

Image fusion based on complex-shearlet domain with guided filtering

Shuaiqi Liu^{1,2} · Mingzhu Shi³ · Zhihui Zhu⁴ · Jie Zhao^{1,2}

Received: 17 January 2015 / Revised: 23 June 2015 / Accepted: 28 June 2015 /
Published online: 5 July 2015
© Springer Science+Business Media New York 2015

Abstract Combined the advantages of time-frequency separation of complex shearlet (CST) with the feature of guided filtering, a new image fusion algorithm based on CST domain and guided filtering is proposed. Firstly, CST is utilized for decomposition of the source images. Secondly, two scale guided filtering fusion rule is applied to the low frequency coefficients. Thirdly, larger sum-modified-Laplacian with guided filtering fusion rule is applied to the high frequency coefficients. Finally, the fused image is gained by the inverse CST. The algorithm can not only preserve the information of the source images well, but also improve the spatial continuity of fusion image. Experimental results show that the proposed method is superior to other current popular ones both in subjective visual and objective performance.

Keywords Image fusion · CST · Guided filtering · SML

This work was supported in part by Natural Science Foundation of China under Grant No. 61401308, Natural Science Foundation of Hebei Province under Grant No. 2013210094, Natural Science Foundation of Hebei University under Grant No. 2014-303, Science and technology support project of Baoding City under Grant No. 15ZG016, Open laboratory project of Hebei University under Grant No. sy2015009 and sy2015057.

✉ Shuaiqi Liu
shdkj-1918@163.com

Mingzhu Shi
shimingzhu1@163.com

Zhihui Zhu
zzhu@mines.edu

Jie Zhao
jzhao_hbu@163.com

- ¹ College of Electronic and Information Engineering, Hebei University, Baoding 071000, China
- ² Key Laboratory of Digital Medical Engineering of Hebei Province, Baoding 071000, China
- ³ College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
- ⁴ Department of Electrical Engineering and Computer Science, Colorado School of Mines, Golden, CO 80401, USA

1 Introduction

Generally, image fusion can be divided into three levels in ascending order: pixel level fusion, feature level fusion and decision level fusion (Pajares and Cruz 2004). In this paper, we focus on pixel-level fusion technique. Image fusion methods based on pixel-level can be divided into two categories: spatial domain algorithms and transform domain algorithms. The spatial domain algorithms mainly include weighted average (Liu et al. 2015), principal component analysis (Jia 1998) and so on. The transform domain algorithms are mainly based on the wavelet domain and multi-scale geometric transform domain, such as image fusion algorithm based on wavelet domain proposed in Pajares and Cruz (2004), image fusion algorithm based on contourlet transform (CT) proposed in Zhang and Guo (2009), Liu et al. (2011), Qu et al. (2009), and image fusion algorithm based on non-subsampled contourlet transform (NSCT) and pulse couple neural network (PCNN) proposed in Qu et al. (2008). In order to obtain a good image fusion effect, fusion algorithm needs to satisfy three characteristics. Firstly, fused image requires retaining the most information of the source images. Secondly, artificial textures should not be introduced to fused image, which means that fused image should have spatial smoothness. Finally, fusion algorithm should have robustness so that it can be applied to different source images. Currently, the transform domain algorithm cannot only capture details of the source images, but also suit different source images, so it is the mainstream of current fusion algorithm.

The image fusion algorithm based on wavelet transform plays an important role in traditional transform domain image fusion algorithms. But the traditional discrete wavelet transform (DWT) cannot represent image optimally and do not hold the feature of translation invariant. In order to better represent two-dimensional image which contains line or surface singularities, many scholars have put forward a number of methods. CT proposed in Do and Vetterli (2005) is one of the most influential transforms. Because CT is computationally simple and can represent images sparsely, it has been widely applied to image fusion algorithms, such as in Zhang and Guo (2009), Liu et al. (2011), and Qu et al. (2009). However, using CT for image decomposition will generate frequency aliasing and translational variability (Eslami and Radha 2004). The wavelet-contourlet proposed in Eslami and Radha (2004) has achieved good effect in the field of image processing by using DWT overcome the effect of frequency aliasing in scale decomposition. However, due to the downsampling in the processing of scale decomposition and directional decomposition, wavelet-contourlet transform is not translation invariant. Then in Cunha et al. (2006), the authors put forward a method to construct NSCT by utilizing the non-subsampled Laplacian transform and non-subsampled directional filter. Although NCST has translation invariance and can overcome the pseudo Gibbs phenomena, the redundancy of transform (which increases rapidly as scale increases) leads to complex computation. NSCT does not comply with multi-resolution theory and cannot conduce to mathematical theory analysis.

Recently, Guo and Labate constructed shearlet transform (ST) by an affine system with composite dilations, which can sparsely represent an image and produce optimal approximation (Easley et al. 2008; Kutyniok et al. 2011; Lim 2010). Compared to CT and NSCT, ST fits tight frame theory, has strict mathematical derivation, and its sampled directional filter does not produce a pseudo-Gibbs phenomenon caused by sampling (Easley et al. 2008). In addition, its discrete form is very easy to implement and the computational complexity is greatly reduced. Therefore ST has also been widely applied to image fusion. For example, the adaptive fusion algorithm based on ST was proposed in Miao et al. (2011), a region saliency fusion algorithm based on ST domain was proposed in Miao et al. (2011), the image fusion algorithm based on ST and PCNN was proposed in Geng et al. (2012). Among them,

the algorithm was proposed in Geng et al. (2012) has achieved good results. However, these algorithms can neither overcome the shortcoming of translation invariance of ST nor solve the second problem of image fusion; in other words it fails to make effective use of the spatial continuity of image.

In order to overcome the two shortcomings mentioned above, we adopt the CST proposed in Liu et al. (2013, 2014) to do transform domain analyzing, and utilize the idea of fusion algorithm based on two scale guided filtering proposed in Li et al. (2013) to enhance spatial continuity of fusion image based on transform domain. CST is constructed by dual-tree complex wavelet Kingsbury (1999) and shear directional filter. Compared to the non-subsampled shearlet transform (NSST), it realize shift invariance with limited redundancy, which results in accelerating computational speed and enhancing the number of directional decomposition of high frequency coefficients to get more sparse representation of the image than other multi-scale geometric transform (Liu et al. 2013, 2014). So CST is used to decompose images in our method. After getting the fused maps of high-frequency coefficients, unlike other transform domain fusion algorithms, we take the following steps instead of directly taking it through the weighted sum to obtain fused image. Firstly, guided filtering which is edge preserving filtering is imported to increase the spatial continuity of fused maps of high-frequency coefficients. Secondly, two scales guided filtering fusion algorithm proposed in Li et al. (2013) is used to fuse low frequency coefficients. Finally fused image is obtained by inverse CST. In this paper, the properties of shift invariant of CST are utilized to overcome pseudo-Gibbs phenomenon of the fused image. Through the guided filtering, the spatial smoothness of fused image is enhanced too. As both of CST and guided filtering having low computational complexity, the whole algorithm has low computational complexity. Compared the proposed algorithm to the fusion algorithms proposed in Que et al. (2009, 2008), Miao et al. (2011), Geng et al. (2012) and Li et al. (2013), experimental results demonstrate that visual effect has been significantly improved, and the objective evaluation criteria has been improved as well.

