 5 is the optimal value of s_1, respectively. And transform M_A, M_B and M_F into one-dimension sequence $M_A(\tau)$, $M_B(\tau)$ and $M_F(\tau)$ $(\tau = 1, 2, ..., 25)$. Denote $M_A(\tau)$, $M_B(\tau)$ as the comparative sequences and $M_F(\tau)$ as the referential sequence.

Step 2 Calculate the values of grey absolute relation ξ_{FA}, ξ_{FB} between $M_F(\tau)$ and $M_A(\tau)$, $M_B(\tau)$, respectively.

Step 3 Obtain the initial map $Z(x_1, y_1)$, a binary image, by comparing ξ_{FA} and ξ_{FB} according to Eq. (7).

$$Z(x_1,y_1) = \begin{cases} 1 & if\ \xi_{FA} \geq \xi_{FB} \\ 0 & if\ \xi_{FA} < \xi_{FB} \end{cases} \tag{7}$$

where logical's 1 represents that block at the location (x_1, y_1) in image A is from focus region and presents white in the initial decision map. On the contrary, logical's 0 represents block at the location (x_1, y_1) in image B is from focus region and presents black in the initial decision map. For example, we select clean multi-focus image named lab as the experiment object to detect focus region by using the proposed method (as showed in Fig. 3a, b are focus on the left and right, respectively). The initial decision map is displayed in Fig. 3c. It is considered in Fig. 3c that the white and black regions represent the focus regions of Fig. 3a, b, respectively.

Step 4 However, the binary image Z cannot fully reflect the focused characteristics of the source image. What's more, the complexity of the details of image would cause the value of grey absolute relation in focused areas to not always be greater. Thus, there are thin protrusions, thin gulfs, narrow breaks, small holes, etc. in Z. According to the theory of imaging, the areas, either in focus or out of focus, are always continuous in the interior of these regions (Li and Yang 2010). Thus, the defects mentioned above should be removed from Z for obtaining great fused images. To remove these defects, the morphological opening and closing operators with small structure element are applied. Opening, denoted as $Z \cdot G$, is simply erosion of Z by the structure

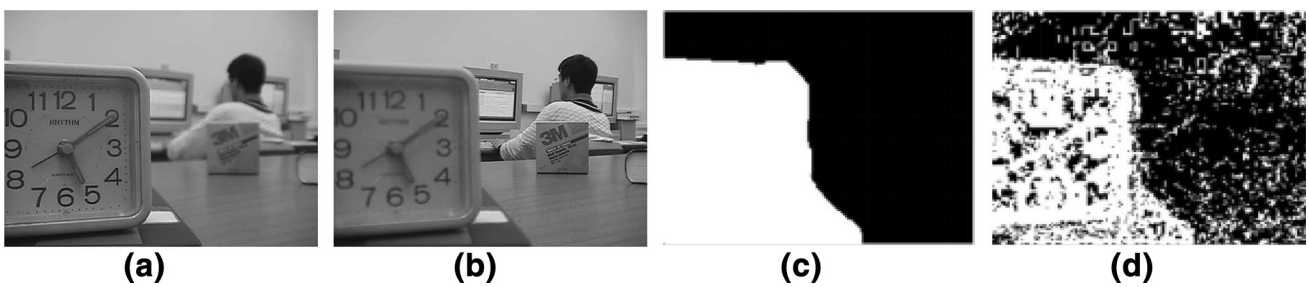

Fig. 3 **a** lab1; **b** lab2; **c** the decision map; **d** the initial map

element G, followed by the result by G . It can remove thin connections and thin protrusions. However, closing can join narrow breaks and fills long thin gulfs. It is dilation followed by erosion, and can be denoted as $Z \cdot G'$. Unfortunately, holes larger than G' cannot be removed only by using opening and closing operators. In fact, small holes are always misjudged. Hence, a threshold T' is set to remove the holes smaller than the threshold (Stathaki (2008)). Opening and closing are again implemented to smooth object contours. In the proposed focused region detection method, the structure element T' is a 4×4 matrix with logical's 1 and the threshold T' is determined according to the experimental results. The initial decision map Z becomes continuous by the process of the morphological opening and closing. The result is named as the decision map Z'. The decision map of source images 'lab' is showed in Fig. 3d. Obviously, the focus regions are always continuous except for the border regions, which indicates the proposed method is effective and reliable in processing multi-focus images.

3.3 Fuse the final fused image

According to the theory analysis, the white region and the black region in the decision map Z' separately represent different focus region of two input images. Therefore, the fused image can be directly and successfully obtained with the guidance of the decision map. However, the mathematical morphology technique can only deal with the continuous area very well, and cannot acquire desired performance at the boundary of focus regions. In order to enhance the quality of the boundary areas in the fused image, a window with fixed size around some pixel is introduced to determine where the corresponding combination pixel comes from.

As we know, the size of the decision map Z' is only one to five times of source images. First of all, it needs to be enlarged to be the same size with source images. We define the enlarged decision map as Z'' and the window matrix (assuming that the horizontal and vertical radius of which are m_1 and n_1, respectively) around pixel (x_1, y_1) as S. Therefore, the final fused image can be obtained as follows.

$$R(x_1,y_1) = \begin{cases} I_A(x_1,y_1) & \text{if } Z''(x_1,y_1) = 1 \text{ and } S(x_1,y_1) = m_1 n_1 \\ I_B(x_1,y_1) & \text{if } Z''(x_1,y_1) = 0 \text{ and } S(x_1,y_1) = 0 \\ F(x_1,y_1) & \text{if } 0 < S(x_1,y_1) < m_1 n_1 \end{cases} \tag{8}$$

where $S(x_1,y_1) = \sum_{m=-m_1}^{m_1} \sum_{n=-n_1}^{n_1} Z''(x_1+m, y_1+n)$, and $I_A(x_1,y_1)$, $I_B(x_1,y_1)$, $F(x_1,y_1)$ and $R(x_1,y_1)$ represent the pixels at the location (x_1,y_1) of image A, image B, initial fused image and final fused image, respectively.

In Eq. (8), $Z''(x_1,y_1) = 1$ and $S(x_1,y_1) = 0$ indicate that the pixel located at (x_1,y_1) come from continuous area in the decision map and the corresponding pixel of image A should be selected to the fused image.

