

## Research Article

# Image Fusion Based on Nonsampled Contourlet Transform and Saliency-Motivated Pulse Coupled Neural Networks

**Liang Xu, Junping Du, and Qingping Li**

*Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China*

Correspondence should be addressed to Junping Du; [junpingdu@126.com](mailto:junpingdu@126.com)

Received 26 March 2013; Accepted 26 May 2013

Academic Editor: H. K. Leung

Copyright © 2013 Liang Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the nonsampled contourlet transform (NSCT) domain, a novel image fusion algorithm based on the visual attention model and pulse coupled neural networks (PCNNs) is proposed. For the fusion of high-pass subbands in NSCT domain, a saliency-motivated PCNN model is proposed. The main idea is that high-pass subband coefficients are combined with their visual saliency maps as input to motivate PCNN. Coefficients with large firing times are employed as the fused high-pass subband coefficients. Low-pass subband coefficients are merged to develop a weighted fusion rule based on firing times of PCNN. The fused image contains abundant detailed contents from source images and preserves effectively the saliency structure while enhancing the image contrast. The algorithm can preserve the completeness and the sharpness of object regions. The fused image is more natural and can satisfy the requirement of human visual system (HVS). Experiments demonstrate that the proposed algorithm yields better performance.

## 1. Introduction

Due to a tremendous growth in the application of image sensors, image fusion technique has huge potential for growth and has been used successfully in many fields, such as remote sensing, medical imaging, defense surveillance, and computer vision [1–3]. The aim of image fusion is to combine several source images (obtained from different sensors and view points) into a fused image, which contains all important contents from source images and expresses more abundant information in a scene. In practical applications, the direct obtained image is not able to satisfy the requirements because of many factors, for example, the limitations of sensors, varying illumination, occlusions and angles, and so forth. Image fusion technique can solve effectively the problems by taking advantage of multiple-source information producing the fused result which satisfies perception system.

According to the level, image fusion approaches can be generally classified into three types: pixel level, feature level, or decision level [4]. According to whether the fusion methods need the assistant of multiscale transform (MST) tools or not, they can also be categorized into two main

classes [5]: MST-based and non-MST-based approaches. A variety of MST tools have been developed for image fusion. The earliest and the most popular MST tools are pyramid [6, 7] and wavelet [8, 9] transform. They are directly constructed by combination of two 1D transforms, so they are not the true 2D transforms. To improve the accuracy of decomposition and reconstruction, the more advanced MST tools have been proposed, such as ridgelets [10], contourlets [11, 12], curvelets [13, 14], and NSCT [15]. The approaches are the true 2D geometric MST tools, which can achieve the decomposition and reconstruction of image signals perfectly and satisfy the requirement of image fusion. In addition, for non-MST-based methods, Piella [16] performs the image fusion by a variational model, and the fused result contains the geometry structure of all the inputs and enhances the contrast for visualization. Ludusan and Lavielle [17] propose a variational approach based on error estimation theory and partial differential equations for concurrent image fusion and denoising of multifocus images.

The combination strategy of the decomposed coefficients is another key step in the MST-based fusion approaches. Fusion strategies can mainly be divided into three categories:

pixel-based, window based, and area based [18]. The simplest pixel-based fusion rule directly selects the fused coefficients using single pixel, but the method is easily influenced by noise. Window based and area based fusion rules take advantage of the local characteristics of neighborhood pixels and, thus, are superior to pixel-based rules [19].

The existing image fusion approaches do not take fully into account the characteristics of HVS, which the HVS tends to focus on the most relevant saliency regions in a scene. According to the visual perception mechanism, the fused image should improve the quality of object areas in a scene. The goal of the proposed algorithm is to preserve the completeness, saliency, and sharpness of object areas and satisfy the requirements of HVS. Consequently, based on NSCT and saliency-motivated PCNN, the paper proposes a novel image fusion algorithm. The visual saliency model and PCNN are two very important tools in image processing. The former is inspired by the behavior and the neuronal architecture of the early primate visual system; the latter is a visual cortex-inspired neural network and characterized by the global coupling and pulse synchronization of neurons. The saliency map produced by the visual saliency model as input to motivate PCNN is used as the fusion rule which can preserve the saliency objects from source images leading to more abundant content contained in a fused image.

