

MULTI-FOCUS IMAGE FUSION BASED ON NONSUBSAMPLED SHEARLET TRANSFORM AND PULSE COUPLED NEURAL NETWORK WITH SELF-SIMILARITY AND DEPTH INFORMATION

Liu Shuaiqi^{a,b}, Shi Mingzhu^c, Zhao Jie^{a,b}, Geng Peng^d
and Zhang Zhong^c

^aCollege of Electronic and Information Engineering, Hebei University, Baoding
071000, P. R. China

^bKey Laboratory of Digital Medical Engineering of Hebei Province, Baoding
071000, P. R. China

^cCollege of Electronic and Communication Engineering, Tianjin Normal University,
Tianjin 300387, P. R. China

^dSchool of Information Science and Technology, Shijiazhuang Tiedao University,
Shijiazhuang, P. R. China

Abstract

Combined with the shared similarity among multiple source images and depth of field in a camera, a new image fusion algorithm based on nonsubsampling shearlet transform (NSST) domain and pulse coupled neural network (PCNN) is proposed. First, NSST is utilized for decomposition of the source images, the low frequency coefficients are fused by weight votes in the structure-driven regions by using shared similarity and depth of field (SSSID), and the high frequency coefficients are fused by fired map of PCNN which motivated by larger

*Corresponding author.

E-mail address: shqliu@hbu.edu.cn (Liu Shuaiqi).

Copyright © 2015 Scientific Advances Publishers

2010 Mathematics Subject Classification: 80.

Submitted by Jianqiang Gao.

Received January 26, 2015

sum-modified-Laplacian (SML) based on SSSID, finally, the fusion image is gained by inverse NSST. The algorithm can both preserve the information of the source images well and suppress pixel distortion due to nonlinear operations in transform domain. Experimental results demonstrate that the proposed method outperforms the state-of-the-art transform domain and PCNN fusion methods.

Keywords: image fusion, NSST, self-similarity, depth information, PCNN.

1. Introduction

Image fusion technique aims at generating a clear and acceptable to human vision panoramic image by fusing comprehensive and redundant information from a set of source images of the same scene [1]. The source images can not reflect the panoramic scene from limited focal length of the camera. So fused image will be more suitable for machine processing and human perception than any individual source image [2]. Generally, multi-focus image fusion can be divided into three levels: pixel level fusion, feature level fusion, and decision level fusion [3], and in our paper, we focus on pixel-level fusion technique. Image fusion method at pixel-level can be divided into two categories: spatial domain algorithms and transform domain algorithms. The spatial domain algorithms mainly include principal component analysis [4, 5], guided filtering based method [6] and so on. The transform domain algorithms are mainly based on multiresolution geometric analysis (MGA) tool domain and they can be divided into local energy function (such as image fusion algorithm based on wavelet [7, 8], ripplelet [9], contourlet [10, 11], shearlet [13], surfacelet [15], trained dictionaries [16, 17]) and artificial neural network (ANN) (such as contourlet with PCNN [12], shearlet with PCNN [14]) by different fused rules.

MGA tools mainly contain discrete wavelet transform (DWT), contourlet transform, non-subsampled contourlet transform (NSCT) [12], and shearlet transform (ST) [14]. They are all successfully applied to image fusion. Contourlet transform is computationally simple and can represent images sparsely, it has been widely applied to image fusion

algorithms, such as in [10-12, 19]. However, contourlet transform is translational variability [19]. NSCT constructed by non-subsampled Laplacian transform and non-subsampled directional filter can overcome this disadvantage, and NSCT-based image fusion method arrives satisfactory fusion performance. However, the high computational complexity of NSCT cause it difficult to be applied to real time applications. Unlike to NSCT, ST [20-23] has a rigorous mathematical framework and smaller support sizes of shear filter can accelerate the computation speed [23-25]. But ST also lack shift-variance and this will lead some artifacts when fused by using ST. NSST is proposed to overcome the disadvantages of NSCT an ST [20]. Compared with current MGA tools, NSST is not only shift-invariance, but also has better sparse representation ability and much faster computational speed. So, image fusion based on NSST can achieve better fused performance.