Alternatively, $Z''(x_1,y_1) = 0$ and $S(x_1,y_1) = 0$ show that the pixel located at (x_1,y_1) come from continuous area in the decision map and the corresponding pixel of image B should be selected to the fused image.

Otherwise, $0 < S(x_1,y_1) < m_1 n_1$ means that the pixel at (x_1,y_1) is from border region in the decision map. We choose the corresponding pixel of the initial fused image as the one of the fused image to guarantee the perfect quality.

4 Experiments results and analysis

To evaluate the fusion performance and robustness of the proposed method, two sets of clean multi-focus images, two sets of mis-registration images and a set of images with noise are used to the experiments, respectively. They are shown in Fig. 4a, b, Fig. 5a, b, Fig. 6a, b) and Fig. 7a, b), respectively. To demonstrate the advantages of both great fusion performance and low computational complexity of the proposed method, the proposed method is compared with state-of-art multi-scales analysis methods, such as the Curvelet-based method (Ma et al. 2012), the lifting stationary wavelet transform (LSWT)-based method (Chai et al. 2011), the nonsubsampled contourlet transform (NSCT)-based method (Yang et al. 2014) and the nonsubsampled shearlet tansform (NSST)-based method. All of them can effectively distinguish and extract the detail features of multi-focus images. For the NSST-based method, three-levels NSST with 10,10,18,18 directions respectively and maxflat fitler are applied to each set of images, and lowpass subband and highpass subbands are fused based on average scheme and maximum-absolute-value scheme, respectively. In this paper, all the experiments are performed in MATLAB 7.0 on a 1.66 GHz Genuine Intel(R) CPU with 1.00 GB RAM.

The first set of experiments are realized on two sets of clean multi-focus images. Fig. 4a is one of source images 'clock', and focus on the left. The small clock close to the optics lens is in focus and clear visually, and, on the contrary, the large clock which keeps away from the optics lens is out of focus and blurring visually. The focus characteristic of Fig. 4b is completely opposite to that of Fig. 4a. Fig. 4c–g are the fusion results based on five types of different fusion methods mentioned above, respectively. It is recognized that the brightness of Fig. 4g is slightly better than other four fused images and "AT&T" letters in Fig. 4c is blurring. For better comparison, Fig. 4h–l display the differences between Fig. 4c–g and Fig. 4b, respectively. In difference images, the less residual features there are, the more successfully the corresponding fusion algorithm transfers useful information into fused

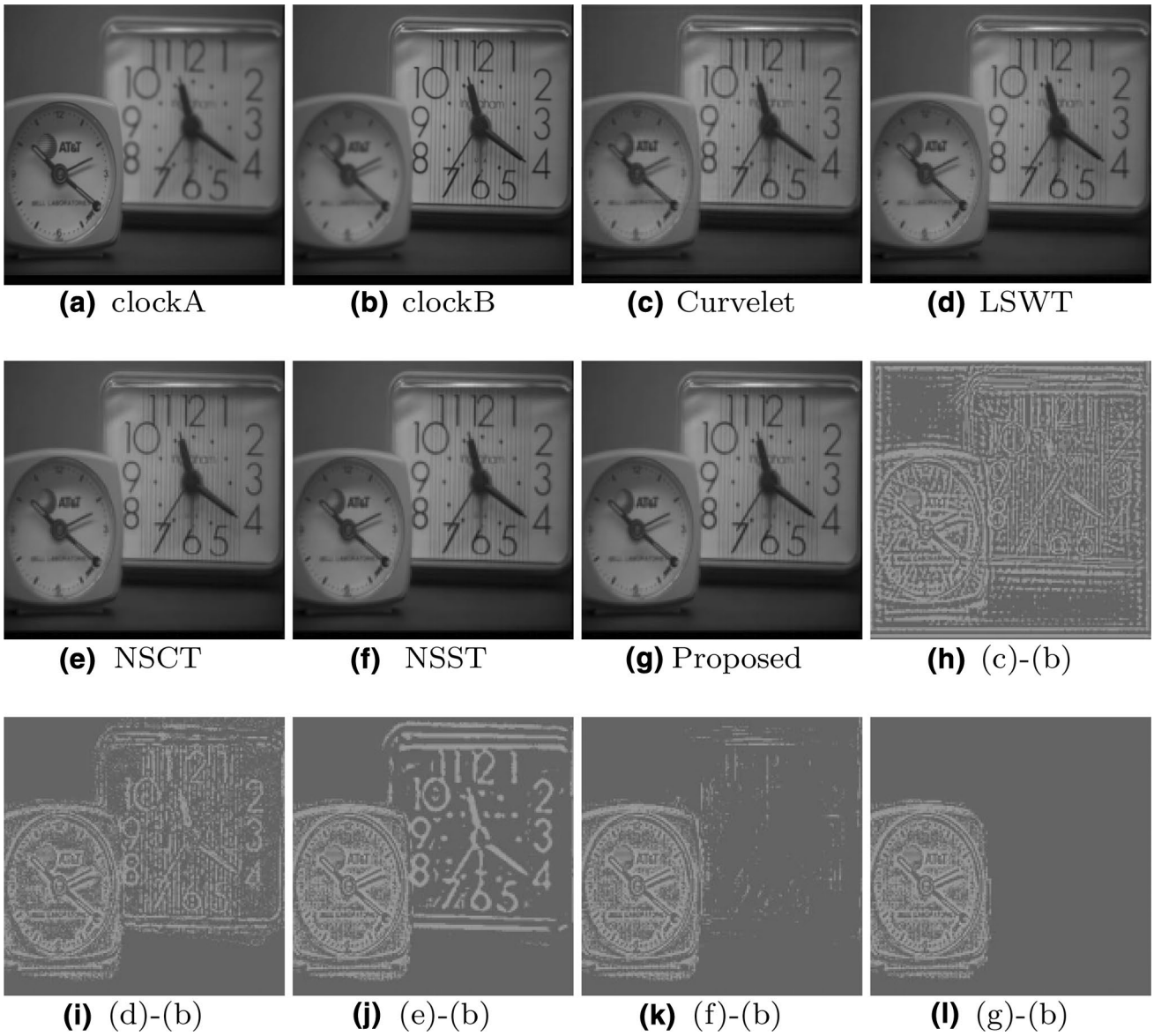

Fig. 4 Source images 'clock': **a** focus on the left and **b** focus on the right; fused images based on different fusion and methods: **c**–**g**; differences between fused images and (**b**): **h**–**l**