This paper is organized as follows. In Sect. 2, the construction of CST is reviewed. In Sect. 3, the working principle of the guided filtering in keeping the edge is introduced. In Sect. 4 the proposed image fusion algorithm is described. In Sect. 5, we demonstrate the feasibility of the proposed algorithm through experimental analysis. In the last section, we summarize the advantages and disadvantages of the algorithm.

2 Complex shearlet transform

2.1 The construction of CST

For any $(\xi_1, \xi_2) \in \hat{R}^2, j \geq 0, l = -2^j, 2^j - 1$ and $k \in \hat{R}^2, d = 0, 1$, the Fourier transform of the ST can be expressed as:

$$\hat{\psi}_{j,l,k}^{(d)}(\xi) = 2^{\frac{3j}{2}} V(2^{-2j}\xi) W_{j,l}^{(d)}(\xi) e^{-2\pi i \xi A^{-j} B^{-l} k} \tag{1}$$

where $V(2^{-2j}\xi)$ is the Fourier coefficient of the multi-scale analysis, $W_{j,l}^{(d)}, d = 0, 1$ is a window function which denotes multi-direction decomposition. So the ST of $f \in L^2(R^2)$ can be computed by:

$$\langle f, \psi_{j,l,k}^{(d)} \rangle = 2^{\frac{3j}{2}} \int \hat{f}(\xi) \overline{V(2^{-2j}\xi) W_{j,l}^{(d)}} e^{-2\pi i \xi A^{-j} B^{-l} k} d\xi \tag{2}$$

A new image representation method named CST based on the summarization of the advantages of dual tree complex wavelet transform (DTCWT) and ST is proposed in (Liu et al. 2013, 2014). CST is implemented as follows. Given an image $f \in L^2(R^2)$, let $\hat{f}[k_1, k_2]$ denote its 2D discrete Fourier transform coefficients. Brackets $[\bullet, \bullet]$ denote arrays of indices, and parentheses (\bullet, \bullet) denote function evaluations. Instead of Laplace transform, DTCWT is used to compute $\hat{f}(\xi_1, \xi_2) V(2^{-2j}\xi_1, 2^{-2j}\xi_2)$ at j -th scale. Then, we can decompose the father sub-band coefficient $f_a^{j-1}[n_1, n_2]$ into a low-pass sub-band coefficients $f_a^j[n_1, n_2]$ and six high-pass sub-band coefficients $f_{d(\theta)}^j$ ($\theta = 1 \sim 6$ denotes different directions, $N_a^j = 2^{-j+1}N$ and $N_d^j = 2^{-j}N$ are the size of $f_a^j[n_1, n_2]$ and $f_{d(\theta)}^j[n_1, n_2]$ at j -th scale respectively).

Let $\hat{\delta}_p$ represent the discrete Fourier transform of the delta function in the pseudo-polar grid, and φ_p denote the mapping function from the Cartesian grid to the pseudo-polar grid. Then, we can calculate the CST coefficients $\hat{f}_{d(\theta)}^j[n_1, n_2]\hat{w}_{j,l}^s[n_1, n_2]$ in the Cartesian grid (Liu et al. 2014), where

$$\hat{w}_{j,l}^s[n_1, n_2] = \varphi_p^{-1} \left(\hat{\delta}_p [n_1, n_2] \tilde{W} \left[2^j n_2 - l \right] \right) \tag{3}$$

where \tilde{W} is a frequency-based Meyer window function. After calculating $\hat{f}_{d(\theta)}^j[n_1, n_2]\hat{w}_{j,l}^s[n_1, n_2]$, the inverse Fourier transform is used to obtain the coefficients of CST. Original image can be reconstructed perfectly through simple sum (Liu et al. 2013, 2014). CST is described as follows (Liu et al. 2014):

$$DST_\psi(f) = f_a^N + \sum_{j=0}^N \sum_{l=-2^j}^{2^j-1} f_d^j * w_{j,l}^s \tag{4}$$

Because DTCWT is translational invariant, CST is also translational invariant through translational invariant shear directional filter (Liu et al. 2014). In addition, DTCWT can produce six high-pass sub-bands. After directional filter is applied to each high-pass sub-band, the coefficients of CST are sparser and more conducive for image fusion.

2.2 The performance of CST in image fusion

To verify the effect of CST, ST in fused algorithm based on regional variance in ST domain (ST-RV) in Miao et al. (2011) is replaced by CST (called CST-RV) to fuse common multi-focus fusion test images. Because the fusion rules are the same in ST-RV and CST-RV, the fusion effect just depends on the contribution of CST and ST. In the experiment, all the decomposition scale is 4, and the directions in every scale are [6 10 10 8]. Figure 1a, b are classical multi-focus fusion images. The size of both images is 512×512 . The fused images and difference images (fused images minus the source image) are shown in Fig. 1c–h.

The difference images by fused image minus sources images can highly indicate how many information of sources images are got by fused algorithm. In general, the more information it keeps, the better the algorithm is. The blur edge which is marked by rectangle in Fig. 1d, e is caused by the pseudo Gibbs effect. This obviously reduces the visual effect of image fusion. For clearly showing the performance of CST, the region marked by rectangle is amplified in Fig. 2.

Comparing Fig. 2a, b, we can find that CST-RV can retain more textures in the fused image. Figure 2c, d show that CST-RV can effectively suppress artificial texture in the fused image. What is more, CST is computationally efficient enough to work as a tool in image fusion. The computational complexity of operation of CST is about twice of the ST (the time

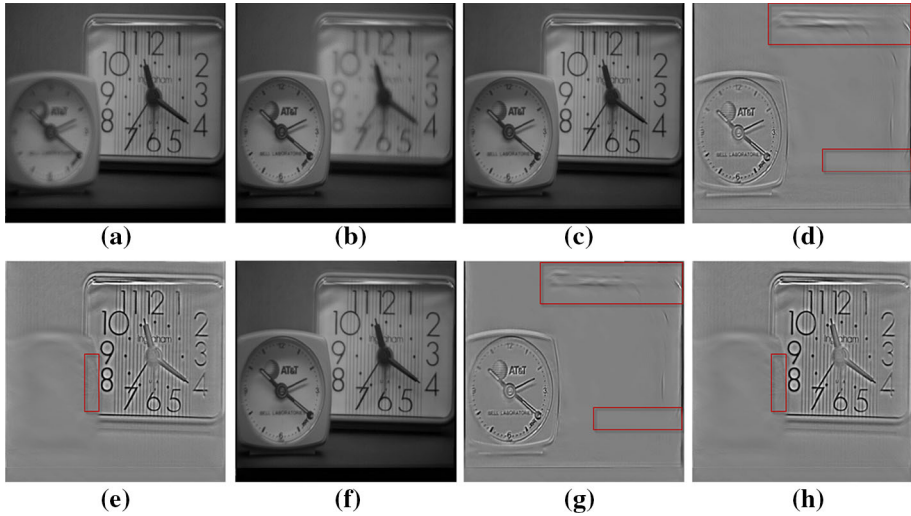


Fig. 1 The performance of CST. **a** Right focus image, **b** left focus image, **c** fused by ST-RV, **d** difference image of (c) minus (a), **e** difference image of (c) minus (b), **f** fused by CST-RV, **g** difference image of (f) minus (a), **h** difference image of (f) minus (b)