In addition, a good fusion method not only relies on the transform but also depend on fused rules. The region fusion rules combine the feature level fusion and the pixel level fusion together, which can achieve better fusion effect than others. So, now the region-based fusion rules has been widely studied, but due to most MGA tools fusion methods, region-based fusion rules just be applied to high frequency, ignore the low frequency fused rules. And the shape of a region will seriously affect fused images. For example, the commonly used isotropic square regions in the spatial domain will easily lead to blocky artifacts in blurred areas around the edge [18]. Guo et al. proposed structure-driven regions by using the self-similarity of a single image and the shared similarity among multiple source images in [18]. And they also proposed a new method which can refine the image clarity by automatically estimating depth information of blurred images. So, in our paper, instead the simple fused rules, structure-driven region fused rule called SSSID is applied to fuse the low frequency coefficients. However, MGA tools fusion methods also lead to pixel distortion due to nonlinear operations in the transform domain [18]. On the contrary, much less distortions are introduced if linear fusion

rules in the spatial domain or ANN [18]. PCNN is a simplified model which is suited to realizing synchronous pulse excitation behaviour. This model has the characteristics of spatial proximity and similar characteristics aggregation, which has been efficiently applied to image processing in applications, such as image segmentation, image fusion, etc. However, the value of single pixel in spatial or transform domain is used to motivate one neuron of PCNN. In fact, humans are often sensitive to edges, directional features, etc. But, a simple use of single pixels is not enough. So, PCNN is applied to fuse the high frequency coefficients motivated by larger SML with SSSID.

This paper is organized as follows. In Section 2, we give a brief review of NSST theory. In Section 3, we introduce the working principle of SSSID. In Section 4, we give a brief review of PCNN. In Section 5, we give the steps of the new image fusion algorithm based on NSST with PCNN and SSSID. In Section 6, we demonstrate the experimental results of the proposed and the comparisons with other typical fusion methods. And in the last section, we explore some conclusions.

2. Non-Subsampled Shearlet Transform

Shearlet Transform (ST) is a new multi-scale geometric analysis algorithm which inherits advantages of contourlet and curvelet transform [23]. ST theory is based on composite wavelets. In dimension $n = 2$, the affine systems with composite dilations are the collections of the form [20, 21]:

$$\Omega_{\mathbf{AB}}(\psi) = \left\{ \begin{array}{l} \psi_{j,l,k}(x) = |\det \mathbf{A}|^{j/2} \psi(\mathbf{B}^l \mathbf{A}^j x - k) \\ : j, l \in \mathbb{Z}, k \in \mathbb{Z}^2 \end{array} \right\}, \quad (1)$$

where $\psi \in L^2(\mathbb{R}^2)$, \mathbf{A} , \mathbf{B} are 2×2 invertible matrices and $|\det \mathbf{B}| = 1$.

When for any $f \in \mathbb{R}^2$, $\Omega_{\mathbf{AB}}(\psi)$ forms a Parseval frames (also called tight frame), we call the elements of $\Omega_{\mathbf{AB}}(\psi)$ composite wavelet. And the

dilations A^j are associated with scale transformations, while the matrices B^l are associated to area-preserving geometrical transform, such as rotations and shear [20-22]. Normally, $A = A_0 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$ is the anisotropic dilation matrix, and $B = B_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is shear matrix in (1), then, we can construct a tiling of the frequency like Figure 1.

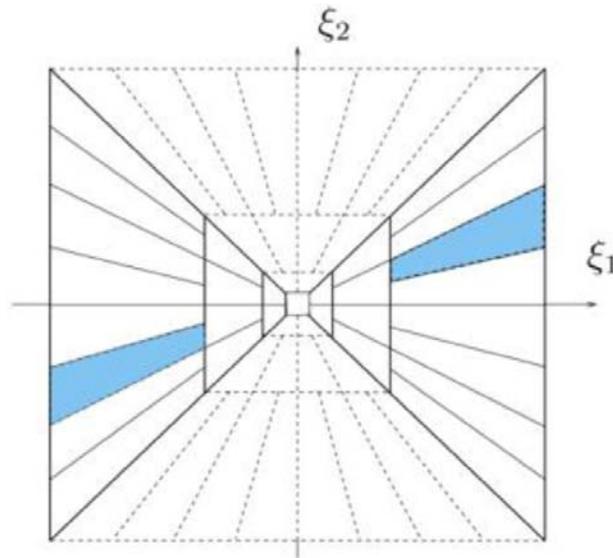


Figure 1. The tiling of the frequency plane induced by the ST.

