Optik 125 (2014) 2274–2282

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A novel image fusion algorithm based on nonsubsampled shearlet transform

Ming Yin [a,*], Wei Liu [a,*], Xia Zhao [b], Yanjun Yin [c], Yu Guo [a]

[a] School of Mathematics, Hefei University of Technology, Hefei 230009, China
[b] Academic Affairs Division, Tongling University, Tongling 244000, China
[c] School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

ARTICLE INFO

Article history:
Received 15 May 2013
Accepted 13 October 2013

Keywords:
Image fusion
Nonsubsampled shearlet transform
Singular value decomposition
Sum-modified-Laplacian

ABSTRACT

To overcome the shortcoming of traditional image fusion method based on multi-scale transform, a novel adaptive image fusion algorithm based on nonsubsampled shearlet transform (NSST) is proposed. Firstly, the NSST is utilized to decompose the source images on various scales and in different directions, and the low frequency sub-band and bandpass sub-band coefficients are obtained. Secondly, for the low frequency sub-band coefficients, the singular value decomposition method in the gradient domain is used to estimate the local structure information of image, and an adaptive 'weighted averaging' fusion rule based on the sigmoid function and the extracted features is presented. To improve the quality of fused image, a novel sum-modified-Laplacian (NSML), which can extract more useful information from source images, is employed as the measurement to select bandpass sub-band coefficients. Finally, the fused image is obtained by performing the inverse NSST on the combined coefficients. The proposed fusion method is verified on several sets of multi-source images, and the experimental results show that the proposed approach can significantly outperform the conventional image fusion methods in terms of both objective evaluation criteria and visual quality.

1. Introduction

With the development of digital image processing technology, data fusion technology and multi-scale analysis theory, the image fusion technology has become a research hotspot. Image fusion is a process in which the different source images from multiple sensors are combined to form a new one which meets the specific needs, so as to effectively merge the advantages of different source images and improve the capability of analyzing and extracting image information [1]. And it has been successfully applied to many areas, such as military affairs, computer vision, medical imaging and remote sensing.

During the last decade, researchers put forward a lot of fusion methods, which can be roughly divided into two groups. One kind is selecting the obviously characteristic pixels in the spatial domain to compose fused images [2,3]. Most of these methods are selecting the fused pixels according to a certain measurement, and the measurement is obtained by calculating a block which is often in the form of window with a certain size. Thus it may easily produce block effect which influences the quality of the fused image. Another kind is mainly a multi-resolution-analysis-based image fusion method. Multi-resolution analysis tools such as traditional pyramid transform and discrete wavelet transform have been widely applied to image fusion [1,4–7]. But most of the traditional pyramid transforms have no directionality, and cannot get more edges and detail information, thus it is unable to achieve good fusion effect. Compared with the traditional pyramid transform, the discrete wavelet transform has many advantages such as localization and direction, and it can well represent the detail information of the image, thus the DWT-based fusion methods are generally superior to the traditional pyramid-based methods.

Unfortunately, 1-D wavelet transform is good at catching point singularities, while the traditional 2-D wavelet basis functions are commonly made up of two 1-D wavelet basis functions tensor product, and its directional selectivity is very limited. Therefore, it cannot provide optimal sparse representation for the contour and edge information of the image. In addition, the DWT is not shift-invariant, and DWT-based fusion methods will probably introduce some false information and serious degraded appearance in the fused images once source images are mis-registration. In order to better represent high order singular features, researchers put forward some multi-scale geometric analysis tools such as Ridgelet transform [8], Curvelet transform [9], and Contourlet transform [10]. Contourlet transform [10] is a new image analysis tool, which is anisotropic and has good directional selectivity. So it can accurately represent the image edges information in different scale and

* Corresponding authors. Tel.: +86 0551 2902592.
E-mail addresses: ymhfut@126.com (M. Yin), lw_feixi@163.com (W. Liu).

0030-4026/$ – see front matter © 2013 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.ijleo.2013.10.064

different direction frequency sub-bands, and some fusion algorithms based on the Contourlet transform have been proposed in recent years [11,12]. The sub-sampling operation is employed in the implementation of Contourlet transform. So it is not shift-invariant and sub-band spectrum aliasing phenomenon takes place, which weakens the directional selectivity of Contourlet transform. To overcome this disadvantage, Cunha et al. [13] proposed nonsubsampled Contourlet transform (NSCT) which is an improved version of Contourlet transform, and it can achieve good fusion effect in NSCT domain [14,15]. But the computational efficiency of NSCT is low. In 2005, Labate et al. proposed a new multi-scale geometric analysis tools: shearlet [16] which is optimally sparse in representing images. The decomposition of shearlet is similar to Contourlet transform, but an important advantage of shearlet over Contourlet transform is that there are no restrictions on the number of directions for the shearing. In addition, the inverse shearlet transform only requires a summation of the shearing filters rather than inverting a directional filter banks, thus the implementation of shearlet is more efficient computationally. Due to these advantages with the shearlet, it has been applied in many image processing fields, such as image denoising [17], edge detection [18], and image fusion [19–21]. The shearlet transform also uses the sub-sampling operation, so it is not shift-invariant and it will produce pseudo-Gibbs phenomenon around the singular point of the denoised and fused images. The nonsubsampled shearlet transform [17] by cascading of nonsubsampled pyramid filter banks and shearing filter banks has all the advantages of the shearlet transform and does not need the implementation of the sub-sampling operation, thus it is shift-invariant, and the impacts of mis-registration on the fused images can be reduced effectively. Therefore, the nonsubsampled shearlet transform is more suitable for image fusion.

A good multi-scale decomposition fusion method not only relies on the transform but also depends on how to select the coefficients in transform domain. Usually, researchers pay more attentions to the high frequency fusion rule. For the low-pass sub-band, the general fusion rule is 'averaging' scheme. The realization of 'averaging' scheme is simple, but it may significantly reduce the contrast of the fused image, thus some information from the source images will be lost. For the high frequency sub-band, the typical fusion rule is approximate 'absolute-maximum-selecting' scheme, during which some 'redundant' information between source images are easily lost and some 'artifacts' are introduced into the fused image.

According to aforementioned analysis, a novel fusion algorithm based on the NSST is proposed in this paper. Source images are decomposed into different sub-bands with NSST. When selecting the low-pass sub-band coefficients, the image local structure information is estimated with the singular value decomposition method in the gradient domain, and the extracted features can effectively reflect edge information of low frequency sub-band. Then the measurement is combined with the sigmoid function, and the adaptive weighted values can be obtained. For the bandpass sub-bands, by the multi-direction of NSST, and we present a novel sum-modified-Laplacian as a clarity index. By compares the clarity indexes of source images and selects the greater clarity index, and the bandpass sub-band coefficients of the corresponding position are considered fused coefficients. Experimental results show that the proposed method performs very well. For the mis-registration source image, this method also has good robustness.