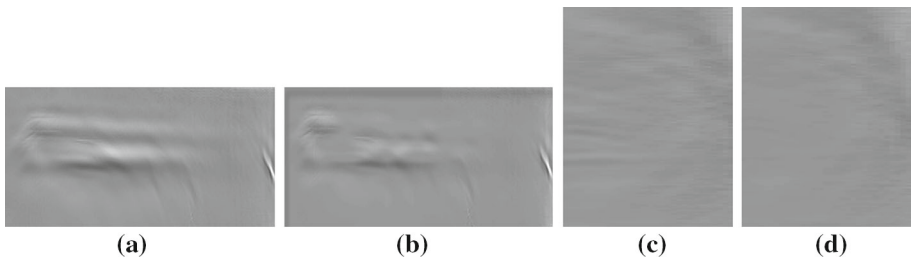


Fig. 2 Enlarged images of rectangles marked in Fig. 1. **a** Enlarged image of above rectangle in Fig. 1d, **b** enlarged image of above rectangle in Fig. 1g, **c** enlarged image of rectangle in Fig. 1e, **d** enlarged image of rectangle in Fig. 1h

complexity of CST and ST are all $O(N \log N)$; Liu et al. 2014). But it is much lower than NSST (Liu et al. 2014), and is the one- K th of ST with cycle spinning. K denotes the number of spinning.

3 Guided filtering

Recently, the research of edge preserving guided filtering has become a hotspot in image processing (He et al. 2013; Farbman et al. 2008). The guided filtering (He et al. 2013), the latest weighted least squares filtering (Farbman et al. 2008) and other edge-preserving smoothing filtering can do image filtering without blurring strong edges and hence they avoid introducing artificial texture in the process of filtering. The guided filtering is a linear edge-preserving filtering algorithm and its computation time does not depend on the size of filtering kernel. Therefore it can be well applied in the field of image processing. Compared to other edge preserving filtering algorithms, the guided filtering is very fast (its time complexity is O

(n) [He et al. 2013](#)), and it is easily applied to image fusion (details see [Li et al. 2013](#)). Hence guide filtering is selected to enhance the spatial continuity in our image fusion algorithm.

3.1 The construction of guided filtering

The process of using the guided filtering involves a guidance image I , an input image p which needs filtering, and an output image q . Image I and p can be identical. The key assumption of the guided filtering is a local linear model between the guidance I and the filtering output q :

$$q_i = a_k I_i + b_k, \quad \forall i \in \omega_k \tag{5}$$

where i is pixel index, a_k, b_k are some linear coefficients, and ω_k is a local square window of radius $(2r + 1)$ centered at the pixel k in the guidance image I .

Then the image filtering problem of edge preserving is converted into the optimal problem of minimizing the difference between q and p and satisfying the linear condition in (5) at the same time. That is:

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2) \tag{6}$$

Here ε denotes a regularization parameter (fuzzy parameter). The solution to formula (6) can be given by linear regression ([Draper and Smith 1981](#)).

$$a_k = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon} \tag{7}$$

$$b_k = \bar{p}_k - a_k \mu_k \tag{8}$$

Here, μ_k and σ_k^2 denote the mean and variance of I in local window ω_k , $|\omega|$ is the number of pixels in ω_k , and \bar{p}_k is the mean of p in ω_k .

In order to keep the value of q_i invariant when it is computed in different windows, after working out a_k and b_k for all patches ω_k in the image, we use average filter to average all the possible values in local window. That is:

$$q_i = \frac{1}{|\omega|} \sum_{k:i \in \omega_k} (a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i \tag{9}$$

where \bar{a}_i and \bar{b}_i denote the mean of a_k and b_k in local window ω_k , namely

$$\bar{a}_i = \frac{1}{|\omega|} \sum_{k \in \omega_i} a_k, \quad \bar{b}_i = \frac{1}{|\omega|} \sum_{k \in \omega_i} b_k \tag{10}$$

For simplicity, this paper use $G_{r,\varepsilon}(p, I)$ to denote guided filtering. Here, r denotes the size of filter kernel, p denotes the input image, and I denotes the guidance image.

3.2 The performance of guided filtering in image fusion

In multi-focus image fusion, spatial distorting is most likely to occur in the junction of clear part and the blur part of the source images. To describe the action of guided filtering clearly, we consider an extreme case that a clear image is split to two half images such as Fig. 3a, b. In these images, one half has a very clear image and the other half image is very blur (the values of these pixels are all the same to 0). We can observe that guided filtering can heavily improve the spatial continuity by surveying the difference images. The fused method