Let $\psi_{j,l,k}^{(\theta)}(x) = 2^{\frac{3j}{2}} \psi^{(\theta)}(B_1^l A_1^j x - k)$. That is, each element $\hat{\psi}_{j,l,k}$ is supported on a pair of trapezoids, of approximate size $2^{2j} \times 2^j$, oriented along lines of slope $l2^{-j}$, see Figure 2.

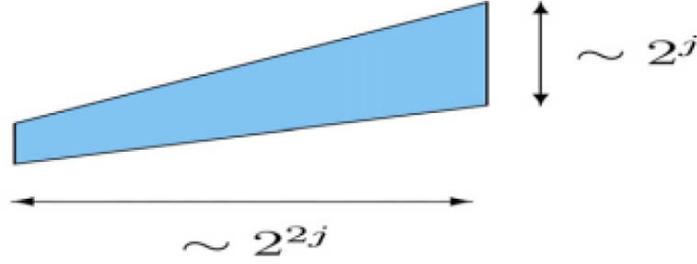


Figure 2. The frequency support of ST.

If $f \in L^2(\mathbb{R}^2)$, then its continuous ST is defined as

$$SH_{\psi} = \langle f, \psi_{j,l,k}^{(\theta)} \rangle, \quad (2)$$

where $j \geq 0$, $l = -2^j, 2^j - 1$, $k \in \mathbb{Z}^2$, $\theta = 0, 1$.

The discretization process of NSST [21] is composed of two phases: multi-scale factorization and multi-orientation factorization. Non-subsampled pyramid is utilized to complete multi-scale factorization, which can produce $(k + 1)$ sub-images which consist of one low-frequency image and k high-frequency images whose sizes are all the same as the source image, where k denotes the number of decomposition levels. The multi-orientation factorization in NSST is realized via improved shearing filters. For shearing filter used in ST, NSST makes the standard shearing filter map from pseudo-polar coordinate systems into Cartesian system to satisfy the property of shift-invariance [21].

3. Shared Similarity Regions and Depth Information

Adaptive fusion regions by using image similarity information are proposed in [18]. Depth information also is utilized to refine spatial continuity. In this section, we give reviews SSSID.

First, shared similarity regions are defined as follows. The adaptive region is generated by using the shared similarity of source images. Images are first divided into multiple overlapped square patches, and similar patches are searched [18, 26]. Given a reference patch $P_r \in R^{m \times m}$ and a region $W(\mathbf{r}) \in R^{m \times m}$ centered at pixel \mathbf{r} , the similarity of any candidate patch $P_q \in R^{m \times m}$ to the P_r is defined as

$$\eta_q = \|P_q - P_r\|_F, \quad (3)$$

where $\|A_M\|_F = \sqrt{\sum_{i=1}^I \sum_{j=1}^J |a_{ij}|}$ denotes the Frobenius norm of matrix

$A_M = (a_{ij})$. By sorting the η_q by the descending order for all the patches in this region, the most k similar patches to P_r are found and the collection of this patches are expressed as $L_W(\mathbf{r}) = \{P_{q_1}, P_{q_2}, \dots, P_{q_k}\}$.

Similar patches shared by both $L_{W^A}(\mathbf{r})$ and $L_{W^B}(\mathbf{r})$ are

$$L_W^s(\mathbf{r}) = L_{W^A}(\mathbf{r}) \cap L_{W^B}(\mathbf{r}), \quad (4)$$

where $L_W^s(\mathbf{r})$ is one of adaptive regions for fusion, $W^A(\mathbf{r})$ and $W^B(\mathbf{r})$ denote the same region of source images f^A and f^B , respectively [18]. The adaptive region is composed of 4 similar patches as shown in Figure 3.