The rest of the paper is organized as follows. Section 2 reviews basic NSCT theory in brief. Section 3 presents the proposed image fusion algorithm in detail. Section 4 demonstrates and discusses the experimental results. Section 5 concludes.

## 2. Nonsampled Contourlet Transform

In this section, we briefly review the theory and properties of NSCT, which will be used in the rest of this paper (see [15] for details).

NSCT is a kind of overcomplete transform and is a shift-invariant version of contourlet transform. NSCT has some excellent properties in the process of image decomposition, including shift invariance, multiscale, and multidirection. NSCT is used as the MST tool to provide a better representation of the contours and overcome pseudo-Gibbs phenomena. The main components of the NSCT are a nonsampled pyramid filter bank (NSPFB) structure for multiscale decomposition and a nonsampled directional filter bank (NSDFB) structure for directional decomposition. The NSCT is displayed in Figure 1.

The multiscale property of the NSCT is achieved by using two-channel nonsampled 2-D filter banks (NSFBs), called as NSPFB. The filters for next level are obtained by upsampling the filters of the previous level, by which the multiscale property is obtained without the need for additional filter design. We assume that the NSPFB decomposition is with  $J = N$  levels. At the first level, input images are decomposed by the low-pass filter  $H_0(z)$  and the corresponding high-pass filter  $H_1(z)$ , respectively. The ideal passband support of the low-pass filter at the  $j$ th level is the region  $[-(\pi/2^j), (\pi/2^j)]^2$ . The ideal support of the equivalent high-pass filter is

the complement of the low-pass filter, that is, the region  $[-(\pi/2^{j-1}), (\pi/2^{j-1})]^2 \setminus [-(\pi/2^j), (\pi/2^j)]^2$ . The NSDFB, a shift-invariant directional filter bank (DFB), is obtained by eliminating the downsamplers and upsamplers in the DFB. To achieve multidirection decomposition, the NSDFB is iteratively used. All filter banks in the NSDFB tree structure are obtained from a single NSFB with fan filters. Each filter bank in the NSDFB tree has the same computational complexity as that of the building-block NSFB.

Figure 1 shows the NSCT which is constructed by combining the NSPFB and the NSDFB. The two-channel NSFBs in the NSPFB and the NSDFB satisfy the Bezout identity and are invertible, so the NSCT is invertible. The key of NSCT is the filter design problem of the NSPFB and NSDFB. The aim is to design the filters supporting the Bezout identity and obtaining other useful properties. In addition, for a fast implementation, the mapping approach is used to transform the filter into a ladder or lifting structure. More details can be seen in [15].

## 3. The Proposed Algorithm

In the section, the proposed image fusion algorithm based on NSCT and saliency-motivated PCNN is presented in detail. The main idea is that the visual saliency map is first built on high-pass subband coefficients of the NSCT using the visual attention model (phase spectrum of Fourier transform (PFT) model presented in Section 3.1) and then is combined with source high-pass subband coefficients as input to motivate PCNN. Coefficients with large firing times are employed as the fused high-pass subband coefficients. Low-pass subband coefficients are merged to develop a weighted fusion rule based on firing times of PCNN. PCNN is built in each subband to simulate the biological activity of HVS. The fused image has more natural visual appearance and can satisfy the requirements of HVS. The framework of the proposed algorithm is shown in Figure 2. For the clearness of the presentation, we assume that two registered source images are combined.