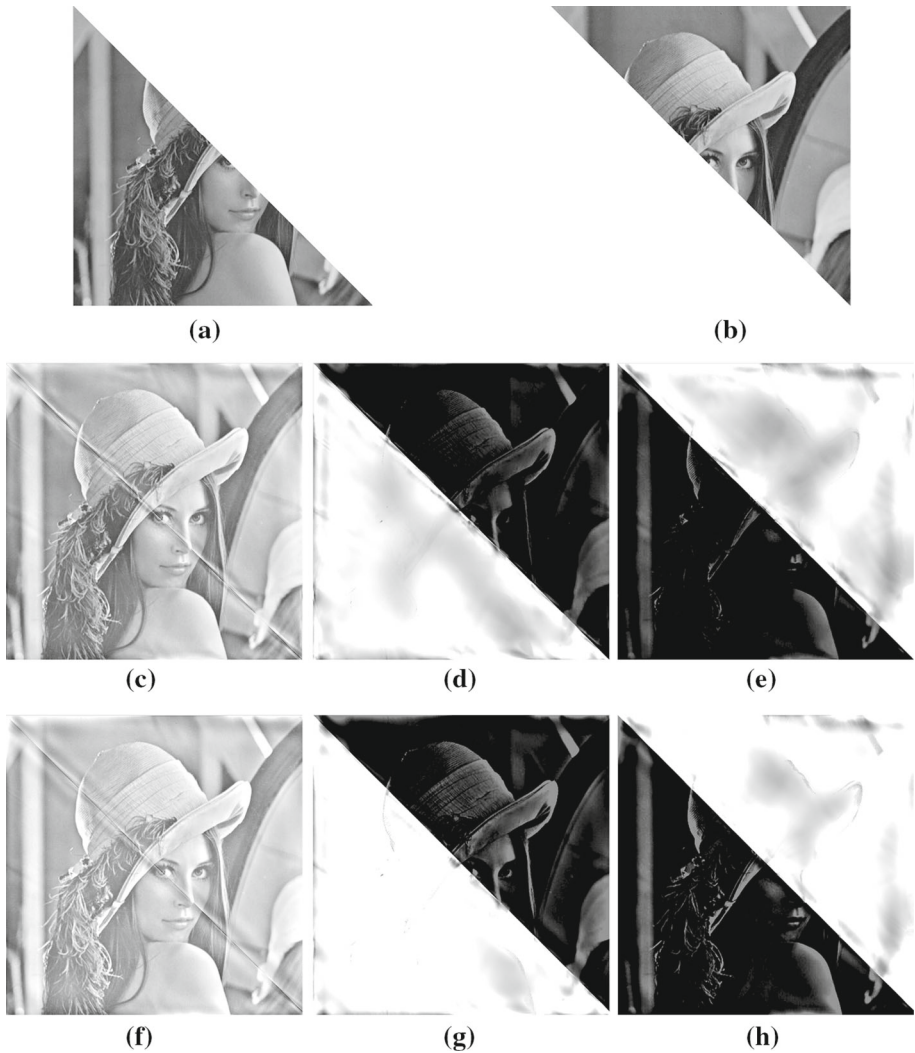


Fig. 3 The performance of guided filtering. **a** *Left* original image, **b** *right* original image, **c** fused image using CST-SML, **d** difference image of (c) minus (a), **e** difference image of (c) minus (b), **f** fused image using GF-CST-SML, **g** difference image of (f) minus (a), **h** difference image of (f) minus (b)

is our method called CST-SML-GF and CST-SML (remove the step of guided filtering in our method), and the decomposition parameters of CST are the same as before. Figure 3a is the left half image, while Fig. 3b is the other half image. Fused images and difference images (fused images minus the source image) are shown in Fig. 3c–h.

Comparison of Fig. 3d, g shows that CST-SML-GF has better spatial continuity. The fused image by CST-SML-GF has little texture and shadow region in the bottom near the diagonal and diagonal left relatively smooth. This performance fully shows that guided filtering can be used to enhance the spatial continuity and suppress artificial texture of fused image. From Fig. 3e, h, we can get the same conclusion.

4 Image fusion algorithm based on CST domain with guided filtering

CST which is translational invariant is utilized to decompose images. It can suppress the pseudo Gibbs effect well without increasing computational complexity. In image fusion, using region energy fused rules to select coefficients at clear parts has good fusion results. Because the human visual system is sensitive to the image edge, direction and texture information instead of a single pixel, the fused rules based on regional energy can satisfy the visual system well. SML proposed in [Qu et al. \(2009\)](#) is a better regional energy function. Compared to other measurements such as, energy of gradient (EOG), spatial frequency (SF), tenengrad, energy of Laplace (EOL), SML can represent the image focusing degree effectively, and SML also has a good effect in selecting the high-frequency coefficients. Combined SML with the context in Sects. 2 and 3, a new fusion method is proposed. Without losing of generality, we suppose that A and B are two images with different focuses to be fused, F is the resulting fused image. Obviously it is easy to extend to multiple image fusion. The proposed image fusion algorithm is described as follows.

Firstly, CST is used to decompose image A and B , and decomposition coefficients are $S_A^{l,d}(k)$ and $S_B^{l,d}(k)$. Generally, the coefficients of CST are $S^{l,d}(k)$, where l and d denote the scale and direction of decomposition respectively. When l is zero, $S^{l,d}(k)$ denote the low frequency coefficients, and otherwise they denote the high frequency coefficients. Let k denote the position of pixel, so SML at pixel k is defined as follows ([Qu et al. 2009](#)):

$$SML^{l,d}(k) = \sum_{i \in \omega_k} [ML^{l,d}(i)]^2 \tag{11}$$

where ω_k denotes a rectangular window centered at the pixel k , let $i = (x, y)$, then:

$$ML^{l,d}(x, y) = |2S^{l,d}(x, y) - S^{l,d}(x - step, y) - S^{l,d}(x + step, y)| + |2S^{l,d}(x, y) - S^{l,d}(x, y - step) - S^{l,d}(x, y + step)| \tag{12}$$

where $step$ denotes a variable spacing between pixels. Generally, $step$ equals to 1.

Secondly, for the low frequency coefficients $S_A^{0,d}(k)$ and $S_B^{0,d}(k)$, dual scale guided filtering fusion rule proposed in [Li et al. \(2013\)](#) is applied to fuse them. The proposed method in [Li et al. \(2013\)](#) utilizes the average filter to get the two-scale representations, which is simple and effective. More importantly, the guided filtering is used in a novel way to make full use of the strong correlations between neighborhood pixels for weight optimization. Approximate information of image is constructed by the low frequency coefficients, so applying the algorithm can get a better fusion effect and greatly increase the spatial continuity of image. We call this fusion rule GFF, and we refer to [Li et al. \(2013\)](#) for details on fusion algorithm. That is,

$$S_F^{0,d}(k) = GFF(S_A^{0,d}(k), S_B^{0,d}(k)) \tag{13}$$

Thirdly, the larger SML with guided filtering fusion rule is applied to high frequency coefficients $S_A^{l,d}(k)$ and $S_B^{l,d}(k)$ (l is larger than zero). Applying (11) to high frequency coefficients, we can get $SML_A^{l,d}(k)$ and $SML_B^{l,d}(k)$. Two matrixes $mapA$ and $mapB$ whose size is as same as the high frequency coefficients are initialized to zeros. The two matrixes are calculated by (14).

$$mapA(k) = \begin{cases} 0 & SML_A^{l,d}(k) > SML_B^{l,d}(k) \\ 1 & SML_A^{l,d}(k) \leq SML_B^{l,d}(k) \end{cases}, \quad mapB(k) = 1 - mapA(k) \tag{14}$$

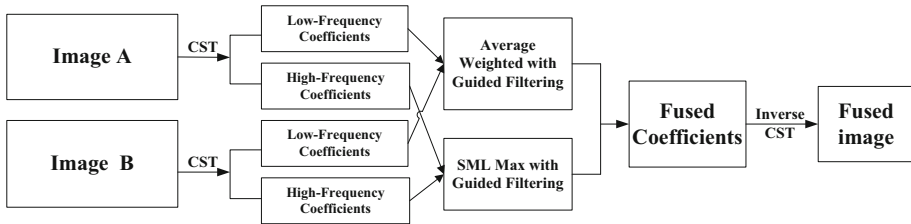


Fig. 4 Framework of image fusion algorithm based on CST domain with guided filtering

Certainly, in order to enhance the spatial continuity of high frequency coefficients, guided filtering is used to filter $mapA$ and $mapB$ by the guided image $S_A^{l,d}$ and $S_B^{l,d}$, as shown in (15).

$$mapA = G_{r,\varepsilon} \left(mapA, S_A^{l,d} \right), \quad mapB = G_{r,\varepsilon} \left(mapB, S_B^{l,d} \right) \tag{15}$$

Then $mapA$ and $mapB$ should be normalized. The fused high frequency coefficients $S_F^{l,d}(k)$ can be obtained by the following formula.

$$S_F^{l,d}(k) = mapA \times S_A^{l,d}(k) + mapB \times S_B^{l,d}(k) \tag{16}$$

Finally, the fused image F is obtained by fused low frequency coefficients $S_F^{0,d}(k)$ and the fused high frequency coefficients $S_F^{l,d}(k)$ through inverse CST (Since the CST coefficients of fused image constructed by the original CST coefficients of image A and image B are also complex number, we can get the fusion image by inverse CST). In conclusion, the framework of the proposed fusion algorithm is shown in Fig. 4.