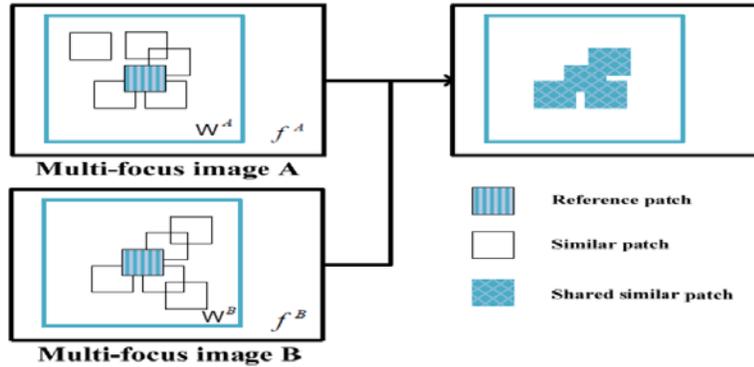


Figure 3. An adaptive region of shared similar patches.

Then, depth information is introduced to refine spatial continuity in [18]. $d(\mathbf{r})$ denotes the distance between the focal plane and the object whose location corresponds to the image pixel \mathbf{r} . Rays from a point of another object at $d(\mathbf{r})$ will reach multiple sensor points and result in a blurred image. Let $s(\mathbf{r})$ denote the diameter of the circle of confusion. In [27], a defocus map estimation method is automatically estimating by the assumption of the linear response camera $d(\mathbf{r}) \propto s(\mathbf{r})$. The whole estimation process is shown in [18, 27]. In the follow of our paper, we will use the same fusion rules method as in [18] to fuse the low-frequency coefficients of NSST, and use SSSID and PCNN to fuse the high-frequency coefficients of NSST.

4. Pulse Coupled Neural Network

PCNN is a feedback network which consists of many PCNN neurons. Each PCNN neuron includes three parts: the receptive field, modulation field, and pulse generator. When the PCNN is used in digital image processing, each pixel corresponds to a PCNN neuron. This paper applies a simplified PCNN model and discrete form [14]:

$$\left\{ \begin{array}{l} F_{i,j}^{l,k}(n) = I_{i,j}^{l,k}(n), \\ L_{i,j}^{l,k}(n) = e^{-\alpha L} L_{i,j}^{l,k}(n-1) + V_L \sum_{p,q} W_{i,j,pq}^{l,k} Y_{i,j,pq}^{l,k}(n-1), \\ U_{i,j}^{l,k}(n) = F_{i,j}^{l,k}(n) \cdot (1 + \beta_{i,j}^{l,k}(n) \cdot L_{i,j}^{l,k}(n)), \\ \Theta_{i,j}^{l,k}(n) = e^{-\alpha \Theta} \Theta_{i,j}^{l,k}(n-1) + V_{\Theta} Y_{i,j}^{l,k}(n), \\ Y_{i,j}^{l,k}(n) = \begin{cases} 1 & U_{i,j}^{l,k}(n) > \Theta_{i,j}^{l,k}(n), \\ 0 & \text{otherwise,} \end{cases} \end{array} \right. \quad (5)$$

where n denotes the current iterations. $F_{i,j}^{l,k}$ and $L_{i,j}^{l,k}$ denote the feeding input and linking input, respectively. $I_{i,j}^{l,k}$ is the external stimulation input (this is the NSST coefficient located at the pixel (i, j) in l -th scale

and k -th direction sub-band). $\beta_{i,j}^{l,k}$ is the linking strength. $U_{i,j}^{l,k}$ is the internal activity of neuron, and $\Theta_{i,j}^{l,k}$ is the dynamic threshold. $W_{i,j,pq}^{l,k}$ is the synaptic gain strength and subscripts p and q are the size of linking range. V_L is the amplitude gain of the linking field and V_Θ is the amplitude gain of the threshold adjuster. α_L and α_Θ are the time constant of the linking field and the threshold adjuster, respectively. $Y_{i,j}^{l,k}$ stands for the pulse output of neuron. If $U_{i,j}^{l,k} > \Theta_{i,j}^{l,k}$, then the neuron generates a pulse $Y_{i,j}^{l,k}(n) = 1$, also called one firing time. After n iterations, the firing frequency map consisted of the total firing times of PCNN neuron can be regarded as the output of PCNN. Due to the high frequency coefficients of NSST represent the detail components of image, such as edge, line, region boundary, etc. So, PCNN-based fused rule with SSSID is applied to the high frequency coefficients processing.