The algorithm first decomposes source images into the low-pass subband and high-pass directional subband coefficients by the NSCT. The coarsest subband contains the main energy from source images and denotes the abundant structural information. Therefore, an adaptive weighted average fusion rule based on the firing times of PCNN is developed to merge the low-pass subband. High-pass directional subbands contain the abundant detail contents of source images, so we create a maximum selection fusion principle based on saliency-motivated PCNN for selecting the fused coefficients. The final fused image is reconstructed by applying the inverse NSCT on the merged coefficients.

*3.1. Images Decomposition and Saliency-Motivated PCNN.* The decomposition of source images employs NSCT presented in Section 2. Input images  $A$  and  $B$  are decomposed into different scale and direction subbands using NSCT. The subbands  $\{C_{j_0}^A(x, y), C_{j,l}^A(x, y)\}$  and  $\{C_{j_0}^B(x, y), C_{j,l}^B(x, y)\}$  are

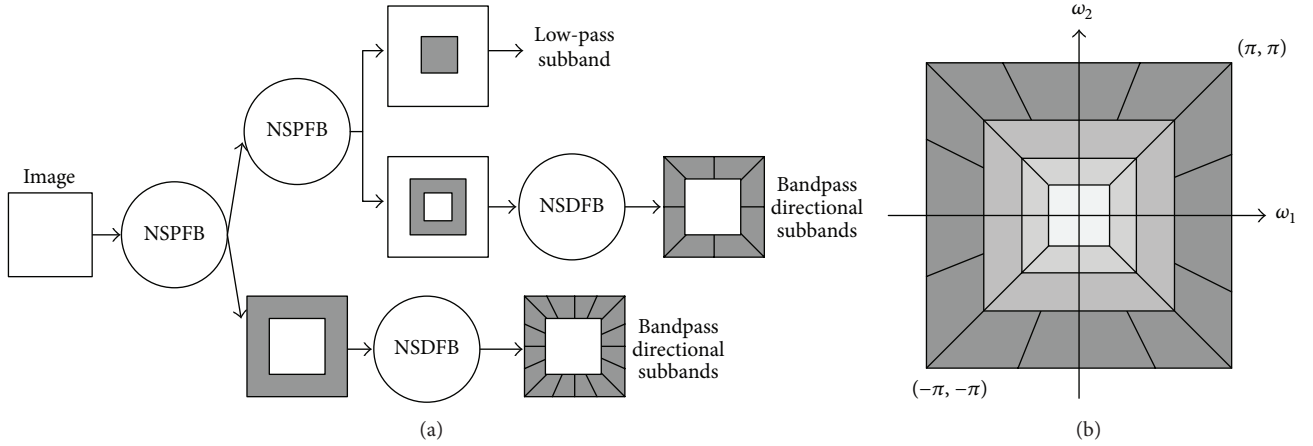


FIGURE 1: Nonsampled contourlet transform. (a) NSFB structure that implements the NSCT and (b) idealized frequency partitioning obtained with the proposed structure.