The main structure of the paper is as follows: Section 2 reviews the theory of the NSST in brief. Section 3 illustrates the main framework of the proposed method and describes the fusion algorithm using NSST in detail. Experimental results and the performance of proposed method are discussed in Section 4. Conclusions are presented in the last section.

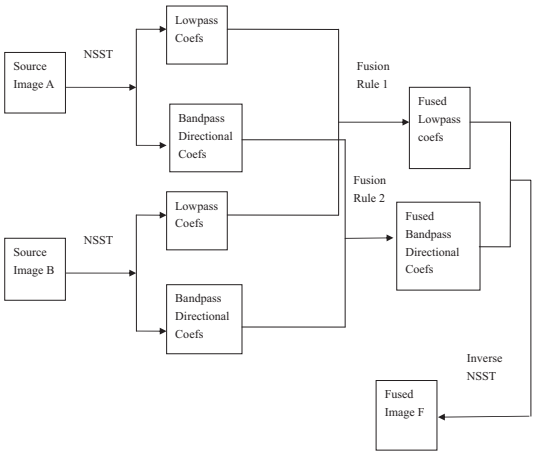

Fig. 1. Block diagram of image fusion based on the NSST.

2. The nonsubsampled shearlet transform

In this section, we review the theory of the NSST in brief which will be used in the subsequent sections.

The implementation of the NSST [17] is divided into two steps: nonsubsampled pyramid filter banks [13] and shearing filter banks, which have many advantages, such as strong selectivity of anisotropic directionality, shift-invariance and stability, and it is the true 2-D sparse representation for images with edges. Nonsubsampled pyramid filter banks realize multi-scale partition of NSST. We can get $k+1$ sub-bands which are the same size with the original image when performing scale k nonsubsampled pyramid filter banks decomposition. Shift-invariant shearing filter banks accomplish the directional localization, and through translation window function in the grid false polarization we can realize shearing filters of standard shearlets. The implementation process needs the sampling operation, so standard shearlets lack shift-invariance.

The NSST maps the standard shearing filters from false polarization grid system into a Cartesian coordinate system. Through the inverse Fourier transform, we can prove its operation can be completed directly through the 2-D convolution. This transform avoids the sampling operation and possesses shift-invariance. In this paper, the main work is to probe into the image fusion method, and more details of NSST can be found in [17].

3. The NSST-based image fusion method

3.1. The framework of NSST-based fusion algorithm

To simplify the discussion, we assume the fusion process is to generate a composite image F from a pair of source images denoted by A and B. Fig. 1 illustrates the block diagram of the proposed image fusion algorithm. Firstly, we assume here that source images have been well registered. The general procedure of the proposed algorithm can be summarized as follows:

(1) The source images A and B are decomposed by the NSST to obtain one low-frequency sub-band and a series of directional bandpass sub-bands, respectively.
(2) Employ different fusion rules to merge the low-frequency sub-band coefficients and the directional bandpass sub-bands coefficients of source images, respectively, where the fusion rules of low-frequency sub-band and directional bandpass sub-bands are all directly applied to the NSST coefficients. The NSST coefficients of the fused image F are thus obtained.

(3) Apply the inverse NSST to the fused coefficients to obtain the fused image F.

We will discuss in detail the image fusion algorithm using the NSST in the following chapter.

3.2. Selection principle for low-pass sub-band coefficients

The lowpass sub-band, obtained by the NSST decomposition, contains the main energy and represents the approximation component of the source image. At present, the processing of lowpass sub-band coefficients uses averaging rule [22] which is simple. However, this will reduce the fused image contrast and lose some useful information of the source image. According to the findings in the physiological and psychological research, the human vision system is highly sensitive to edges, but insensitive to real luminance at independent positions. The accurate calculation of clarity index of low frequency coefficients decides the quality of the fused image. In past decades, researchers put forward many clarity indexes in space domain, such as energy of image gradient, spatial frequency, Laplacian energy, and so on. They all can measure the variation of pixels. The pixels with greater values of these measurements, when source images are compared, are considered from clear parts and selected as the fused pixels. Since sub-bands of source images in NSST domain can be viewed as images and the variation of sub-bands also exists, it is reasonable to utilize these measurements as the clarity indexes to select coefficients from the clear parts of source images. In order to improve the quality of the fused image, the clarity index should be able to highlight the distinguishing feature of low frequency coefficients. Therefore, we construct a structure descriptor by using the results which are obtained using singular value decomposition in the gradient field, and the response values of local structure descriptor as the characteristic values of low frequency sub-band coefficients, namely, the corresponding clarity indexes.

3.2.1. Singular value decomposition in the gradient field

Takeda et al. constructed Steering Gaussian kernel by using singular value decomposition method in gradient field [23]; Thaipanich et al. improved the performance of the non-local means by using this method [24]. For an image f, the main steps for singular value decomposition method in gradient field can be summarized as follows.

(1) Calculate gradient value of f;
(2) Group local area gradient values into matrix $\boldsymbol{G}$ of size $N \times 2$;
(3) Compute matrix $\boldsymbol{G}$'s truncated singular value decomposition, the specific form is

$$\boldsymbol{G} = [\nabla f^T(1)\nabla f^T(2)\cdots\nabla f^T(N)]^T = \boldsymbol{USV}T$$

where $\nabla f^T(i) = [(\partial f(i)/\partial x)(\partial f(i)/\partial y]^T$ is the gradient of image f at point i, N is the number of pixels in local area, $\boldsymbol{U}$ is an $N \times N$ orthogonal matrix, $\boldsymbol{S}$ is an $N \times 2$ matrix that contains singular values and $\boldsymbol{V}$ is an 2×2 orthogonal matrix which gives the dominant orientation of the gradient field, and the second column of the matrix $\boldsymbol{V}$, $v_2 = [v_1, v_2]^T$, is a singular vector corresponding to the smallest nonzero singular.