5. Image Fusion Based on SSSID and PCNN

NSST owned translational invariant which can suppress the pseudo Gibbs effect well without increasing computational complexity is utilized to decompose images. Low-frequency coefficients of NSST can reserve the main spatial information of the source images, so we can directly use SSSID to get the low-frequency fused image. SSSID (different as [18], the depth information should be measured by the source images) is a better representation of the regional energy function of the image edge details, and has a good effect in motivating by SSSID. Therefore, PCNN with SSSID is utilized to smooth spatial continuity of fused image. The proposed image fusion algorithm is described as follows. Without losing of generality, A and B are two images with different focuses to fuse, Fr is fused image. Obviously, it is easy to promote it into multiple image fusion.

First, we utilize NSST to decompose image A and B , with decomposition coefficients $C_A^{l,k}(\mathbf{r})$ and $C_B^{l,k}(\mathbf{r})$. Generally, the coefficients of NSST are $C^{l,k}(\mathbf{r})$, where l and k denote the scale and direction of decomposition, respectively. When l is zero, they denote low-frequency coefficients, and otherwise they denote high-frequency coefficients. \mathbf{r} denotes the position of pixel, so the SML at the pixel \mathbf{r} is defined as follows [6]:

$$SML^{l,\theta}(\mathbf{r}) = \sum_{i \in W(\mathbf{r})} [ML^{l,\theta}(i)]^2. \quad (6)$$

Let $\mathbf{i} = (x, y)$, then

$$\begin{aligned} ML^{l,d}(x, y) = & |2C^{l,\theta}(x, y) - C^{l,\theta}(x - step, y) - C^{l,\theta}(x + step, y)| \\ & + |2C^{l,\theta}(x, y) - C^{l,\theta}(x, y - step) - C^{l,\theta}(x, y + step)|, \end{aligned} \quad (7)$$

where $step$ denotes a variable spacing between pixels. Generally, $step$ equals to 1. Thus, the depth information of original sources images can be used as a refinement for the clarity metric on low-frequency coefficients of NSST.

$$CM(\mathbf{r}) = SML^{l,\theta}(\mathbf{r}) \times [d(\mathbf{r})]^{-2}, \quad (8)$$

CM denotes the clarity metric and f_{Low}^F is the fused image of low-frequency coefficients, and clarity is measured in each adaptive region. If the clarity in an adaptive region $L_w^s(\mathbf{r})$ of f_{Low}^A is larger than that of f_{Low}^B , one vote will be assigned to all the pixels $\mathbf{j} \in L_w^s(\mathbf{r})$ in this adaptive region of f_{Low}^A , meaning

$$\nu^A(\mathbf{j}) = \nu^A(\mathbf{j}) + 1, \quad \mathbf{j} \in L_w^s(\mathbf{r}), \quad (9)$$

where $\nu^A(\mathbf{j})$ denotes the counter at the spatial location \mathbf{j} . The initial values of $\nu^A(\mathbf{j})$ and $\nu^B(\mathbf{j})$ are 0 and voting stops until all the adaptive

regions are compared. Finally, one can have counter maps $\nu^A(\mathbf{j})$ and $\nu^B(\mathbf{j})$ that are in the same size as source images [18]. In the following, $\nu^A(\mathbf{r})$ and $\nu^B(\mathbf{r})$ are used to denote the final vote for pixels at location \mathbf{r} . At last, use weightings to compose the pixel according to [18].

$$f_{Low}^F(\mathbf{r}) = \frac{\nu^A(\mathbf{r})}{\nu^A(\mathbf{r}) + \nu^B(\mathbf{r})} f_{Low}^A(\mathbf{r}) + \frac{\nu^B(\mathbf{r})}{\nu^A(\mathbf{r}) + \nu^B(\mathbf{r})} f_{Low}^B(\mathbf{r}). \quad (10)$$

Then, for the high-frequency decomposition coefficients $C_A^{l,\theta}(\mathbf{r})$ and $C_B^{l,\theta}(\mathbf{r})$ (l is greater than zero), we apply PCNN with SSSID fusion rule to fuse them. f_{High}^F is the fused image of high-frequency coefficients. PCNN fired by SSSID is applied to fuse the high-frequency coefficients. That means SSSID is the input of the PCNN and makes use of the global coupling and pulse synchronization to select the high frequency coefficients of the fused image. That is, in (5), let $F(\mathbf{r}) = CM(\mathbf{r})$ ($\mathbf{r} = (i, j)$). Then according to (5), compute the total firing times of each high frequency coefficients $T_{\mathbf{r}}^{l,k}(n) = T_{\mathbf{r}}^{l,k}(n-1) + Y_{\mathbf{r}}^{l,k}(n)$. After N iterations, the fused high frequency coefficients can be determined by (11).