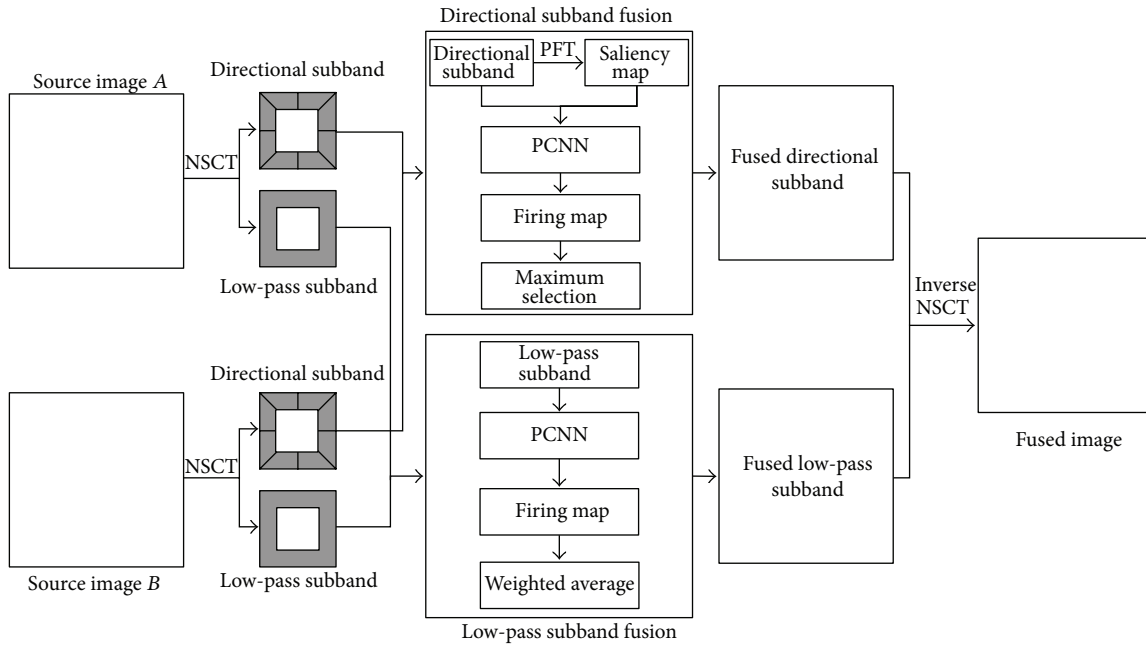


FIGURE 2: Architecture of the proposed algorithm.

obtained, where  $C_{j_0}(x, y)$  denotes the low-pass subband coefficients of the input images at the coarsest scale and  $C_{j,l}(x, y)$  denotes the high-pass directional subband coefficients at the  $j$ th scale and in the  $l$ th direction.

The following the proposed saliency-motivated PCNN model is discussed. Eckhorn develops a novel biological neural network, called PCNN which is based on the experimental observations of synchronous pulse bursts in cat and monkey visual cortices [20]. PCNN is a feedback network and each PCNN neuron consists of three parts: receptive field, modulation field, and pulse generator [21]. In image processing, PCNN is a single-layer and a two-dimensional connection neural network [22, 23] shown in Figure 3.

In this paper, let  $C_{j,l}(x, y)$  denotes the coefficient located at  $(x, y)$  in the  $j$ th scale at the  $l$ th direction.  $C_{j,l}(x, y)$  in each subband is inputted to PCNN to motivate the neurons and generate pulse of neurons with (1). Firing times  $T_{xy}^{j,l}$  are then computed as in (2):

$$\begin{aligned}
 F_{xy}^{j,l} [n] &= C_{j,l}(x, y) \\
 L_{xy}^{j,l} [n] &= \exp(-\alpha_L) L_{xy}^{j,l} [n-1] \\
 &\quad + V_L \sum_{pq} W_{xy,pq}^{j,l} Y_{xy,pq}^{j,l} [n-1] \\
 U_{xy}^{j,l} [n] &= F_{xy}^{j,l} [n] * (1 + \beta L_{xy}^{j,l} [n])
 \end{aligned}$$

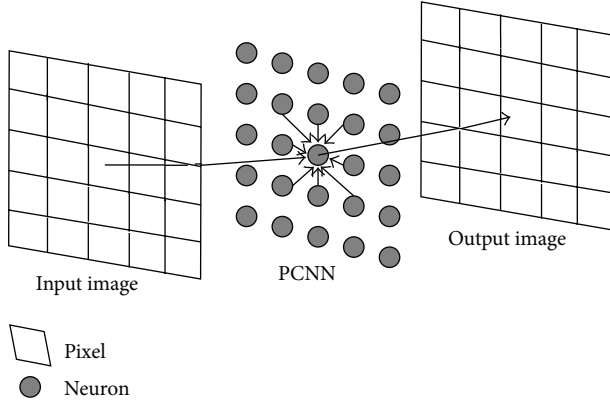


FIGURE 3: Connection model of PCNN neuron.