5 Experimental results

5.1 Experiment setups

To verify the effect of our algorithm, we analyze the comparison of our algorithm with other fusion algorithms. The proposed fusion method has been evaluated on five pairs are shown in Fig. 5, which are publicly available online (<http://www.imagefusion.org>). The size of all images is 512×512 .

Image standard deviation (Std), average gradient (Avg), $Q^{AB/F}$ metric (Qu et al. 2009), mutual information (MI) (Li et al. 2013) and spatial frequency (SF) (Li et al. 2013) are employed as objective criteria. Std and Avg reflect the energy concentration of image, $Q^{AB/F}$ measures the amount of edge information transferred from the source images to fused images, MI essentially computes how much information that is transferred to fused image, and SF measures the overall activity of image spatial domain, which can reflects the ability of image expression on tiny details contrast. The larger five index values are, the clearer fused image we get, the better fusion performance a method has.

In the experiment, the decomposition level of CST is 4 as well, with 6, 10, 10, 18 directional, and the parameters of the high frequency fusion rule proposed in this paper are set as $r = 3, \varepsilon = 1$. The comparison fusion algorithms are as follows: the fusion method of choosing larger SML based on CT proposed in Qu et al. (2009) (CT-SML), the fusion method based on NSCT and PCNN proposed in Qu et al. (2008) (NSCT-PCNN), the adaptive fusion algorithm based on ST proposed in Miao et al. (2011) (ST-AF), the fusion method based

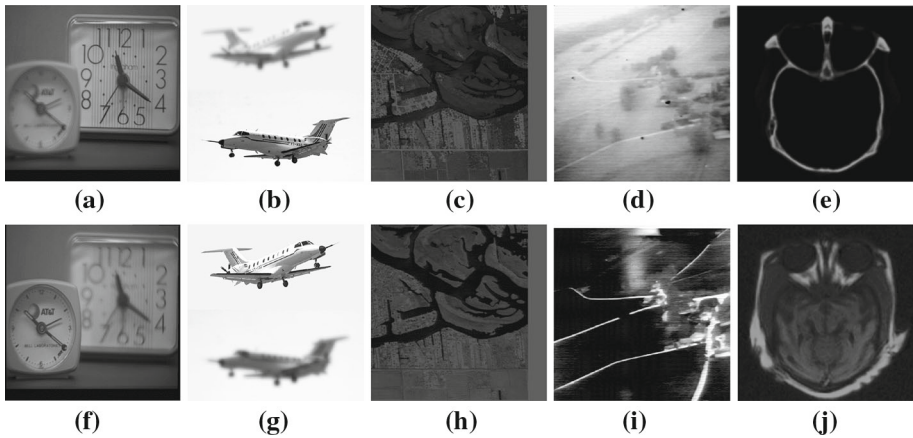


Fig. 5 Test images for fused algorithm. **a, f** are images of “Clock” with different focus. **b, g** are images of “Airplane” with different focus. **c, h** are of “Remote sensing” with different sensor. **d, i** are images of “Navigation” with different sensor. **e, j** are “Medical” of CT and MRI

on ST and PCNN proposed in [Geng et al. \(2012\)](#) (ST-PCNN), and the fusion method based on two scale guided filtering proposed in [Li et al. \(2013\)](#) (GFF). Default parameters in the shared source codes are used in these fused methods. To verify the effect of each part of our algorithm, fused method based on Guided filtering in ST domain (ST-GF) and fused method based on Guided filtering in DTCWT domain (DTCWT-GF) are also compared to our method (CST-SML-GF). In the experiment, the decomposition scale of DTCWT is 4.

5.2 Results on multi-focus images

Multi-focus image fusion is very important in image fusion. Figure 5a, f show a pair of images called “Clock”. The eight methods above are utilized to fuse this pair of images respectively. The fused images and difference images (fused images minus the source image) are shown in Fig. 6a–x.

Comparing the top left corner and the area around number 8 of the large clock in fused images, we see that the proposed method which uses image spatial continuity can avoid introducing some artificial texture into the fused image which exists in CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, ST-GF and DTCWT-GF. While the proposed method is poorer in these areas compared with the GFF, obviously the gray level is clearer, which can be seen from the difference images. This is mainly because we introduce the CST for scale decomposition to utilize its good time-frequency localization characteristic to improve the image gray layering and clarity. Overall, comparison of the difference image shows that the proposed algorithm has the best visual appearance, the least resulting artificial textures and it also has a significant suppression of artificial textures, which is mainly due to preserving the spatial continuity by using guided filtering.

Besides the subjective visual appearance, the five objective criteria mentioned above are used to investigate the performance of different transform methods. From Table 1, we can see that the proposed algorithm is the highest in all the criteria, which fully shows that the proposed algorithm does not only consider the information of two differently focused images, but also fully retain the spatial information of both. The highest $Q^{AB/F}$ and SF show that our method improves the ability of preserving image information and edge information. Because

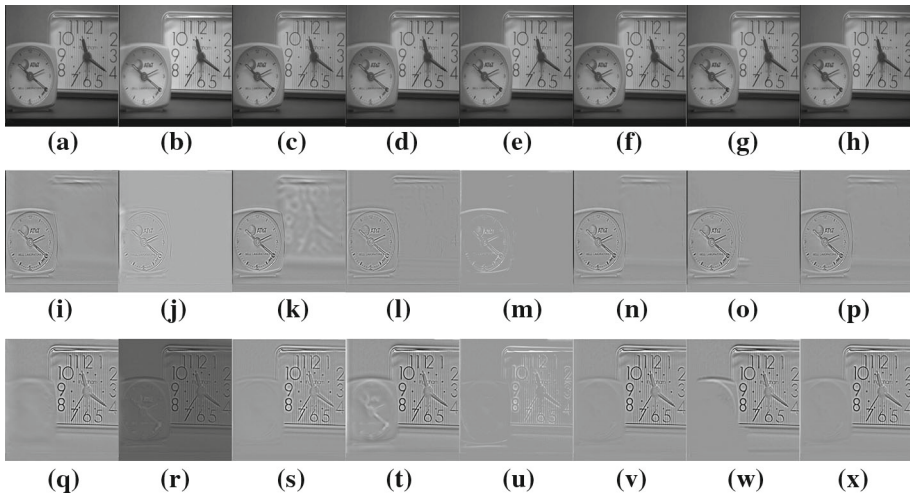


Fig. 6 Fusion results using “Clock” images. **a–h** are the fused images by using CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF, DTCWT-GF and CST-SML-GF. **(i, q), (j, r), (k, s), (l, t), (m, u), (n, v), (o, w)** and **(p, x)** are difference images (fused images minus Fig. 5a, f)

Table 1 Objective criteria by each fused method in the fusion of Fig. 6

The fusion methods	Std	Avg	$Q^{AB/F}$	MI	SF
CT-SML	41.1849	0.0112	0.6783	6.7086	8.0840
NSCT-PCNN	40.6143	0.0110	0.6775	7.4948	7.9091
ST-AF	40.8756	0.0112	0.6670	6.7204	8.0705
ST-PCNN	40.3870	0.0111	0.6900	6.7927	8.0202
GFF	40.4602	0.0111	0.7157	7.3364	8.0058
ST-GF	40.8594	0.0110	0.6902	7.5035	7.9173
DTCWT-GF	40.8514	0.0110	0.6861	7.3106	7.9205
CST-SML-GF	41.8595	0.0112	0.7275	7.8925	8.0842

The bold can show the best objective criteria

guided filtering in the transform domain is enhances spatial continuities which help the fused method can preserve more texture and spatial information of source images, the proposed algorithm has highest MI.