images (Li et al. 2012). In Fig. 4h–k, one can clearly see that the outline of the large clock are extremely distinct, which illustrates that the fusion algorithms mistakenly recognize and extract useless information to fused image. For example, in Fig. 4j, the Arabic numerals in the large clock can be found easily. Though Fig. 4k contains less residual information than the former three difference images, there are still a few spots in the right areas. However, Fig. 4l is smooth and leaves little residual information in focus region, which demonstrates that the proposed method has strong ability to distinguish features and successfully transfer more useful information into fused images. In other words, the proposed method is much more reliable and effective in fusing clean multi-focus images than the four kinds of compared algorithms.

To testify the performance of the proposed method, clean input images 'pepsi', as shown in Fig. 5a, b, are applied to the second experiment. The focus characteristic of this set of images is similar to that of the first set of ones. As displayed in Fig. 5c–g, they are the fusion results based on five kinds of different fusion methods, respectively. We can find that these letters in the middle and upper part look quite blurred visually in Fig. 5c–e, especially Fig. 5e. Both Fig. 5f, g are clearer and more natural than other three images, but the later looks much brighter than the former visually. In order to further comparison, the differences between Fig. 5c–g and

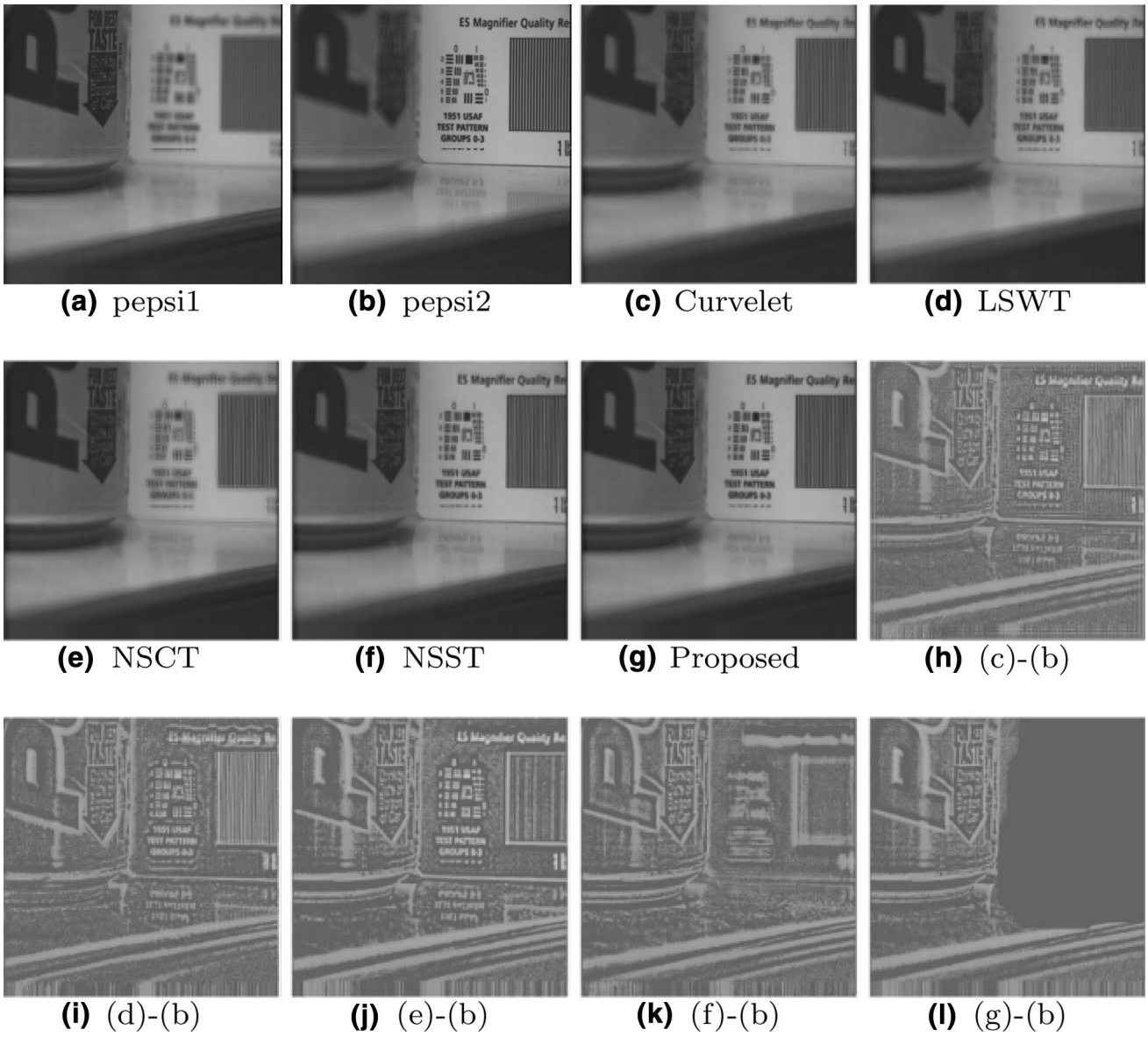

Fig. 5 Source images 'pepsi': **a** focus on the left and **b** focus on the right; fused images based on different fusion and methods: **c**–**g**; differences between fused images and (**b**): **h**–**l**

b are separately presented in Fig. 5h–l. It can be found easily that this set of difference images have similar situation to the first set of ones.