3.2.2. Local structure descriptor

The different areas of image contain different content; generally, it can be divided into flat areas, edge and texture direction consistency areas and abundant details areas (such as the cross edge). The singular values λ_1, λ_2 which reflect the size of the energy change in the direction of eigenvector. In the flat areas, the gray values vary little and singular values $\lambda_1 \approx \lambda_2 \approx 0$; in edge and texture direction consistency areas, the gray values which cross the direction of edge have big change, and singular values $\lambda_1 > \lambda_2 \approx 0$; in abundant details areas, the gray value of each direction changes, and singular values $\lambda_1 > \lambda_2 > 0$. Because the singular value decomposition has a strong antinoise ability and the noise has no any direction, singular values λ_1, λ_2 can well represent the image's content of the local area. The measure of energy of the local direction is defined as

$$e(i) = \lambda_1(i) + \lambda_2(i) \tag{1}$$

The size of e can reflect the basic structural mode of local image area. If the value of e is large, the current area contains rich details information; if the value of e is small, then the current region is smooth. Therefore, e can be used as a structure descriptor of the local structural information of image, and the value of e can represent the feature of image's content. Fig. 2 is Barbara and noisy images and the corresponding extractive feature images using Eq. (1). As can be seen in Fig. 2, the extractive feature image of noisy image which corrupted by simulated additive Gaussian white noise with standard deviation equal to 10 is very similar to the original image, and the feature images can well reflect the contents of different areas of the images, such as flat, edge (the table leg) and abundant details areas (the scarf). Since the singular value decomposition has a strong antinoise ability, the proposed local structure descriptor which is constructed by using singular value decomposition method in the gradient field is robust and efficient.

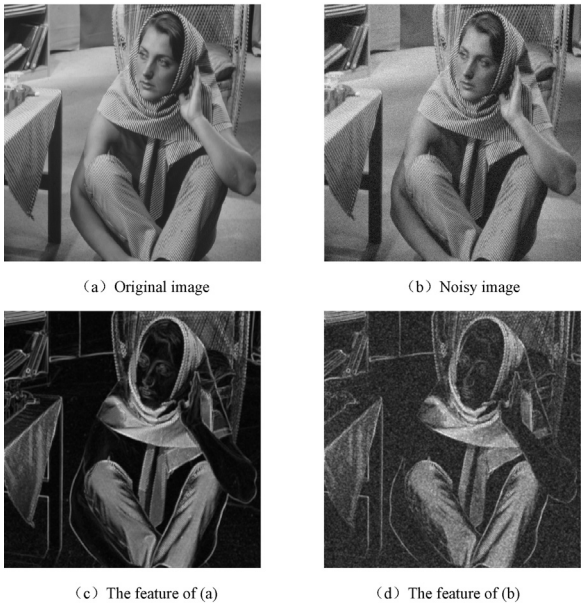

Fig. 2. The comparison of the extractive feature images of Barbara and its noisy image.

3.2.3. The adaptive 'weighted averaging' fusion rule

Since the low frequency sub-band's information reflects the general view features of the original image and it is the approximation component of the source image, the local structure descriptor which is defined in above can describe the content of the low frequency sub-band, and it is considered to be a clarity index. The averaging rule is simple, but it is not adaptive in fusing source images. Therefore, in this paper, we use an adaptive 'weighted averaging' fusion rule, and the weighted ω are determined by a sigmoid function. The fused low frequency sub-band coefficient is

$$L^F(i,j) = \omega(i,j) \times L^A(i,j) + (1 - \omega(i,j)) \times L^B(i,j)$$

where $L^A(i,j)$, $L^B(i,j)$ are NSST low frequency sub-band coefficients of source images at position(i,j), respectively, and $L^F(i,j)$ is low frequency coefficient of the fused image. In this paper, we construct a new sigmoid function

$$\omega(i,j)=\begin{cases}1-\dfrac{\sum_{k=1}^{k}(-1)^{k-1}(x)^{K-k}}{x^K+\sum_{k=1}^{K}(-1)^{k-1}(x)^{K-k}} & \text{if } 0<x<1\\ \dfrac{\sum_{k=1}^{K}(-1)^{k-1}(x)^{k}}{1+\sum_{k=1}^{K}(-1)^{k-1}(x)^{k}} & \text{if } 1\le x\end{cases} \quad (2)$$

$$\text{and } x=\frac{e^A(i,j)}{e^B(i,j)}$$

where K is shrink factor of the sigmoid function and an odd number and $K>1$, $e^A(i,j)$, $e^A(i,j)$ are the corresponding clarity index values of low-frequency coefficients at position (i,j), respectively. As can be seen from Eq. (2), when $K\to+\infty$, the proposed algorithm is equivalent to the choose-max scheme. For the same K, when $(e^A(i,j))(e^B(i,j))\to 0$, the weight coefficient $\omega(i,j)$ is closer to zero, and the fused coefficient mainly comes from the source image B; otherwise, when $(e^A(i,j))(e^B(i,j))\gg 1$, the weight coefficient $\omega(i,j)$ is closer to one, and the fused coefficient mainly comes from the source image A; when $(e^A(i,j))(e^B(i,j))\to 1$, the weight coefficient $\omega(i,j)$ is closer to 1/2, and the proposed algorithm is equivalent to the average fusion method. Therefore, the proposed fusion rule can be dynamic in selecting weighted value according to the low frequency sub-band's feature, thus the fusion rule is adaptive.

3.3. Selection principle for directional band-pass sub-band coefficients

Under the assumption that image detail information are contained in the directional band-pass sub-bands in NSST domain. Since the sum-modified-Laplacian can well reflect the feature information of the edges of the images and properly represent the focus features and sharpness of the images, Huang and Jing used the sum-modified-Laplacian for image fusion and achieved good results [2]. In order to obtain fused images which have better visual effect and more detail information, we design a new fusion scheme based on a novel sum-modified-Laplacian. The sum-modified-Laplacian is defined as follows:

$$\begin{aligned}SML^{l,k}(i,j)&=\sum_{a=-P}^{P}\sum_{b=-Q}^{Q}\left[ML^{l,k}(i+a,j+b)\right]^2\\ ML^{l,k}(i,j)&=\left|2MP^{l,k}(i,j)-MP^{l,k}(i-step,j)-MP^{l,k}(i+step,j)\right|\\ &\quad+\left|2MP^{l,k}(i,j)-MP^{l,k}(i,j-step)-MP^{l,k}(i,j+step)\right|\end{aligned} \quad (3)$$

where *step* is variable spacing between coefficients and in this paper *step* always equals one, $MP^{l,k}(i,j)$ denotes the directional band-pass sub-bands coefficient located at (i,j) in the lth scale and kth direction sub-band, $SML^{l,k}(i,j)$is the corresponding sum-modified-Laplacian, the parameters P and Q determine the square-shaped window with size $(2P+1)\times(2Q+1)$ are used to compute the measurement, and the size may be 3×3 or 5×5. The sum-modified-Laplacian which is defined in Eq. (3) neither outstands the importance of center pixel, nor considers the distance between the center pixel and neighborhood pixels. When some long-distance pixels have big contribution to the sum-modified-Laplacian, the center pixel's information may be lost, which may lead to unsatisfactory fusion effect. The expression of novel sum-modified-Laplacian is:

$$NSML^{l,k}(i,j)=\sum_{a=-P}^{P}\sum_{b=-Q}^{Q}w(a,b)\left[ML^{l,k}(i+a,j+b)\right]^2 \quad (4)$$

where $w(a,b)$ is the weight of $ML^{l,k}(i+a,j+b)$, $NSML^{l,k}(i,j)$ denotes the novel sum-modified-Laplacian of multi-scale products coefficient located at (i,j) in the lth scale and kth direction sub-band. From Eq. (4), we can find that the choice of weights $w(a,b)$ has a big impact on the calculation of salience measure and fusion effect. To simplify the discussion, $W(2p+1)\times(2Q+1)$ is noted as the weight matrix which needs to satisfy the following conditions:

(1) Highlight the center pixel.
(2) The closer pixels to center pixel, the greater the weight.
(3) $\sum_a\sum_b w(a,b)=1$.