$$C_A^{l,\theta}(\mathbf{r}) = \begin{cases} 1, & \text{if } T_{\mathbf{r}}^{l,k}(\mathbf{r}) \geq T_{\mathbf{r}}^{l,k}(\mathbf{r}), \\ 0, & \text{if } T_{\mathbf{r}}^{l,k}(\mathbf{r}) < T_{\mathbf{r}}^{l,k}(\mathbf{r}). \end{cases} \quad (11)$$

Finally, the fused low frequency coefficients $C_A^{l,\theta}(\mathbf{r})$ and the fused high frequency coefficients $C_A^{l,\theta}(\mathbf{r})$ are utilized for reconstructing the fused image F_r through inverse NSST. In conclusion, the framework of the proposed fusion algorithm is shown in Figure 4.

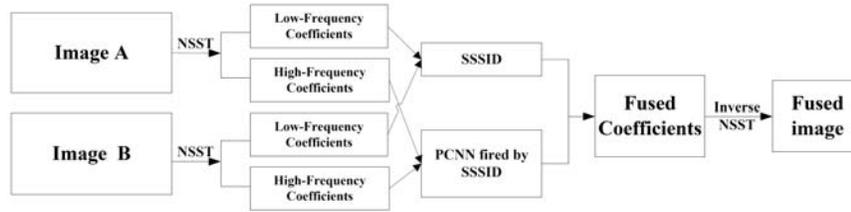


Figure 4. Framework of image fusion algorithm based on NSST-PCNN-SSSID.

6. Experimental Results

In order to evaluate the performance of the proposed fusion method, we use some pair multi-focus images to test it from visual appearance and objective criteria. Fusion method based on MGA tools, such as multi-focus image fusion method of ripplelet transform based on cycle spinning [9], the fusion is also performed using the fusion method of choosing larger SML based on contourlet domain proposed in [11], the fusion method based on the combination of NSCT and PCNN proposed in [12], the fusion method based on the combination of ST and PCNN proposed in [14] are the comparison methods. $Q^{AB/F}$ metric [11], mutual information (MI) [11] are employed as objective criteria. $Q^{AB/F}$ measures the amount of edge information transferred from the source images to the fused images. MI essentially computes how much information from source images is transferred to the fused image. The better fusion performance a method has the larger two index values. In the experiment, default parameters in the shared source codes are used. The parameters of PCNN in our method are the same to [12]. Experiments are implemented on an Intel I5 3.4GHz with 4GB RAM. The simulation software is Matlab 2014b. The group of source images shown in Figure 5 can be downloaded at website <http://www.imagefusion.org>.

Figure 5(a) and 5(f) are images of clock with focus on the right and on the left in typical multi-focus images for image fusion test. Figure 5(b)

and 5(g) are flowers images with different focus, Figure 5(c) and 5(h) are disk images with different focus, Figure 5(d) and 5(i) are lab images with different focus, Figure 5(e) and 5(j) are pepsi images with different focus.