Next test images are Fig. 5b, g which are airplane images with different focus. The eight methods above are utilized to fuse this pair images respectively. The fused images and difference images are shown in Fig. 7a–x.

Comparison of the area around the head and tail of the two airplanes in fused images shows that the proposed method overcomes introducing some artificial texture into the fused image by using image spatial continuity which exists in CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF and DTCWT-GF. The difference images which the fused image obtained by each method minus Fig. 5b, g also show that our method is better than others. It shows that the proposed algorithm has the best visual appearance and the least resulting artificial textures, which means that the proposed algorithm has a significant suppression of artificial textures.

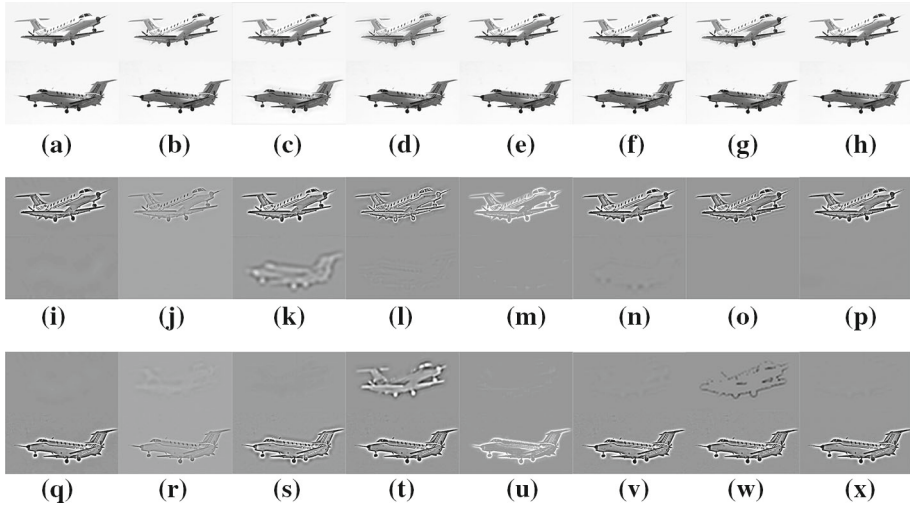


Fig. 7 Fusion results using “Airplane” images. **a–h** are the fused images by using CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF, DTCWT-GF and CST-SML-GF. **(i, q), (j, r), (k, s), (l, t), (m, u), (n, v), (o, w)** and **(p, x)** are difference images (fused images minus Fig. 5b and g)

Table 2 Objective criteria by each fused method in the fusion of Fig. 7

The fusion methods	Std	Avg	$Q^{AB/F}$	MI	SF
CT-SML	56.7416	0.0261	0.7939	5.3533	18.8566
NSCT-PCNN	56.3129	0.0261	0.7937	5.3227	18.7903
ST-AF	56.0713	0.0260	0.7872	4.7371	18.7271
ST-PCNN	54.8607	0.0256	0.7793	5.0406	18.4354
GFF	56.8756	0.0261	0.7975	5.4177	18.8529
ST-GF	56.8675	0.0261	0.7973	5.5714	18.8488
DTCWT-GF	56.7965	0.0259	0.7949	5.5685	18.6748
CST-SML-GF	56.8756	0.0261	0.7988	5.5735	18.8576

The bold can show the best objective criteria

The five objective criteria above are used to investigate the performance of different transform methods. As shown in Table 2, the proposed algorithm has the highest objective criteria.

5.3 Results on remote sensing and navigation images

Next, our method is tested on remote sensing and navigation images. Figure 5c, h and d, i are typical test images. CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF and DTCWT-GF are applied to the two pair images respectively. The fused images and difference images are shown in Figs. 8a–x and 9a–x.

Figure 8 shows that the proposed method suppresses artificial texture of fused image which exists in others methods. Similarly, difference images show that the proposed method is better than others.

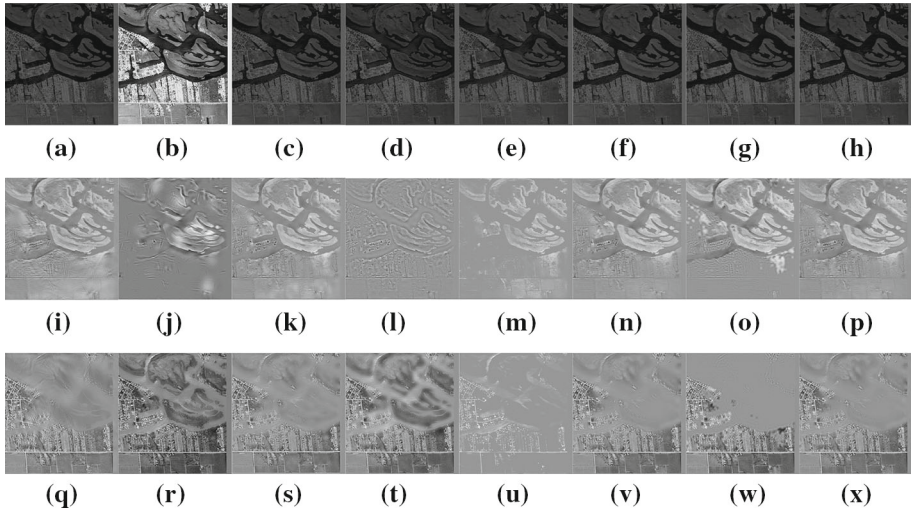


Fig. 8 Fusion results using “Remote sensing” images. **a–h** are the fused images by using CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF, DTCWT-GF and CST-SML-GF. **(i, q), (j, r), (k, s), (l, t), (m, u), (n, v), (o, w)** and **(p, x)** are difference images (fused images minus Fig. 5c, h)