Chai et al. only gave the specific form matrix weight which size is 3×3, but they did not give the general construct method of weight matrix [25]. Below, we give the construction method:

$$\begin{cases}W_{(2p+1)\times(2Q+1)}=[w(a,b)]_{(2p+1)\times(2Q+1)}=\dfrac{1}{\sum_{a=-P}^{p}\sum_{b=-Q}^{Q}\hat{W}(a,b)}\hat{W}_{(2p+1)\times(2Q+1)}\\ \hat{W}_{(2p+1)\times(2Q+1)}=[\hat{W}(a,b)]_{(2p+1)\times(2Q+1)}=X^TY\end{cases} \quad (5)$$

where $X=[1,2,\ldots,P+1,\ldots,21]_{1\times(2P+1)}$ and $Y=[1,2,\ldots,Q+1,\ldots,21]_{1\times(2Q+1)}$.

$$\text{When } P=Q=1,\quad W_{3\times3}=\frac{1}{16}\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix};$$

$$\text{When } P=Q=2,\quad W_{5\times5}=\frac{1}{81}\begin{bmatrix}1&2&3&2&1\\2&4&6&4&2\\3&6&9&6&3\\2&4&6&4&2\\1&2&3&2&1\end{bmatrix}.$$

Thus we found a general calculation method of weight matrix. We know that, multi-directional selectivity is an advantage of the NSST. In order to make full use of this advantage, it should be considered in calculation of salience measure. The priority should be given to the pixels which have abundant edge information which contribute a lot to novel sum-modified-Laplacian. Therefore, we should choose different weight matrixes for different directional band-pass sub-bands and properly enlarge the weight in different directions. In this paper, we use 4, 4, 8, 8 directions in the scales from coarser to finer, respectively. So the concrete weight matrixes will be listed in the following. For more directional decomposition, the weight matrixes can be deduced in the same way. For the directional band-pass sub-bands with 4 directions, the novel weight matrix is:

$$\begin{cases}W^{\theta}_{(2p+1)\times(2Q+1)}=[w^{\theta}(a,b)]_{(2p+1)\times(2Q+1)}=\dfrac{1}{\sum_{a=-P}^{p}\sum_{b=-Q}^{q}\hat{W}^{\theta}(a,b)}\hat{W}^{\theta}_{(2p+1)\times(2Q+1)}\\ \hat{W}^{\theta}_{(2p+1)\times(2Q+1)}=[\hat{w}^{\theta}(a,b)]_{(2p+1)\times(2Q+1)}=X^TY+\delta^{\theta}_{(2p+1)\times(2Q+1)}\end{cases}$$

where the definitions of X and Y are similar to Eq. (5). $W^{\theta}_{(2p+1)\times(2Q+1)}$ $(\theta=0,1,2,3)$ denotes the novel weight matrix. $\theta=0,1,2,3$ represents the 4 directions that are 0°, 45°, 90°, 135°. We take $P=Q=1$. The matrix $\delta^{\theta}_{(2p+1)\times(2Q+1)}$ is edge weight increase operator. In this paper, the $\delta^{\theta}_{(2p+1)\times(2Q+1)}$ is given as:

$$\delta^0_{3\times3} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \delta^1_{3\times3} = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 3 & 0 & 0 \end{bmatrix}, \quad \delta^2_{3\times3} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \quad \delta^3_{3\times3} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

For the directional band-pass sub-bands with 8 directions, we take $P = Q = 2$, and the corresponding edge weight increase operators are:

$$\delta^0_{5\times5} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 6 & 3 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \delta^1_{5\times5} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 \\ 0 & 3 & 0 & 3 & 0 \\ 7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \delta^2_{5\times5} = \begin{bmatrix} 0 & 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\delta^3_{5\times5} = \begin{bmatrix} 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \end{bmatrix}, \quad \delta^4_{5\times5} = \begin{bmatrix} 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \end{bmatrix}, \quad \delta^5_{5\times5} = \begin{bmatrix} 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 \end{bmatrix}$$

$$\delta^6_{5\times5} = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{bmatrix}, \quad \delta^7_{5\times5} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 7 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus the novel sum-modified-Laplacian can better reflect the local distinguishing feature of band-pass sub-band image in different directions. According to Eq. (4), we can calculate the novel sum-modified-Laplacian of directional band-pass sub-bands coefficients of source images, namely, $NSML_A^{l,k}(i,j)$, $NSML_B^{l,k}(i,j)$, which denote the novel sum-modified-Laplacian of $MP_A^{l,k}(i,j)$, $MP_B^{l,k}(i,j)$, respectively. Therefore, the proposed selection rule for directional band-pass sub-band coefficients is defined as:

$$H_F^{l,k}(i,j) = \begin{cases} H_A^{l,k}(i,j), & \text{if } NSML_A^{l,k}(i,j) \geq NSML_B^{l,k}(i,j) \\ H_B^{l,k}(i,j), & \text{if } NSML_A^{l,k}(i,j) < NSML_B^{l,k}(i,j) \end{cases}$$

It means coefficients with maximum novel sum-modified-Laplacian measurement are selected as the coefficients of fused image.