$$\theta_{xy}^{j,l} [n] = \exp(-\alpha_\theta) \theta_{xy}^{j,l} [n-1] + V_\theta Y_{xy}^{j,l} [n-1]$$

$$Y_{xy}^{j,l} [n] = \begin{cases} 1, & \text{if } U_{xy}^{j,l} [n] > \theta_{xy}^{j,l} [n] \\ 0, & \text{otherwise,} \end{cases}$$

$$T_{xy}^{j,l} [n] = T_{xy}^{j,l} [n-1] + Y_{xy}^{j,l} [n]. \quad (2)$$

In (1), the coefficient  $C_{j,l}(x, y)$  is assigned to the feeding input  $F_{xy}^{j,l}$ . The linking input  $L_{xy}^{j,l}$  is equal to the sum of neurons firing times in linking range, where  $\alpha_L$  indicates the decay constants and  $V_L$  is the amplitude gain.  $W_{xy,pq}$  is the weighted coefficient ( $p$  and  $q$  point out the size of linking range in PCNN). The internal state signal  $U_{xy}^{j,l}$  is obtained by modulating  $F_{xy}^{j,l}$  and  $L_{xy}^{j,l}$ , where  $\beta$  is the linking strength.  $\theta_{xy}^{j,l}$  is the threshold, where  $\alpha_\theta$  and  $V_\theta$  are the decay constants and the amplitude gain, respectively.  $n$  denotes the iteration times. If  $Y_{xy}^{j,l} = 1$ , the neuron will generate a pulse, called one firing. If  $Y_{xy}^{j,l} = 0$ , the neuron will not generate a pulse. In applications,  $T_{xy}^{j,l} [n]$  defined in (2) are often used to indicate the total firing times in  $n$  iteration. The firing times are employed to represent image information.

The saliency maps  $S_{j,l}^A(x, y)$  and  $S_{j,l}^B(x, y)$  are computed on the high-pass directional subbands  $C_{j,l}^A(x, y)$  and  $C_{j,l}^B(x, y)$ , which denotes the  $j$ th scale and  $l$ th direction. The saliency maps are used as the importance indicator of the coefficients for preserving important information of source images.

Phase spectrum of Fourier transform (PFT) proposed in [24] is employed as a saliency detection model for grayscale image. PFT showed that the saliency map can be easily computed by the phase spectrum of an image's Fourier transform when its amplitude spectrum is at nonzero constant value. Only the phase spectrum is used to reconstruct an image which reflects the saliency information of the source image. The implementation of PFT model consists of three steps. An image is first transformed into frequency domain using Fourier transform, and the amplitude and phase spectrums are then obtained. Finally, the saliency map is obtained by

inverse Fourier transform on only the phase spectrum. Given an input image  $I(x, y)$ , three steps have the corresponding equations as follows:

$$F(u, v) = F(I(x, y)) \quad (3)$$

$$P(u, v) = P(F(u, v)) \quad (4)$$

$$S(x, y) = g * \left\| F^{-1} \left\{ \exp^{i \cdot P(u,v)} \right\} \right\|^2, \quad (5)$$

where  $F$  and  $F^{-1}$  denote Fourier transform and inverse Fourier transform.  $P(F)$  is the phase spectrum of  $I$  and  $g$  is a 2D Gaussian filter. The saliency value in location  $(x, y)$  is computed using (5).

PFT model is a simple and efficient saliency detection method. An example of the PFT saliency detection is shown in Figure 4. Figures 4(a) and 4(c) are two multifocus source images which show complementary focus point regions. Figures 4(b) and 4(d) are the corresponding saliency maps which indicate different saliency regions of source images. We can observe that the saliency maps present focus point areas in source images.