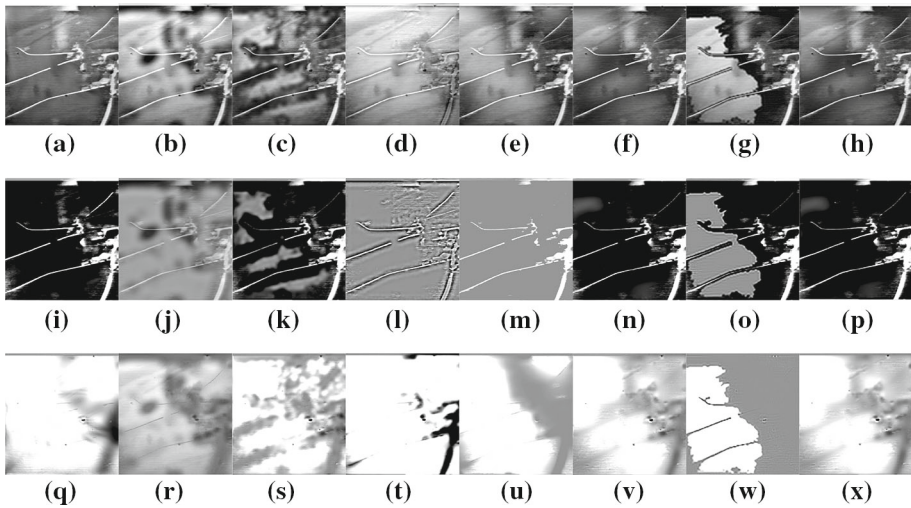


Fig. 9 Fusion results using “Navigation” images. **a–h** are the fused images by using CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF, DTCWT-GF and CST-SML-GF. **(i, q), (j, r), (k, s), (l, t), (m, u), (n, v), (o, w)** and **(p, x)** are difference images (fused images minus Fig. 5d, i)

Comparing the clarity of navigation and the area around the stones in the river in fused images in Fig. 9, we can see that the proposed algorithm and GFF have the best visual appearance. Similarly, the difference images show that the proposed method is better than others. It shows that the proposed algorithm is also suitable for fusing remote sensing and navigation images.

Table 3 Objective criteria by each fused method in the fusion of Fig. 8

The fusion methods	Std	Avg	$Q^{AB/F}$	MI	SF
CT-SML	22.3367	0.0143	0.5807	3.3400	10.3032
NSCT-PCNN	21.3312	0.0137	0.6380	4.2417	9.8478
ST-AF	21.7066	0.0146	0.6112	3.5579	10.6325
ST-PCNN	21.8665	0.0146	0.6005	3.6088	10.6906
GFF	20.7646	0.0134	0.6180	3.9427	9.6821
ST-GF	21.5162	0.0136	0.6174	3.6288	9.8271
DTCWT-GF	21.6604	0.0134	0.6157	3.9273	9.6781
CST-SML-GF	22.7151	0.0147	0.6638	4.6152	10.7150

Table 4 Objective criteria by each fused method in the fusion of Fig. 9

The fusion methods	Std	Avg	$Q^{AB/F}$	MI	SF
CT-SML	50.8261	0.0238	0.6342	1.8244	17.1938
NSCT-PCNN	50.3312	0.0237	0.6280	2.2417	16.8478
ST-AF	55.6532	0.0239	0.5599	2.2981	17.2400
ST-PCNN	53.1247	0.0213	0.5194	2.4512	15.3323
GFF	56.7465	0.0227	0.6613	2.0999	16.3669
ST-GF	54.9362	0.0238	0.6640	1.5554	17.1744
DTCWT-GF	56.5303	0.0238	0.6461	2.5408	17.1247
CST-SML-GF	56.8418	0.0239	0.6966	2.5420	17.2381

The bold can show the best objective criteria

From the objective criteria shown in Tables 3 and 4, we find the proposed algorithm has the highest objective criteria. So again it demonstrates that it is suitable for remote sensing and navigation images fusion.

5.4 Results on medical images

Finally, our method is tested on medical images. Figure 5e, j are medical images test sets. CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF and DTCWT-GF are applied to this pair images respectively. The fused images and difference images are shown in Fig. 10a–x.

Comparing the fused images of each algorithm, we can see the proposed fusion algorithm and ST-AF, ST-PCNN, GFF, DTCWT-GF have better visual appearance. These algorithms preserve the texture information of source images well. Similarly, comparing the difference images which the fused image obtained by each method minus Fig. 5e, j respectively, we can see although the proposed method is worse than GFF and DTCWT-GF, it is better than other algorithms. It shows that the proposed algorithm doesn't have the best fusion effect, but it can still be used in the application of the medical image fusion. From the objective criteria shown in Table 5, the proposed algorithm also has the highest objective criteria. Though our method's subjective visual performance is worse than GFF and DTCWT-GF, its objective performance is the best. For further research, our method can be applied to medical image.

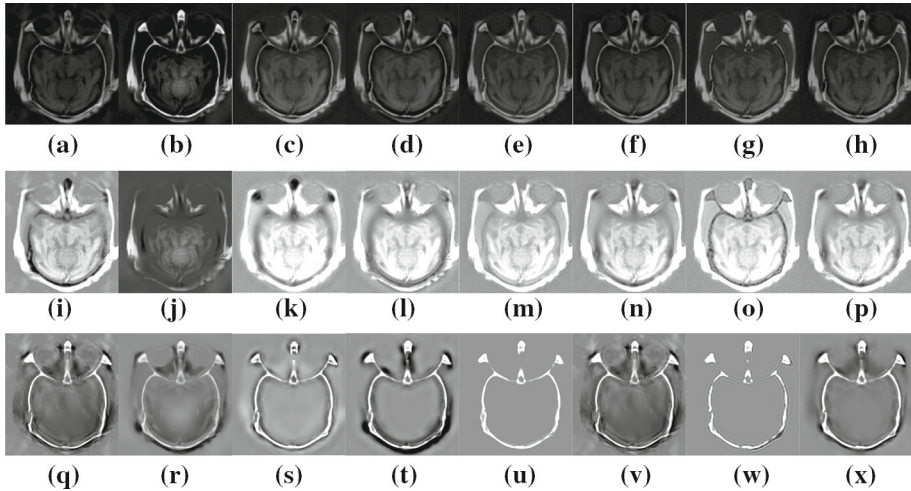


Fig. 10 Fusion results using “Medical” images. **a–h** are the fused images by using CT-SML, NSCT-PCNN, ST-AF, ST-PCNN, GFF, ST-GF, DTCWT-GF and CST-SML-GF. **(i, q), (j, r), (k, s), (l, t), (m, u), (n, v), (o, w)** and **(p, x)** are difference images (fused images minus Fig. 5e, j)

Table 5 Objective criteria by each fused method in the fusion of Fig. 10

The fusion methods	Std	Avg	$Q^{AB/F}$	MI	SF
CT-SML	27.4671	0.0255	0.5901	1.6195	9.1569
NSCT-PCNN	24.3072	0.0245	0.5254	1.1221	8.7962
ST-AF	29.8417	0.0269	0.7248	2.6951	9.7923
ST-PCNN	27.5633	0.0261	0.6761	2.2828	9.5430
GFF	29.9454	0.0250	0.7433	3.2586	8.9935
ST-GF	29.2278	0.0260	0.7414	2.3933	9.3573
DTCWT-GF	29.2706	0.0249	0.7152	3.4216	8.9500
CST-SML-GF	29.9976	0.0269	0.7799	3.4216	9.7923

The bold can show the best objective criteria

In terms of the visual appearance and objective criteria of the above five categories of fused image, although the proposed algorithm’s visual appearance is slightly poorer than GFF in medical image fusion, but in other types of image fusion it occupies great advantages. The proposed fusion algorithm can not only preserve the detail information of source image effectively, but also suppress the artificial textures in fused images, and most importantly, it is robust for different types of image fusion. In general, the proposed algorithm is a fairly good and worth extending algorithm for image fusion.