4. Experiments and performance analysis

To verify the performance of the proposed fusion scheme, we made two sets of fusion experiments. In the first set experiment, four pairs of the multi-focus images are used as the source images. In order to show the advantage of the proposed fusion method, we compared the proposed method with five fusion methods, namely GP-based, DWT-based, Contourlet-based, ST-based, NSST-simple-based method. In all of these methods, the low-pass sub-band coefficients and the band-pass sub-band coefficients are simply merged by the 'averaging' scheme and the 'absolute maximum choosing' scheme, respectively. The second set experiment

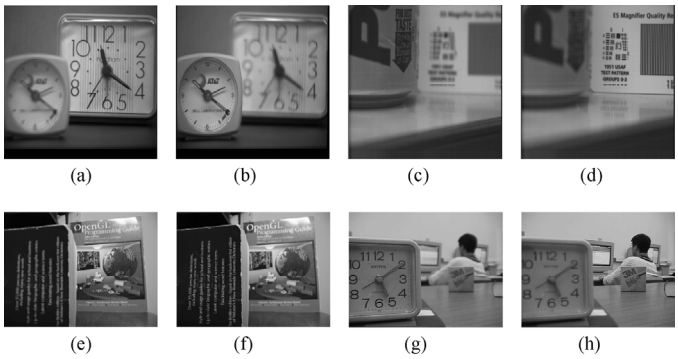

Fig. 3. Source images for multi-focus fusion. (a) and (b), (c) and (d), (e) and (f), (g) and (h) are the pairs.

is performed on source images from different type of sensors. The proposed method is compared with GP-based, DWT-based, Contourlet-based, ST-based, MP-NSML-max method in which the low-pass sub-band coefficients are simply merged by the 'averaging' scheme and the band-pass sub-band coefficients are merged by the proposed rule, and Low-Adaptive-Mean method in which the low-pass sub-band coefficients are merged by the proposed adaptive 'weighted averaging' fusion rule and the band-pass sub-band coefficients are simply merged 'absolute maximum choosing' scheme. In experiments, besides visual appearance observation, mutual information (MI) [26] and $Q^{AB/F}$ [27] are employed as objective criteria. The two criteria do not require the information of ideal fused image. MI essentially computes how much information from source images is transferred to the fused image, while $Q^{AB/F}$ measures the amount of edge information transferred from the source images to the fused images using a Sobel edge detector. For both criteria, the large the value gets, the better the fusion result is. The images are all decomposed into 4 levels by DWT (with the basis of Haar), Contourlet, ST and NSST. We use 4, 4, 8, 8 directions in the scales from coarser to finer for Contourlet and NSST, respectively. The pyramid filter for Contourlet is set as '9–7' and the directional filter is set as 'pkva'. For the proposed method, the pyramid filter of NSST is set as 'maxflat'.

4.1. Fusion of multi-focus images

Firstly, four pair multi-focus images are used in fusion experiment, the first three pairs of multi-focus images with perfect registration and the last pair of multi-focus image with misregistration. It is easily noticed that there is a slight movement of the student's head. One group focus on the right, and another on the left. Figs. 3–7 illustrate the source images, the fusion results and the difference images obtained by different methods. In order to clearly display the residual part, the difference images in Fig. 5 are the square of the original residual images. And Figs. 8 and 9 are parts of Figs. 5 and 7. For the clear part, the lower residue features in the difference images are, the better the fusion method transfers features of source images to the fused images. From Figs. 4–7, we can see that all the methods can achieve the fusion purpose to a certain degree, and they can inject the information of source images into the fused image. As can be seen in Fig. 8, the fused image which is obtained by GP-based method presents low contrast and fuzzy appearance; the fused images of DWT-based, Contourlet-based and ST-based appear 'ripples' at the edge and false information (e.g. letters become fuzzy); the fused image of NSST-simple-based method can restrain this fuzzy phenomenon to certain extent. It is mainly because the Gradient pyramid, wavelet transform, Contourlet transform and shearlet transform do not have shift-invariant, which leads to the fused image distortion and poor visual effect, while the NSST can

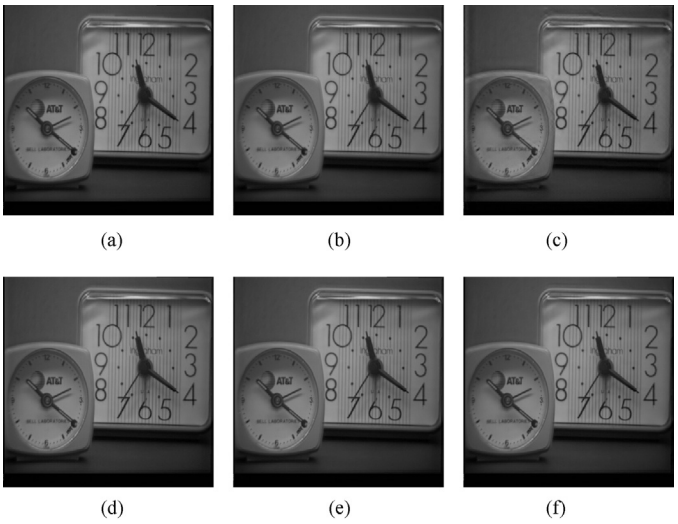

Fig. 4. The fusion results of Fig. 3(a) and (b): (a)–(f) fused images using the GP-based, DWT-based, Contourlet-based, ST-simple-based, NSST-simple-based and the proposed method, respectively.

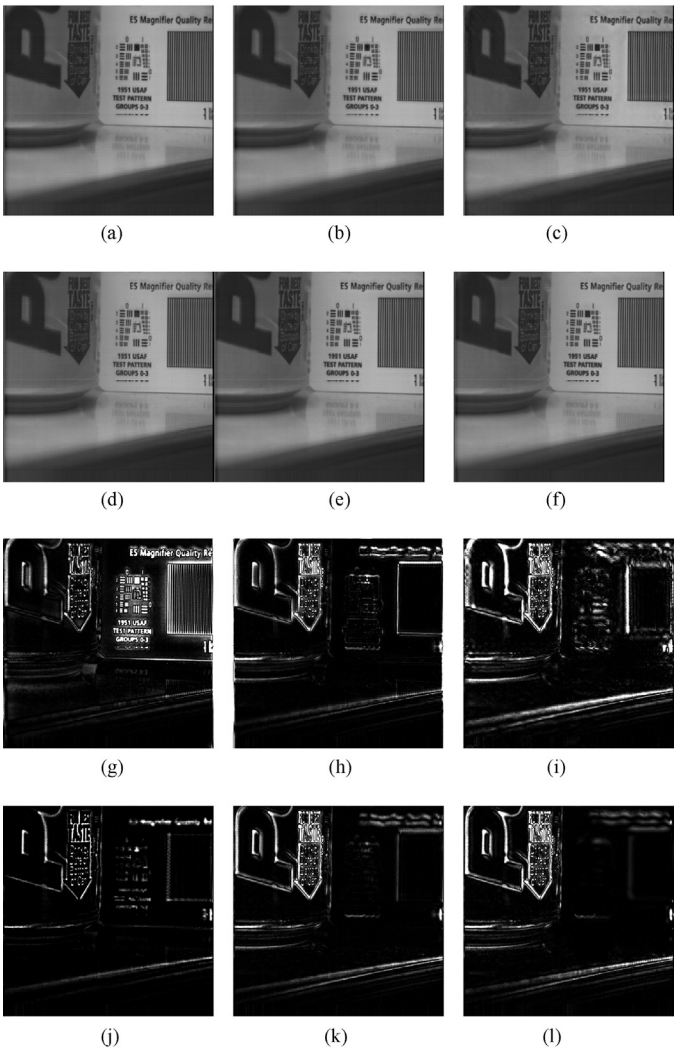

Fig. 5. The fusion results of Fig. 3(c) and (d): (a)–(f) fused images using the GP-based, DWT-based, Contourlet-based, ST-simple-based, NSST-simple-based and the proposed method, respectively; (g)–(l) are the difference images which (a)–(f) minus the source image shown in Fig. 3(d).