Consequently, in this paper, the saliency value of high-pass subbands  $S_{j,l}^{A/B}(x, y)$  can be computed by (5) in which coefficient  $C_{j,l}^{A/B}(x, y)$  replaces  $I(x, y)$  as the input. Instead of using PCNN in NSCT domain directly, the product  $SC_{j,l}(x, y)$  between coefficient  $C_{j,l}^{A/B}(x, y)$  in location  $(x, y)$  of high-pass subbands and its saliency value  $S_{j,l}^{A/B}(x, y)$  are used to motivate PCNN.  $SC_{j,l}(x, y)$  is computed as follows:

$$SC_{j,l}(x, y) = S_{j,l}^{A/B}(x, y) \cdot C_{j,l}^{A/B}(x, y). \quad (6)$$

Then,  $SC_{j,l}(x, y)$  is normalized as  $SC\_Norm_{j,l}(x, y)$  which is inputted to PCNN to motivate neurons. The proposed saliency-motivated PCNN model is defined in (7) by modifying (1):

$$F_{xy}^{j,l} [n] = SC\_Norm_{j,l}(x, y)$$

$$L_{xy}^{j,l} [n] = \exp(-\alpha_L) L_{xy}^{j,l} [n-1] + V_L \sum_{pq} W_{xy,pq}^{j,l} Y_{xy,pq}^{j,l} [n-1]$$

$$U_{xy}^{j,l} [n] = F_{xy}^{j,l} [n] * (1 + \beta L_{xy}^{j,l} [n])$$

$$\theta_{xy}^{j,l} [n] = \exp(-\alpha_\theta) \theta_{xy}^{j,l} [n-1] + V_\theta Y_{xy}^{j,l} [n-1]$$

$$Y_{xy}^{j,l} [n] = \begin{cases} 1, & \text{if } U_{xy}^{j,l} [n] > \theta_{xy}^{j,l} [n] \\ 0, & \text{otherwise.} \end{cases}$$

**3.2. Subband Coefficients Fusion.** The high-pass subbands of NSCT decomposition contain abundant detailed information and indicate the saliency components of images, for example, lines, edges, contours, and so forth. In order to preserve the saliency components in the process of image fusion, we propose the fusion rule based on saliency-motivated



**Step 0:** Given source images  $A$  and  $B$ .  
**Step 1:** Perform decomposition on source images  $A$  and  $B$  using NSCT to obtain the high-pass directional subband coefficients and the low-pass subband coefficients.  
**Step 2:** Merge the high-pass directional subbands with (8), (7) and (2).  
**Step 3:** Obtain the fused low-pass subband with (9), (1) and (2).  
**Step 4:** Construct the fused image by applying the inverse NSCT to the fused subband coefficients.

ALGORITHM 1: Image fusion method with saliency-motivated PCNN.

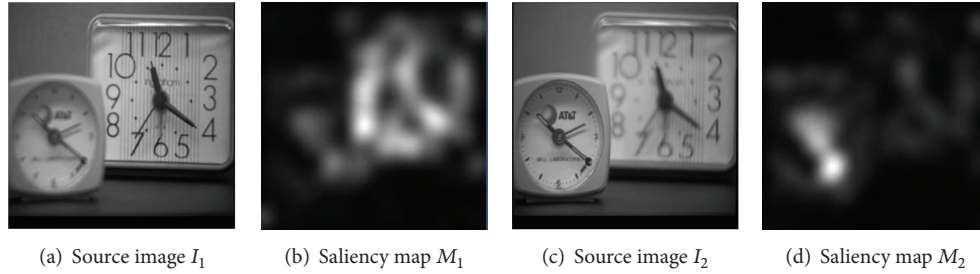


FIGURE 4: The results of saliency detection from two complementary input images. (a) and (c) Multifocus source images (b) and (d) saliency maps from PFT.