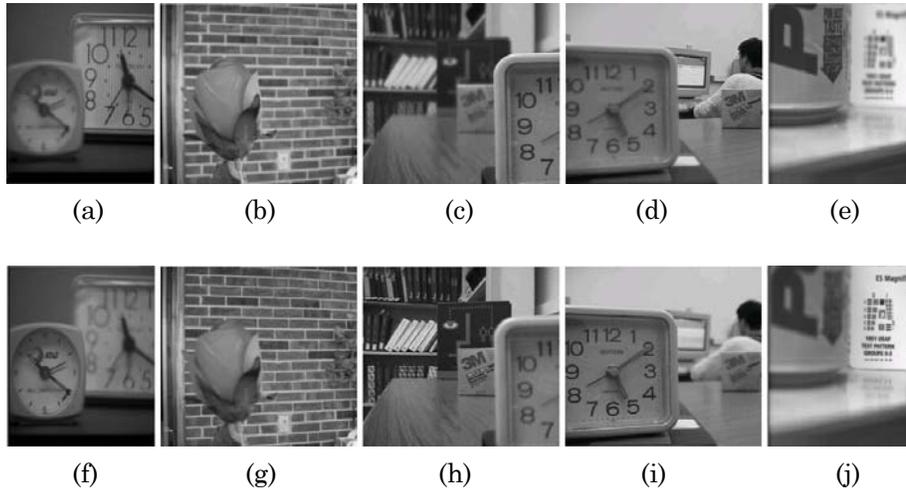


Figure 5. Images with different focus, Figure 5 (a, f), Figure 5 (b, g), Figure 5 (c, h), Figure 5 (d, i) and Figure 5 (e, j) are clock, flowers, disk, lab, and pepsi images with different focus.

We utilize the methods proposed in [9], [11], [12], [14] and the proposed to fuse, respectively. The fused images are shown in Figure 6(a)-(y). Comparing the top left corner and the area around number 8 of the large clock in fused image, we can see the proposed method uses image spatial continuity to avoid introducing some artificial texture into the fused image which exists in [9], [11], [12], and [14]. The other pairs images fusion is the same.



Figure 6. The performance of each fusion method. Figure (a – e), (f – j), (h – o), (p – t), and (u – y) shows the performance of each fusion method on clock, flower, disk, lab, and pepsi.

Overall, all the tests in Figure 6 show 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. As well, this paper also uses the objective criteria “MI” and “ $Q^{AB/F}$ ” to investigate the performance of different transform methods. As shown in Table 1, from which we can see, the proposed algorithm is the highest in the performance of objective criteria. All the above show that the proposed algorithm is very suitable for fairly complex aerial images of different focus to fuse.

Table 1. Objective criteria of different fused methods to Figure 6

The fusion methods	Clock		Flower		Disk		Lab		Pepsi	
	Q ^{AB/F}	MI								
Method in [9]	0.6721	6.6882	0.7072	5.4496	0.7054	6.3327	0.7042	7.0096	0.6477	6.9228
Method in [11]	0.6783	6.7086	0.7250	7.2740	0.7187	7.4027	0.7188	7.4027	0.7856	7.3761
Method in [12]	0.6775	7.4948	0.7137	6.3481	0.7225	6.7524	0.7212	7.8581	0.7774	7.0151
Method in [13]	0.6670	6.7204	0.6522	5.9164	0.7023	6.6419	0.6905	6.9061	0.7326	7.8201
Method in [14]	0.6900	6.7927	0.7373	6.1444	0.7288	7.1365	0.7119	7.1654	0.7537	7.4814
Our method	0.7092	8.0755	0.7439	7.9870	0.7366	7.6685	0.7509	8.1985	0.7905	7.8042

7. Conclusion

This paper presents an image fusion algorithm based on NSST and PCNN. Based on the traditional image fusion algorithm via frequency domain, the algorithm uses structure-driven regions and depth information to fuse the low-frequency coefficients of NSST. And PCNN fired by SSSID be applied to the high-frequency coefficients, which can enhance the spatial characteristics in the process of image fusion. Experimental results demonstrate that the proposed method is better than or close to the current popular image fusion algorithms based on MGA tools in terms of both visual appearance and objective criteria.

Acknowledgements

The authors are grateful to Dr. Qu for sharing the SSSID code. This work was supported in part by Natural Science Foundation of China under grant 61401308, Natural Science Foundation of Hebei Province under grant 2013210094, and Natural Science Foundation of Hebei University under grant 2014-303.

References

- [1] A. A. Goshtasby and S. Nikolov, Image fusion: Advances in the state of the art, *Inf. Fusion* 8(2) (2007), 114-118.
- [2] Y. Liu and Z. Wang, Simultaneous image fusion and denoising with adaptive sparse representation, *IET Image Processing*, 2014.
Online, DOI: 10.1049/iet-ipr.2014.0311
- [3] W. Huang and Z. Jing, Evaluation of focus measures in multi-focus image fusion, *Pattern Recognition Letters* 28(4) (2007), 493-500.
- [4] Y. H. Jia, Fusion of Landsat TM and SAR images based on principal component analysis, *Remote Sensing Technology and Application* 13(1) (1998), 46-49.
- [5] T. Wan, C. C. Zhu and Z. C. Qin, Multifocus image fusion based on robust principal component analysis, *Pattern Recognition Letters* 34(9) (2013), 1001-1008.
- [6] S. T. Li, X. D. Kang and J. W. Hu, Image fusion with guided filtering, *IEEE Transactions on Image Processing* 22(7) (2013), 2864-2875.