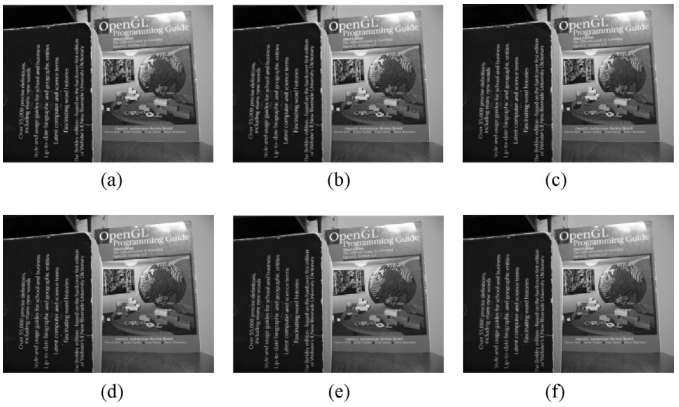

Fig. 6. The fusion results of Fig. 3(e) and (f): (a)–(f) fused images using the GP-based, DWT-based, Contourlet-based, ST-simple-based, NSST-simple-based and the proposed method, respectively.

avoid spectrum aliasing effect in the process of decomposition and reconstruction. The fused image of the proposed method can effectively restrain the fuzzy phenomena, and visual effect is much better. As can be seen in Fig. 5, the GP-based method, DWT-based method, Contourlet-based method and ST-based method cannot effectively extract the source images' features information, and the source images' features information can't be injected into the fused image as well. The difference images keep a lot of information of source images, and fusion is not very successful. The NSST-simple-based method can improve the quality of fused image, but the result is not very ideal yet; compared with the results of the simple fusion methods above, the difference image which is obtained by the proposed method contains little feature (e.g. the right middle part), which illustrates that the proposed method of feature extraction is very effective. Using this method, we can better extract features of source images and transfer them to the fused image, and the useful information of the source images can be well retained. Through the comparison of these fused images, it is known that the proposed image fusion method works well, and each target of foreground and background in the fused images is very clear. The fused images of mis-registration source images are given in Fig. 7. As can be seen from Fig. 7, the proposed method introduces a little of artificial information in the process of the image fusion, and the fused image looks more natural. It also shows that the proposed algorithm injects more information of source images into fused image. Fig. 9 is the part of Fig. 7; it can be obviously found that the fused image

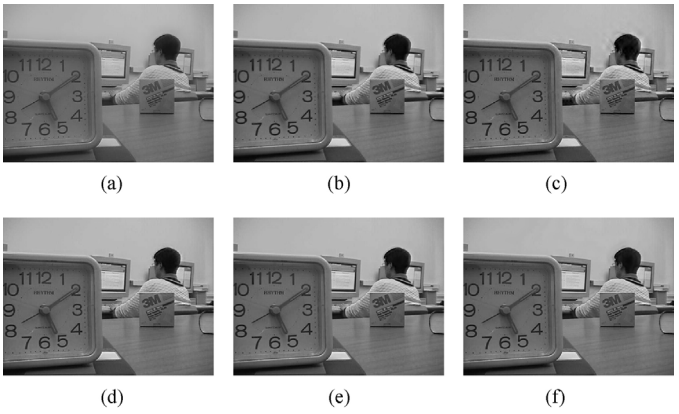

Fig. 7. The fusion results of Fig. 3(g) and (h): (a)–(f) fused images using the GP-based, DWT-based, Contourlet-based, ST-simple-based, NSST-simple-based and the proposed method, respectively.

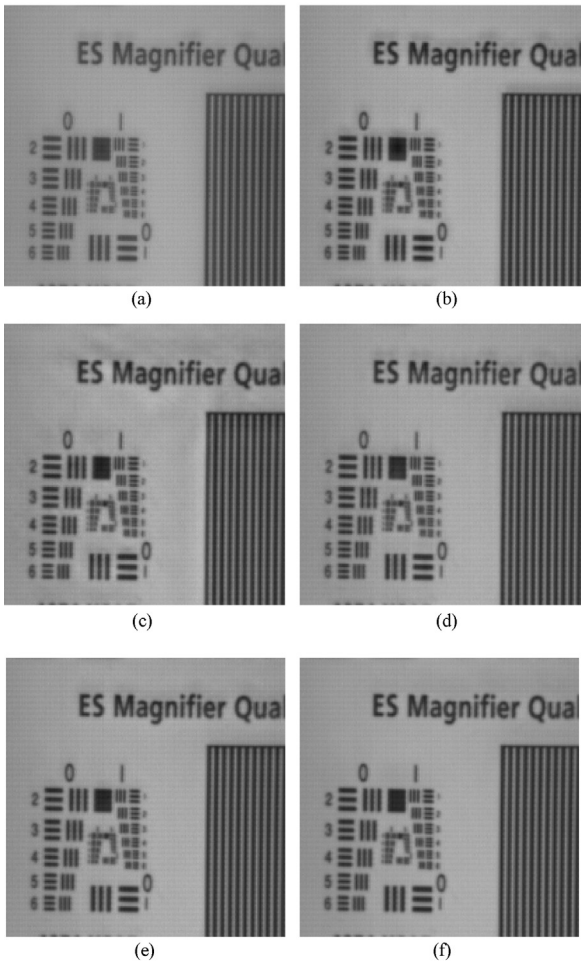

Fig. 8. Parts of the images of Fig. 5.