In this section, two sets of mis-registration images, are selected as input images to demonstrate the robustness and fusion performance of the proposed method. The focus situation for each set of source images is similar to that of the first experiment. For example, Fig. 6a, b are source images 'lab' focused on the left and right, respectively. Figure.6c–g display separately fusion results based on five types of different methods. Obviously, ringing artifact around the left hand of people can be found in Fig. 6c–e. Especially, Fig. 6e looks seriously blurred in the right region. Fig. 6f is clear visually as a whole, but the regions of people still look slightly blurring. However, Fig. 6g is clear and natural in vision. In addition, the brightness is closer to the input images in Fig. 6g than c–f. The difference images between the fused images and input images are showed in Fig. 6h–l. The results are similar to the former two sets of experiments, namely that Fig. 6h–k have plenty of residual information in focus region, whereas Fig. 6l is greatly smooth in the clock region and has little residual information except for the border regions of the clock.

The other source image is named 'flower', as shown in Fig. 7a, b. From Fig. 7c–g, the fusion results based on five

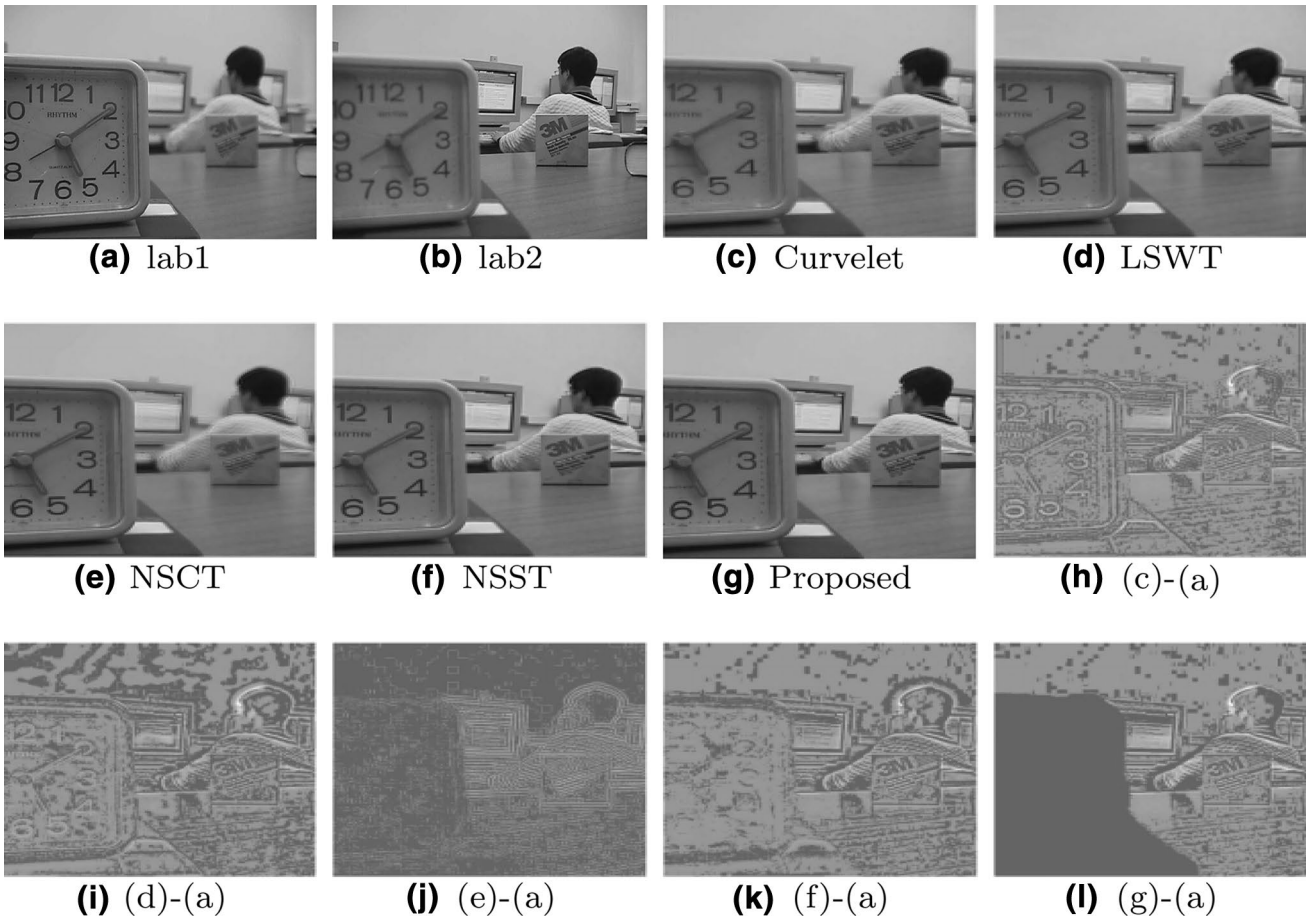

Fig. 6 Source images 'lab': **a** focus on the left and **b** focus on the right; fused images based on different fusion and methods: **c**–**g**; differences between fused images and (**a**): **h**–**l**

different algorithms, it can be observed that the distinctions among five images are extremely weak while the brightness of Fig. 7g outperforms these of other four images. Whereas, the differences between fused images and Fig. 7a focused on the left are also displayed in Fig. 7h–l, which is significantly helpful for comparative analysis. Theoretically, the left regions of the differences should be completely smooth and continuous due to the left focus characteristics of Fig. 7a. However, it can be obviously seen that there exist abundant residual information in the left region of all of the differences except for Fig. 7i, l. On the contrary, in the right region of the differences, more residual information shows that less useless information are successfully transformed into the fused image. Compared to Fig. 7i, there exists more residual information above the clock of Fig. 7g. The two sets of experiments about mis-registration images demonstrate that the proposed method is more effective and has better robustness in fusing multi-focus mis-registration images than other four fusion methods.