- [7] G. Pajares and J. M. Cruz, A wavelet-based image fusion tutorial, *Pattern Recognition* 37(9) (2004), 1855-1872.
- [8] Y. Chai, H. Li and Z. Li, Multifocus image fusion scheme using focused region detection and multiresolution, *Optics Communications* 284(19) (2011), 4376-4389.
- [9] P. Geng, M. Huang and S. Q. Liu et al., Multifocus image fusion method of ripplelet transform based on cycle spinning, *Multimedia Tools and Applications* (2014), 1-11.
- [10] S. Q. Liu, J. Zhao and P. Geng et al., Medical image fusion based on nonsubsampled direction complex wavelet transform, *International Journal of Applied Mathematics and Machine Learning* 1(1) (2014), 21-34.
- [11] X. B. Qu, J. W. Yan and G. D. Yang, Sum-modified-Laplacian-based multi-focus image fusion method in sharp frequency localized contourlet transform domain, *Optics and Processing Engineering* 17(5) (2009), 1203-1212.
- [12] X. B. Qu, J. W. Yan and H. Z. Xiao et al., Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain, *Acta Automatica Sinica* 34(12) (2008), 1508-1514.
- [13] Q. G. Miao, C. Shi and P. F. Xu, A novel algorithm of image fusion using shearlets, *Optics Communications* 284(6) (2011), 1540-1547.
- [14] P. Geng, Z. Wang and Z. Zhang et al., Image fusion by pulse couple neural network with shearlet, *Optical Engineering* 51(6) (2012), 067005-1-067005-7.
- [15] Q. Zhang, L. Wang and Z. Ma et al., A novel video fusion framework using surfacelet transform, *Optics Communications* 285(13) (2012), 3032-3041.
- [16] L. Chen, J. Li and C. L. Chen, Regional multifocus image fusion using sparse representation, *Optics Express* 21(4) (2013), 5182-5197.
- [17] S. B. Gao, Y. M. Cheng and Y. Q. Zhao, Method of visual and infrared fusion for moving object detection, *Optics Letters* 38(11) (2013), 1981-1983.
- [18] D. Guo, J. Yan and X. Qu, High quality multi-focus image fusion using self-similarity and depth information, *Optics Communications* 338(1) (2015), 138-144.
- [19] S. Q. Liu, S. H. Hu and Y. Xiao, SAR image de-noised based on wavelet-contourlet transform with cycle spinning, *Signal Processing* 27(6) (2011), 837-842. (Chinese)
- [20] G. Easley, D. Labate and W. Q. Lim, Sparse directional image representation using the discrete shearlets transform, *Applied and Computational Harmonic Analysis* 25(1) (2008), 25-46.
- [21] G. Kutyniok, J. Lemvig and W. Q. Lim, Compactly supported shearlets are optimally sparse, *Journal of Approximation Theory* 163(11) (2011), 1564-1589.
- [22] W. Q. Lim, The discrete shearlets transform: A new directional transform and compactly supported shearlets frames, *IEEE Trans. Image Proc.* 19(5) (2010), 1166-1180.
- [23] S. Liu, S. Hu and Y. Xiao, Image separation using wavelet-complex shearlet dictionary, *Journal of Systems Engineering and Electronics* 25(2) (2014), 314-321.

- [24] W. Kong, L. Zhang and Y. Lei, Novel fusion method for visible light and infrared images based on NSST-SF-PCNN, *Infrared Physics & Technology* 65 (2014), 103-112.
- [25] S. Q. Liu, S. H. Hu and Y. Xiao et al., Bayesian shearlet shrinkage for SAR image denoising via sparse representation, *Multidimensional Systems and Signal Processing* 25(4) (2014), 683-701.
- [26] X. B. Qu, Y. Hou and F. Lam et al., Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator, *Medical Image Analysis* 18(6) (2014), 843-856.
- [27] S. Zhuo and T. Sim, Defocus map estimation from a single image, *Pattern Recognition* 44(9) (2011), 1852-1858.