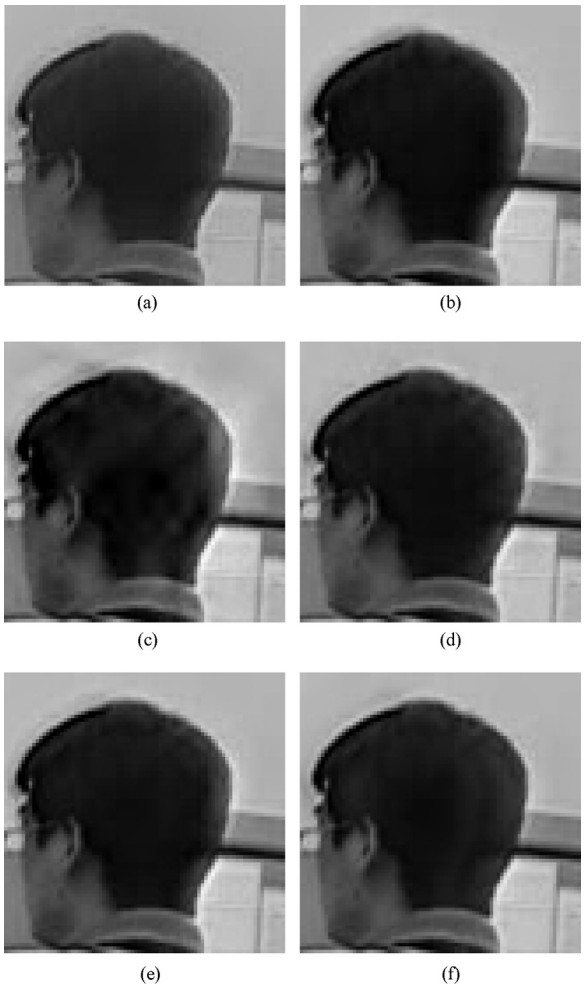

Fig. 9. Parts of the images of Fig. 7.

of Contourlet-based method has serious pseudo-Gibbs phenomena. The fuzzy phenomena of the fused images of GP-based, DWT-based and ST-based have been improved, but it is still not very ideal. The fused images of NSST-based methods are clearer than those from the other methods, which illustrates that shift-invariant can effectively overcome pseudo-Gibbs phenomena and reduce the effect of registration error.

The values of MI and $Q^{AB/F}$ of Fig. 3 are listed in Table 1. It can be seen that the superiority of NSST is more obvious with GP, DWT, Contourlet transform and ST. The proposed scheme almost provides the highest objective data, which illustrates that the proposed algorithm can obtain the fused image which is strongly correlated with the source images, and the edge texture features in the fusion process can be well kept. The objective evaluation results coincide with the visual effect very well. Through the above discussion, we can see that the proposed scheme is effective in the application of multi-focus image fusion.

4.2. Fusion of medical images

Fig. 10 shows the fusion results of two medical images. Fig. 10(a) is a CT image, whose brightness is related to the tissue density. The bone is shown clearly, but soft tissue is invisible. Fig. 10(b) is an MRI image, in which the brightness is related to the number of hydrogen atoms in the tissue. So the soft tissue is shown clearly, but the bone is invisible. From Fig. 10, we can see that all the fused images can retain bony detail information of CT image and the soft tissue information of MRI image. The fused images of GP-based and ST-based methods cannot well inject CT image's information and the contrast is low. The details of the fused image of Contourlet-based method are serious fuzzy. The whole contrast of the fused images of DWT-based and MP-NSML-max methods is enhanced, but some parts are not clear. The overall contrast of the fused images of Low-Adaptive-Mean and the proposed methods is relatively high compared with the above mentioned methods, and the fused images of the two methods well comprehensively integrate the CT and MRI images' information. The edge and contour information of the bone and the soft tissue of the fused images are close to the CT and MRI source images, and the clarity of the fused images is highest. The comparison of MP-SML-Max and the proposed method shows that the fusion rules of low frequency have great effect on the whole visual effect for the medical image fusion.

For comparison, the objective values of Fig. 10 are listed in Table 2. It can be seen that the greatest values demonstrate that the proposed method transferred more information to the fused images than other fusion methods and fusion effect is the best. From the above analysis, we can find that the proposed fusion scheme shows significant improvement over the other methods in the application of the source images from different types of sensors.

Table 1
Comparison on objective criteria using different fusion methods (multi-focus images).

Images	Criteria	GP-based	DWT-based	Contourlet-based	ST-based	NSST-simple-based	Proposed method
Fig. 3(a)	MI	6.2205	6.6066	5.9679	6.5484	6.6816	7.1947
Fig. 3(b)	$Q^{AB/F}$	0.6395	0.6836	0.6192	0.6328	0.6740	0.6974
Fig. 3(c)	MI	6.1351	6.6046	6.2110	6.6880	6.8139	6.9452
Fig. 3(d)	$Q^{AB/F}$	0.7395	0.7494	0.7131	0.7464	0.7681	0.7806
Fig. 3(e)	MI	5.9574	7.8597	7.3350	7.9850	8.0644	8.1741
Fig. 3(f)	$Q^{AB/F}$	0.7696	0.8409	0.8071	0.8293	0.8362	0.8342
Fig. 3(g)	MI	6.0606	6.9978	6.3387	7.0558	7.2159	7.3529
Fig. 3(h)	$Q^{AB/F}$	0.6876	0.7114	0.6687	0.7024	0.7180	0.7366

Table 2
Performance of the different fusion methods on processing Fig. 10.

Images	Criteria	GP-based	DWT-based	Contourlet-based	ST-based	Low-Adaptive-Mean	MP-NSML-max	Proposed method
Fig. 10	MI	2.8883	2.5433	1.1148	2.1275	2.8539	2.1995	2.8926
	$Q^{AB/F}$	0.6836	0.7284	0.3978	0.6947	0.7167	0.6979	0.7275

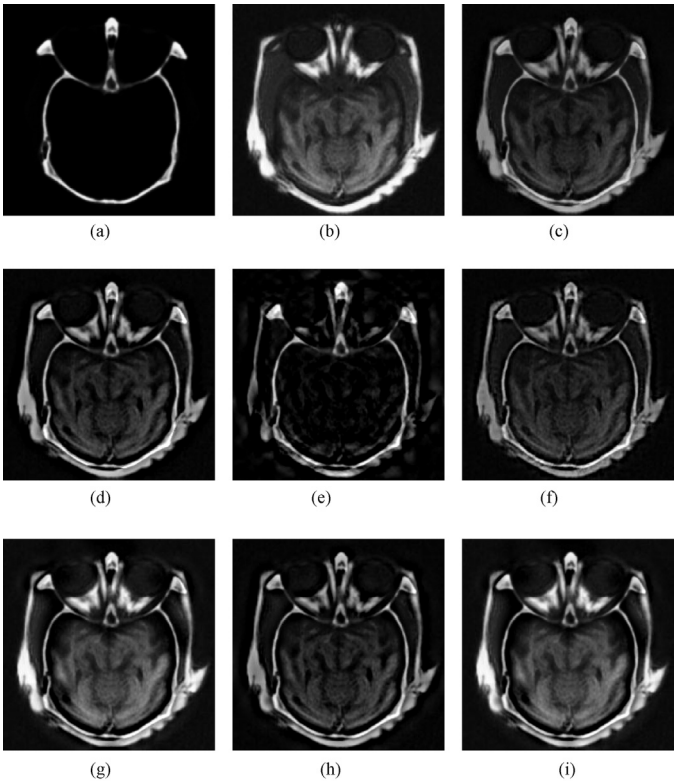

Fig. 10. The medical source images and fusion results: (a) CT image; (b) MRI image; (c)–(h) fused images using the GP-based, DWT-based, Contourlet-based, ST-simple-based, Low-Adaptive-Mean method, MP-NSML-max method and the proposed method, respectively.