The fifth experiment is conducted on source images 'desk' which is a set of images corrupted with noise, as shown in Fig. 8a, b, respectively. This set of noisy images are produced by adding ten percents noise to the corresponding clean images artificially. The fused images based on five kinds of different methods are showed in Fig. 8c–g, respectively. In Fig. 8c–f, when focusing on the book region, it can clearly observed that fused images obtained based on different methods have the same phenomenon that the book region is blurred in vision. In addition, "3M" letters in the left region of Fig. 8c–f are also inconsistently blurring. However, compared with the former four fusion images, Fig. 8(g) is clearer in vision and closer to the input images in term of brightness. Fig. 8h–l present the differences between Fig. 8c–g and a, respectively. One can see that the clock areas in Fig. 8h–k contain the outline of the clock, which shows that corresponding fusion schemes successfully transfer plenty of useless information into fused images. However, Fig. 8l looks extremely smooth and continuous, and, what's more, the texture in defocus region are greatly clear and distinct, which indicates that the proposed method still outperforms the Curvelet-based method, the LSWT-based

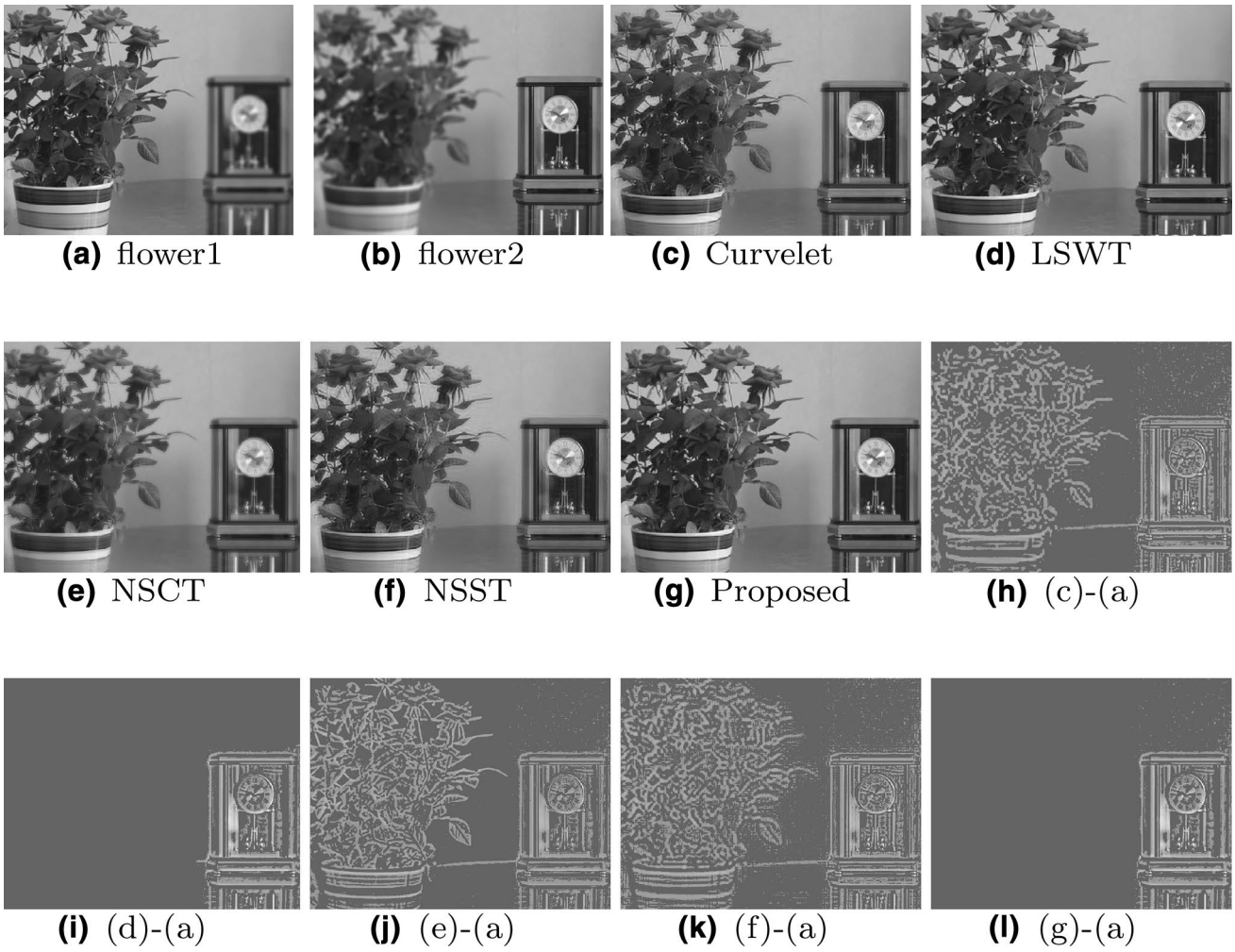

Fig. 7 Source images 'flower': **a** focus on the left and **b** focus on the right; fused images based on different fusion and methods: **c**–**g**; differences between fused images and (**a**): **h**–**l**

method, the NSWT-based method and the NSST-based method in processing multi-focus noisy images.

However, owing to the limitation of human visual system(HVS), it is insufficient to evaluate the fusion performance only from subjective perspective. For further comparison and analysis, it is significantly necessary to introduce objective criteria to estimate the fusion performance quantitatively. Many effective and accurate measurable indicators have been currently proposed, such as the average gradient (AG) (Xydeas and Petrovic 2000), the information entropy (IE) (Ye et al. 2008), the corresponding mutual information (MI) (Liu et al. 2016; Qu and Zhang 2002), the cross entropy (CE) (Jiang 2010; Liu et al. 2016)and the amount of edge information transferred from the source images to the fused image ($Q^{AB/F}$) (Liu et al. 2016; Qu et al. 2008), as well as the multi-scale contrast-based model (MCM) (Lu et al. 0000), the previous five of which are used in this paper. AG reflects the sharpness level of images. IE means the amount of information contained in images. MI indicates the sum of mutual information between every input images and the fused image. CE represents the difference between fused image and source image. $Q^{AB/F}$ shows the number of edge information successfully transferred from the input images into the fused image. As with AG, IE, MI and $Q^{AB/F}$, the greater the value is, the better the performance of fusion image is. Inversely, for CE, the smaller the value is, the better the performance of fusion image is.