PCNN for the high-pass subbands. According to the visual attention mechanism, different regions in an image have varying importance for HVS, so the saliency detection is performed on source images to yield saliency maps which indicate the significance level of every pixel in source images. Based on the characteristics, the PFT model is performed on the high-pass subbands to produce the saliency maps, which indicate the importance level of coefficients. And then, the obtained saliency maps are combined with the corresponding high-pass subband coefficients as the input to motivate PCNN. Coefficients with large firing times are selected as the fused coefficients. In addition, the low-pass subband of NSCT decomposition in the coarsest scale contains the main energy of source images and denotes abundant structural information. The fusion rule of the low-pass subband employs a weighted fusion rule based on firing times of PCNN.

The activity maps of high-pass subbands as the criteria of selecting coefficients are presented by the firing map of saliency-motivated PCNN. The activity level indicates the magnitude of coefficients. The coefficients of greater energy carry more important information, so the coefficients of greater activity level are selected as the fused coefficients. Now, according to (7) and (2), the fused coefficients in location  $(x, y)$  of high-pass subbands denoted by  $F_{j,l}(x, y)$  are defined as follows:

$$F_{j,l}(x, y) = \begin{cases} C_{j,l}^A(x, y), & \text{if } T_{A,xy}^{j,l} > T_{B,xy}^{j,l} \\ C_{j,l}^B(x, y), & \text{otherwise.} \end{cases} \quad (8)$$

The fused coefficients of low-pass subbands denoted by  $F_{j_0}(x, y)$  employ a weighted fusion rule based on firing times

of PCNN on coefficients  $C_{j_0}^A(x, y)$  and  $C_{j_0}^B(x, y)$ , which are defined as follows:

$$F_{j_0}(x, y) = \omega * C_{j_0}^A(x, y) + (1 - \omega) * C_{j_0}^B(x, y) \quad (9)$$

$$\omega = \frac{T_{A,xy}^{j_0}}{T_{A,xy}^{j_0} + T_{B,xy}^{j_0}},$$

where  $\omega$  is the weight of coefficients and  $T_{xy}^{j_0}$  is computed by (1) and (2). Because the low-pass subband at the coarsest scale does not contain the direction, here the symbol  $l$  in (1) and (2) is changed to 0. Specifically,  $C_{j,l}(x, y)$  in (1) is replaced by  $C_{j_0}^{A/B}(x, y)$ , and  $T_{xy}^{j,l}$  in (2) is replaced by  $T_{xy}^{j_0}$ .

Finally, apply the inverse NSCT to the fused coefficients  $\{F_{j_0}(x, y), F_{j,l}(x, y)\}$  and then obtain the fused image  $F$ . At last, the algorithm description of the proposed image fusion approach is shown in Algorithm 1 for better understanding.

## 4. Experiments and Analysis

In this section, the proposed image fusion algorithm based on NSCT and saliency-motivated PCNN (named as NSCT-SPCNN) is tested on several sets of images. The goal of the tests is to validate if the proposed algorithm can be used in the real applications and varying surroundings. For comparison, besides the fusion scheme proposed in this paper, another three fusion algorithms, the Laplacian pyramid transform based (LPT), discrete wavelet transform based (DWT), and NSCT-simple based, are used to fuse the same images. All of these use averaging and absolute maximum selection schemes for merging low- and high-pass subband coefficients, respectively. The decomposition level of all of the transforms is three. Extensive experiments with multifocus image fusion and different sensor image fusion