5. Conclusions

In this paper, NSST is introduced into the image fusion field. As a novel multi-scale geometric analysis tool, it can better sparse represent the source image and effectively reduce registration error influence on fusion performance. Therefore, a novel image fusion scheme based on NSST is proposed in this paper. According to the characteristics of each sub-band, we design fusion rules respectively for each sub-band. For the low-pass sub-band coefficients, the local structure information is estimated using the singular value decomposition method in the gradient domain. And by combining the measurement with sigmoid function, adaptive weighted values can be obtained. For the band-pass sub-bands, we present a select rule by using a novel sum-modified-Laplacian. Several sets of experimental results show that the proposed method can capture the edge detail information of the source image and reduce the 'artificial information' and the impact of registration error, and improve the visual effect of the fused image.

Acknowledgments

The authors would like to thank the anonymous reviewers and the associate editor for their constructive comments and suggestions to our paper. The original test images, which are all available on line at www.imagefusion.org, are kindly provided by Mr. Lewis and Dr. Nikolov, by Manchester University and Dr. Petrovic and by Dr. Toet, respectively. This work was supported by the Natural Science Foundation Project of Anhui Province (No. 1308085MA09) and the Science Foundation Project of Education Department of Anhui Province (No. 2013AJZR0039).

References

[1] R.S. Blum, Z. Liu, Multi-sensor Image Fusion and Its Applications, CRC Press, Boca Raton, 2005, pp. 1–10.
[2] W. Huang, Z.L. Jing, Evaluation of focus measures in multi-focus image fusion, Pattern Recogn. Lett. 28 (2007) 493–500.
[3] S.T. Li, B. Yang, Multi-focus image fusion using region segmentation and spatial frequency, Image Vision Comput. 26 (2008) 971–979.
[4] S.T. Li, J.T. Kwok, Y.N. Wang, Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images, Inform. Fusion 3 (2002) 17–23.
[5] H.H. Wang, A new multiwavelet-based approach to image fusion, J. Math. Imaging Vision 21 (2004) 177–192.
[6] J.J. Lewis, R.J.O. Callaghan, S.G. Nikolov, et al., Pixel-and region-based image fusion with complex wavelets, Inform. Fusion 8 (2007) 119–130.
[7] W. Li, X.F. Zhu, An image fusion algorithm based on second generation wavelet transform and its performance evaluation, Acta Aotomat. Sin. 33 (2007) 817–822.
[8] E. Candes, Ridgelets: Theory and Applications, Department of Statistics, Stanford University, USA, 1998.
[9] J.L. Starck, E.J. Candes, D.L. Donoho, The curvelet transform for image denoising, IEEE Trans. Image Process. 21 (2002) 131–141.
[10] M.N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (2005) 2091–2106.
[11] Q.G. Miao, B.S. Wang, A novel image fusion method using contourlet transform, in: Proc. 2006 International Conference on Communications, Circuits and Systems Processing, 2006, pp. 548–552.
[12] L. Yang, B.L. Guo, W. Ni, Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform, Neurocomputing 72 (2008) 203–211.
[13] A.L. Cunha, J.P. Zhou, M.N. Do, The nonsubsampled contourlet transform: theory, design and application, IEEE Trans. Image Process. 15 (2006) 3089–3101.
[14] Q. Zhang, B.L. Guo, Multifocus fusion using the nonsubsampled contourlet transform, Signal Process. 89 (2009) 1334–1346.
[15] H.F. Li, Y. Chai, Z.F. Li, Multi-focus image fusion based on nonsubsampled contourlet transform and focused regions detection, Optik 124 (2013) 40–51.
[16] D. Labate, W. Lim, G. Kutyniok, et al., Sparse multidimensional representation using shearlets, in: Wavelets XI (San Diego, CA, 2005), SPIE Proc., vol. 5914, SPIE, Bellingham, WA, 2005, pp. 254–262.

[17] G. Easley, D. Labate, W.Q. Lim., Sparse directional image representations using the discrete shearlet transform, Appl. Comput. Harmonic Anal. 25 (2008) 25–46.
[18] S. Yi, D. Labate, G.R. Easley, et al., A shearlet approach to edge analysis and detection, IEEE Trans. Image Process. 18 (2009) 929–941.
[19] C. Deng, S. Wang, X. Chen, Remote sensing images fusion algorithm based on shearlet transform, in: Proceeding of International Conference on Environmental Science and Information Application Technology, ACM, Wuhan, China, 2009, pp. 451–454.
[20] Q.G. Miao, C. Shi, P.F. Xu, et al., A novel algorithm of image fusion using shearlets, Opt. Commun. 284 (2011) 1540–1547.
[21] L. Wang, B. Li, L.F. Tian, Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients, Inform. Fusion (2012), http://dx.doi.org/10.1016/j.inffus.2012.03.002.
[22] Z. Zhang, R.S. Blum, A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application, Proc. IEEE 87 (1999) 1315–1326.
[23] H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction, IEEE Trans. Image Process. 16 (2007) 349–366.
[24] T. Thaipanich, B.T. Oh, P.H. Wu, et al., Improved image denoising with adaptive nonlocal means (ANL-Means) algorithm, IEEE Trans. Consum. Electron. 56 (2010) 2623–2630.
[25] Y. Chai, H.F. Li, J.F. Qu, Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain, Opt. Commun. 283 (2010) 3591–3602.
[26] G.H. Qu, D.L. Zhang, P.F. Yan, Information measure for performance of image fusion, Electron. Lett. 38 (2002) 313–315.
[27] V. Petrovic, C. Xydeas, On the effects of sensor noise in pixel-level image fusion performance, in: Proc. of the Third International Conference on Image Fusion, vol. 2, IEEE Press, 2000, pp. 14–19.