As seen in Table 1, the objective evaluation of different fusion algorithms for different source images have been showed. In Table 1, the optimal value is marked by bold face for each column. It can be found easily that for clean input images 'clock', 'pepsi' and mis-registration images 'lab', 'flower', objective evaluation results are highly similar, that is to say, the values of the former four criteria of the proposed method are always the greatest among five fusion methods. Moreover, the value of the $Q^{AB/F}$ of the proposed

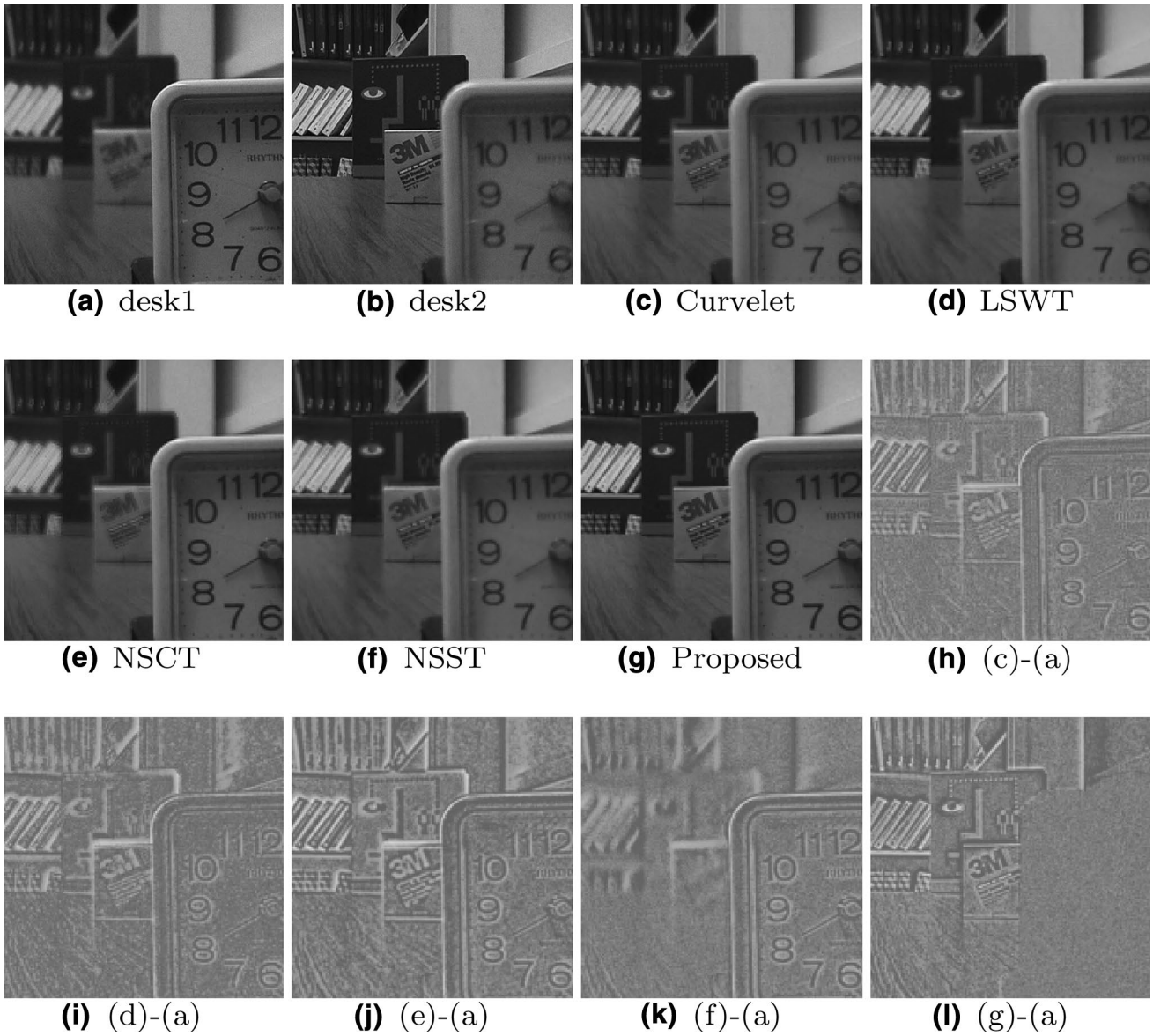

Fig. 8 Source images 'desk': **a** focus on the left and **b** focus on the right; fused images based on different fusion and methods: **c**–**g**; differences between fused images and (**a**): **h**–**l**

method ranks only second to the one of the LSWT-based method for 'clock', the NSST-based method for 'pepsi' and 'lab', the NSCT-based method for 'flower', respectively. The results illustrate that, as with multi-focus clean or mis-registration images, the proposed method can effectively distinguish salient features better and transfer more useful information into fused image than other four fusion methods, which completely coincides with the results of subjective evaluation. Focusing on noisy images 'desk', we can easily recognize that the former three metrics of the proposed method is much greater than ones of other four fusion methods. Though the last two metrics present poor performance, the proposed method is still the best one among five multi-focus fusion methods comprehensively. Hence, the proposed method superiors to other four conventional multi-scales analysis methods in fusing multi-focus noisy images.

With five sets of experiment analyses subjectively and objectively, it could be concluded that the proposed method is effective, robust and reliable, and outperforms the Curvelet-based method, the LSWT-based method, the NSWT-based method and the NSST-based method in processing clean, mis-registration or noisy multi-focus images.

Not only is the proposed method more effective to process multi-focus images than these traditional multi-scales analysis methods, but also has much lower computational