6 Conclusions

This paper presents an image fusion algorithm based on CST with guided filtering. The new algorithm uses the guided filtering to enhance the spatial characteristics of fused image based

on CST domain. It not only makes full use of the excellent time-frequency characteristics of CST, but also uses the energy function in frequency domain to smooth image so that the spatial continuities of fused image are greatly enhanced. Experimental results demonstrate that the proposed method is better than the current popular image fusion algorithms mentioned above in terms of both visual appearance and objective criteria. Moreover, the algorithm is robust for different types of image fusion. The next research direction will be to choose a more adaptive parameter to get better results because the parameters of the guided filtering have great influence on fusion effect.

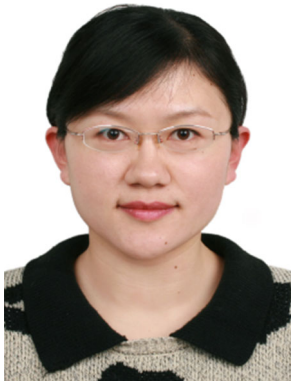
References

- Cunha, A. L., Zhou, J. P., & Do, M. N. (2006). The nonsubsampling contourlet transform: Theory, design and application. *IEEE Transactions on Image Processing*, *15*(10), 3089–3101.
- Do, M. N., & Vetterli, M. (2005). The contourlet transform: An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, *14*(12), 2091–2106.
- Draper, N., & Smith, H. (1981). *Applied regression analysis*. New York: Wiley.
- Easley, G., Labate, D., & Lim, W. Q. (2008). Sparse directional image representation using the discrete shearlets transform. *Applied and Computational Harmonic Analysis*, *25*(1), 25–46.
- Eslami, R., & Radha, H. (2004). Wavelet based contourlet transform and its application to image coding. In *IEEE international conference on image processing*, Singapore (pp. 3189–3192).
- Farbman, Z., Fattal, R., Lischinski, D., & Szeliski, R. (2008). Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics*, *27*(3), 67:1–67:10.
- Geng, P., Wang, Z., Zhang, Z., et al. (2012). Image fusion by pulse couple neural network with shearlet. *Optical Engineering*, *51*(6), 067005-1–067005-7.
- He, K. M., Sun, J., & Tang, X. O. (2013). Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(6), 1397–1409.
- Jia, Y. H. (1998). Fusion of landsat TM and SAR images based on principal component analysis. *Remote Sensing Technology and Application*, *13*(1), 46–49.
- Kingsbury, N. (1999). Image processing with complex wavelets. *Philosophical Transactions: Mathematical Physical and Engineering Sciences*, *357*(1760), 2543–2560.
- Kutyniok, G., Lemvig, J., & Lim, W. Q. (2011). Compactly supported shearlets are optimally sparse. *Journal of Approximation Theory*, *163*(11), 1564–1589.
- Li, S. T., Kang, X. D., & Hu, J. W. (2013). Image fusion with guided filtering. *IEEE Transactions on Image Processing*, *22*(7), 2864–2875.
- Lim, W. Q. (2010). The discrete shearlets transform: A new directional transform and compactly supported shearlets frames. *IEEE Transactions on Image Processing*, *19*(5), 1166–1180.
- Liu, K., Guo, L., & Chen, J. S. (2011). Contourlet transform for image fusion using cycle spinning. *Journal of Systems Engineering and Electronics*, *22*(2), 353–357.
- Liu, S. Q., Hu, S. H., & Xiao, Y. (2013). SAR image de-noising based on complex shearlet transform domain gaussian mixture model. *Acta Aeronautica et Astronautica Sinica*, *34*(1), 173–180. (in Chinese).
- Liu, S. Q., Hu, S. H., & Xiao, Y. (2014). Image separation using wavelet-complex shearlet dictionary. *Journal of Systems Engineering and Electronics*, *25*(2), 314–321.
- Liu, S. Q., Hu, S. H., Xiao, Y., et al. (2014). Bayesian Shearlet shrinkage for SAR image de-noising via sparse representation. *Multidimensional Systems and Signal Processing*, *25*(4), 683–701.
- Liu, S., Zhao, J., & Shi, M. Z. (2015). Medical image fusion based on rolling guidance filter and spiking cortical model. *Computational and Mathematical Methods in Medicine*, *2015*, 1–9.
- Miao, Q. G., Shi, C., & Xu, P. F. (2011). A novel algorithm of image fusion using shearlets. *Optics Communications*, *284*(6), 1540–1547.
- Miao, Q. G., Shi, C., & Xu, P. F. (2011). Multi-focus image fusion algorithm based on shearlets. *Chinese Optics Letters*, *9*(4), 041001.1–041001.5.
- Pajares, G., & Cruz, J. M. (2004). A wavelet-based image fusion tutorial. *Pattern Recognition*, *37*(9), 1855–1872.
- Qu, X. B., Yan, J. W., Xiao, H. Z., et al. (2008). Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain. *Acta Automatica Sinica*, *34*(12), 1508–1514.

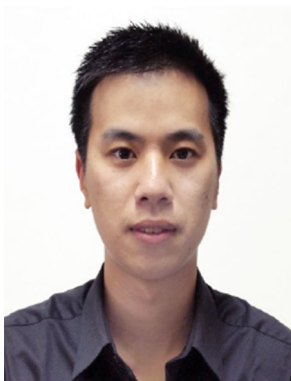
- Qu, X. B., Yan, J. W., & Yang, G. D. (2009). Sum-modified-Laplacian-based multi-focus image fusion method in sharp frequency localized contourlet transform domain. *Optics and Processing Engineering*, 17(5), 1203–1212.
- Zhang, Q., & Guo, B. (2009). Multifocus image fusion using the nonsubsampled contourlet transforms. *Signal Processing*, 89(7), 1334–1346.



Shuaiqi Liu received his Ph.D. degree in Institute of Information Science from Beijing Jiaotong University at 2014 and got B.S. degree in the department of Information and Computer Science from Shandong University of Science and Technology at 2009. At present, he is associate professor in college of Electronic and Information Engineering, Hebei University. His research interests include image processing and signal processing.



Mingzhu Shi received her Ph.D. degree in Optical Engineering from Beijing Institute of Technology at 2014 and got B.S. degree in the School of Optoelectronics Engineering from Xi'an Technological University at 2008. At present, she is a associate professor of Tianjin Normal University. Her research interests include signal processing, photoelectric imaging, image processing.



Zhihui Zhu received the B.Eng. degree in the department of communications engineering from Zhejiang University of Technology at 2012. At present, he is currently working towards the Ph.D. degree at the Colorado School of Mines, USA. His research interests include signal processing using sparse, low-dimensional signal models and low-rate signal sensing and acquisition.



Jie Zhao received Ph.D. degree in the department of Electronic Engineering from Hebei University of Technology at 2008. At present, he is professor in college of Electronic and Information Engineering, Hebei University. His research interests include signal processing, image processing.