Table 1 Run-time(s) and objective evaluation of different fusion algorithm

Dataset	Method	Time (s)	AG	IE	MI	CE	$Q^{AB/F}$
Clock (256 × 256)	Curvelet	9.25	5.3475	7.011	2.4294	0.1976	0.9317
	LSWT	4.485	5.3344	7.0764	2.9956	0.013	**0.9456**
	NSCT	46.547	5.4746	7.0795	3.0121	0.0264	0.9426
	NSST	9.96	5.6642	7.0847	3.278	0.0104	0.9429
	Proposed	**3.983**	**5.677**	**7.0862**	**3.4612**	**0.0092**	0.9433
Pepsi (512x512)	Curvelet	35.469	4.9282	7.1204	2.1561	0.343	0.8428
	LSWT	17.97	4.1369	7.1034	2.2246	0.0276	0.7895
	NSCT	189.678	5.405	7.1205	2.391	0.0316	0.8358
	NSST	69.764	5.6547	7.1118	2.3601	0.0277	**0.8833**
	Proposed	**14.673**	**5.8267**	**7.121**	**3.1146**	**0.0266**	0.851
Lab (480 × 480)	Curvelet	33.074	3.9581	6.812	2.2058	0.2785	0.7882
	LSWT	16.79	3.6448	6.8068	2.1368	0.1661	0.8094
	NSCT	160.24	4.4309	6.8087	2.2357	0.1074	0.7982
	NSST	32.609	4.6332	6.8105	2.3552	0.408	**0.8441**
	Proposed	**13.062**	**4.6383**	**6.823**	**2.9444**	**0.1067**	0.8277
Flower (914 × 677)	Curvelet	45.326	4.4512	7.4086	2.0011	0.0524	0.8082
	LSWT	19.112	4.4364	7.4339	2.9631	0.0731	0.8091
	NSCT	232.063	4.0166	7.3721	2.0938	0.063	**0.8286**
	NSST	91.045	4.6003	7.4168	1.8821	0.0526	0.7861
	Proposed	**18.694**	**4.6414**	**7.456**	**2.9815**	**0.0776**	0.8108
Desk (480 × 480)	Curvelet	34.408	6.5056	7.0905	1.6724	0.3224	0.7582
	LSWT	16.1	6.8869	7.1284	1.5922	0.022	0.7083
	NSCT	33.045	8.3009	7.1761	1.5284	**0.0138**	0.72
	NSST	261.542	8.6256	7.2336	1.6324	0.0216	**0.7837**
	Proposed	**13.289**	**8.7774**	**7.2402**	**1.6821**	0.1064	0.7248

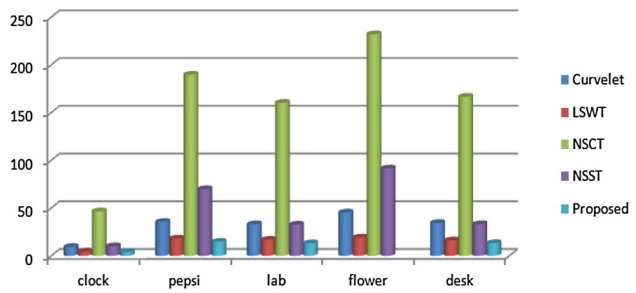

Fig. 9 Run-time(s) of different fusion methods for "clock", "pepsi", "lab", "desk", respectively

complexity. In this paper, we select run-time as evaluation criteria to estimate the complexity of fusion methods. As seen in Table 1, all of run-time results have been shown for each experiments. Similarly, the bold face represents the best performance. In Fig. 9, five varieties of color represent five types of different fusion methods, respectively, and the proposed method is indicated with light blue. It can be found distinctly that the column with green,representing the NSCT-based method, is always the highest among five fusion methods for any set of input images. However, the light blue representing the proposed method is always the lowest. The analysis results show that the proposed method is much more efficient or has lower computational complexity in processing multi-focus images than other four fusion methods.

5 Conclusion

In this paper, an effective multi-focus image fusion algorithm based on Grey Relation of Similarity and morphology is proposed. In the proposed method, the introduction of grey absolute relation enhances the accuracy to distinguish salient features in focus region, which, as a result, improves the fusion quality. Moreover, the low complexity characteristic of grey absolute relation simplifies the calculation of the proposed method. In addition, using decision map obtained by mathematical morphology technique to guide the final image fusion not only guarantees the continuity of focus region, but also promotes the reliability and robustness of the proposed method. Five sets of experiments, realized on different types of multi-focus images, demonstrate subjectively and objectively that the proposed method outperforms the traditional multi-scales analysis methods in processing multi-focus images, such

as the Curvelet-based method, the LSWT-based method, the NSCT-based method and the NSST-based method.

Acknowledgements The authors thank the editors and the anonymous reviewers for their detailed review, valuable comments and constructive suggestions. This work is supported by the National Natural Science Foundation of China (No. 61733004 and 61573134), National Science and Technology Support Program of the Ministry of Science and Technology of China (No. 2015BAF13B00), Natural Science Foundation of Human Province of China (No. 2018JJ3079).

References

Bhateja V, Patel H, Krishn A, Sahu A, Lay-Ekuakille A (2015) Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains. IEEE Sens J 15(12):6783–6790. https://doi.org/10.1109/jsen.2015.2465935

Cao X, Wang RC, Huang HP (2012) Multi-path routing algorithm for video stream in wireless multimedia sensor networks. J Softw 23(01):108–121. https://doi.org/10.3724/SP.J.1001.2012.04070

Cao JW, Lai XP, Chen T, Fan JY (2016) Accurate and efficient scene recognition with compact BoW and ensemble ELM. In: 2016 12th World Congress on Intelligent Control and Automation, pp 2058–2062. https://doi.org/10.1109/WCICA.2016.7578545

Chai Y, Li H, Li Z (2011) Multifocus image fusion scheme using focused region detection and multiresolution. Opt Commun 284(19):4376–4389. https://doi.org/10.1016/j.optcom.2011.05.046

Chen Y, Xiong J, Liu HL, Fan Q (2014) Fusion method of infrared and visible images based on neighborhood characteristic and regionalization in NSCT domain. Optik Int J Light Electron Opt 125(17):4980–4984. https://doi.org/10.1016/j.ijleo.2014.04.006

Deng J (1982) Control problems of grey systems. Syst Control Lett 01(05):9–18

Ellmauthaler A, Pagliari CL, Silva EABD (2013) Multiscale image fusion using the undecimated wavelet transform with spectral factorization and nonorthogonal filter banks. IEEE Trans Images Process 22(03):1005–1017

Fan JY, Chen T, Cao JW (2015) Image tampering detection using noise histogram features. In: 2015 IEEE International Conference on digital signal processing, pp 1044–1048. https://doi.org/10.1109/ICDSP.2015.7252037

Gao GR, Xu LP, Feng DZ (2013) Multi-focus image fusion based on nonsubsampled shearlet transform. IET Image Process 7(06):633–639. https://doi.org/10.1049/iet-ipr.2012.0558

Goyal S, Grover S (2012) Applying fuzzy grey relational analysis for ranking the advanced manufacturing systems. Grey Syst Theory Appl 2(02):284–298. https://doi.org/10.1108/20439371211260243

Guo M, Fu Z, Xi XL (2012) Novel fusion algorithm for infrared and visible images based on local energy in NSCT domain. Infrared Laser Eng 41(08):2229–2235

He GQ, Hao CY, Wang Y (2007) New and better image fusion method based on grey relational analysis and IHS transform. Appl Res Comput 24(07):312–314

Heshmati A, Gholami M, Rashno A (2016) Scheme for unsupervised colourCtexture image segmentation using neutrosophic set and nonsubsampled contourlet transform. IET Image Process 10(06):464–473. https://doi.org/10.1049/iet-ipr.2015.0738

Iovane G, Giordano P, Borysenko SD (2011) Image watermarking via wavelet approach and face biometrics. J Ambient Intell Hum Comput 2(02):91–101. https://doi.org/10.1007/s12652-010-0031-1

Jiang D (2010) Image fusion algorithm and appliction research based on multi-scale transform. Huan University Library, Changsha, pp 25–26

Kakerda RK, Kumar M, Mathur G (2015) Fuzzy type image fusion using hybrid DCTFFT based laplacian pyramid transform. In: International Conference on Communications and Signal Processing, pp 1049–1052. https://doi.org/10.1109/ICCSP.2015.7322661

Kong W, Lei Y, Zhao R (2015) Fusion technique for multi-focus images based on NSCT-CISCM. Optik Int J Light Electron Opt 126(21):3185–3192. https://doi.org/10.1016/j.ijleo.2015.07.142

Li H, Wei S, Chai Y (2012) Multifocus image fusion scheme based on feature contrast in the lifting stationary wavelet domain. EURASIP J Adv Signal Process 01:1–16

Li S, Yang B (2010) Hybrid multiresolution method for multisensor multimodal image fusion. IEEE Sens J 10(09):1519–1526. https://doi.org/10.1109/jsen.2010.2041924

Liu WF, He X (2011) A novel grey relational model. Stat Decis (14):160–161. https://doi.org/10.13546/j.cnki.tjyjc.2011.14.019

Liu SF, Dang YG, Fang ZG (2010) Grey system theory and application, vol 5. Science Press, Beijing, pp 95–101

Liu CP, Long YH, Mao JX (2016) Energy-efficient multi-focus image fusion based on neighbor distance and morphology. Optik Int J Light Electron Opt 127(23):11354–11363. https://doi.org/10.1016/j.ijleo.2016.09.038

Lu X, Lei C, Zeng H (2018) A multi-scale contrast-based image quality assessment model for multi-exposure image fusion. Signal Process 145:233–240. https://doi.org/10.1016/j.sigpro.2017.12.013

Ma M, Wan RY, Yin YL (2012) Multi-focus image fusion based on grey relation of similarity in Curvelet domain. Acta Electronica Sinica 40(10):1984–1988

Mitianoudis N, Stathaki T (2008) Optimal Contrast correction for ICA-based fusion of multimodal images. IEEE Sens J 8(12):2016–2026. https://doi.org/10.1109/jsen.2008.2007678

Nirmala DE, Vignesh RK, Vaidehi V (2013) Multimodal image fusion in visual sensor networks. In: 2013 IEEE International Conference on electronics, computing and communication technologies. https://doi.org/10.1109/CONECCT.2013.6469319

Qu XB, Yan JW, Xiao HZ (2008) Image fusion algorithm base on spatial frequency-motivated pulse coupled neural networks in non-subsampled contourlet transform domain. Acta Autom Sin 34(12):1508–1514

Qu GH, Zhang DL (2002) Information measure for performance of image fusion. Electron Lett 38(07):313–315

Shreyamsha BK, Swamy MNS, Ahmad MO (2013) Multiresolution DCT decomposition for multi focus image fusion. In: 2013 26th Annual IEEE Canadian Conference on Electrical and Computer Engineering. https://doi.org/10.1109/CCECE.2013.6567721

Stathaki T (2008) Image fusion: algorithm and applications, vol 3. Academic Press, New York, pp 367–392

Wang J, Li Q, Jia Z, Kasabov N, Yang J (2015) A novel multi-focus image fusion method using PCNN in nonsubsampled contourlet transform domain. Optik Int J Light Electron Opt 126(20):2508–2511. https://doi.org/10.1016/j.ijleo.2015.06.019

Wang YN, Xiao L, Ling ZG, Yang YM (2015) A method to calibrate vechicle-mounted cameras under urban traffic scenes. IEEE Trans Intell Trans Syst 16(6):3270–3279

Xie X, Liu Q, Hu FP (2015) A study on fast SIFT image mosaic algorithm based on compressed sensing and wavelet transform. J Ambient Intell Hum Comput 6(6):835–843. https://doi.org/10.1007/s12652-015-0319-2

Xydeas CS, Petrovic V (2000) Objective image fusion performance measure. Electron Lett 36(04):308–309

Yang Y, Tong S, Huang S, Lin P (2014) Log-Gabor energy based multimodal medical image fusion in NSCT domain. Comput Math Methods Med 02:835481. https://doi.org/10.1155/2014/835481

Yang Y, Tong S, Huang S, Lin P (2014) Multi-focus image fusion based on NSCT and focused area detection. IEEE Sens J 15(05):1–1. https://doi.org/10.1109/jsen.2014.2380153

Yang YM, Jonathan QM, Wang W (2016) Autoencoder with invertible functions for dimension reduction and image reconstruction. IEEE Trans Syst Man Cybern Syst. https://doi.org/10.1109/TSMC.2016.2637279

Ye CQ, Miao QG, Wang BS (2008) An image fusion algorithm using region segmentation and contourlet transform. Acta Optica Sinica 28(03):447–453

Yin M, Liu W, Zhao X, Yin Y, Guo Y (2014) A novel image fusion algorithm based on nonsubsampled shearlet transform. Optik Int J Light Electron Opt 125(10):2274–2282. https://doi.org/10.1016/j.ijleo.2013.10.064

Yu ZH, Zhang ZM, Chen R (2015) A method for fusion of infrared and visible images based on contourlet transform. Radio Eng 45(08):30–34

Zhao H, Zhao XM, Zhang TQ, Liu Y (2017) A new contourlet transform with adaptive directional partitioning. IEEE Signal Process Lett 24(06):843–847. https://doi.org/10.1109/lsp.2017.2696886