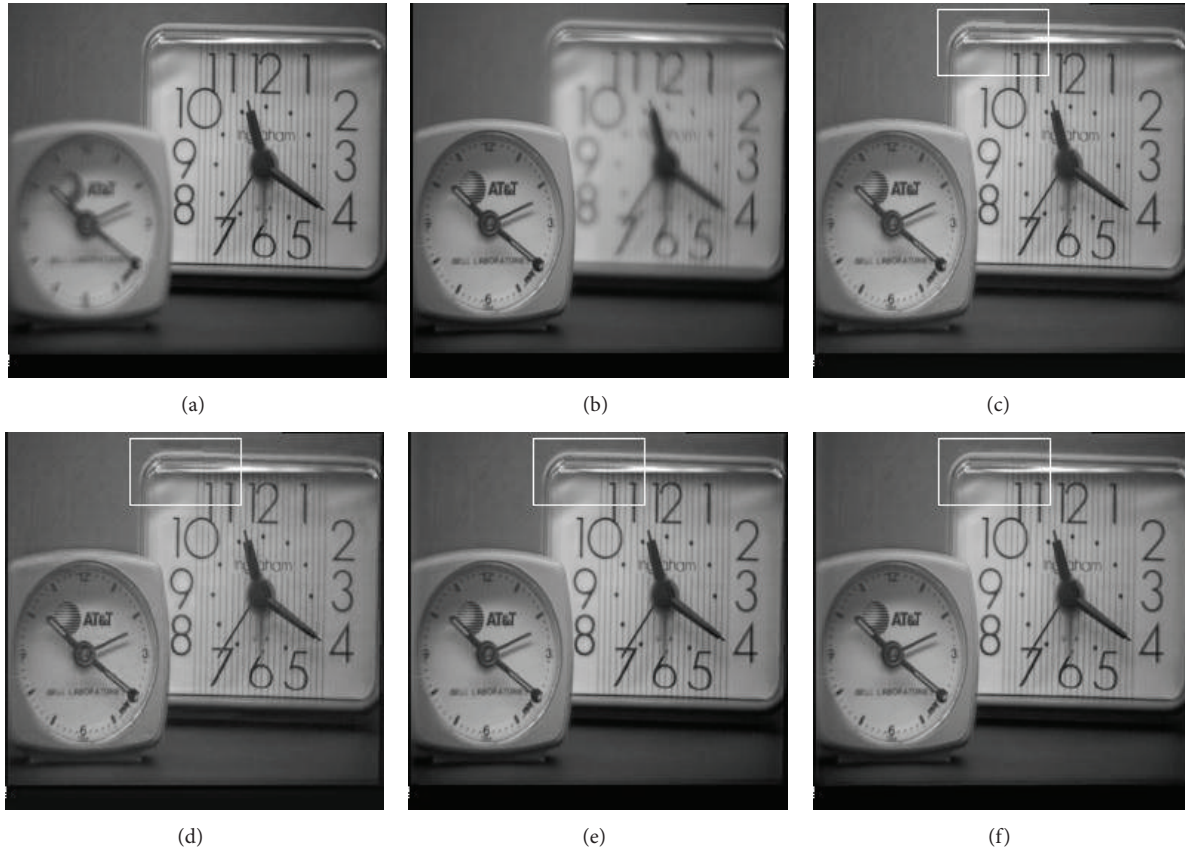


FIGURE 5: “Clock” source images (256 level, size of  $256 \times 256$ ) and fused images: (a) focus on the right; (b) focus on the left and fused images using (c) LPT, (d) DWT, (e) NSCT-simple, and (f) NSCT-SPCNN methods.

have been performed. Here, three groups of different images were tested to evaluate the performance of the proposed algorithm: a set of multifocus images, a set of multimodal medical images, and a set of artificial out-of-focus images. It is assumed that source images have been registered. The fused results were evaluated using subjective visual inspection and objective assessment tools.

**4.1. Visual Analysis.** The first experiment uses two multifocus source images and four fused images produced by LPT, DWT, NSCT-simple, and NSCT-SPCNN methods, shown in Figure 5. Figure 5(a) focuses on the right region. Figure 5(b) focuses on the left region. The fused images contain all of focus point regions of source images and expand effectively the depth of a scene. In Figure 5(f), the saliency value associated with coefficients as the input to motivate PCNN is employed to compute the activity level of coefficients. In this way, the algorithm makes sure that the activity level of the saliency pixel is higher, so that the fused image preserves the saliency regions of source images. The images in Figures 5(c)–5(e) are not clear enough and have lower contrast; artifacts were also introduced. The differences among the fused images are very slight, so it is difficult to evaluate the image quality by direct visual inspection. To observe the image quality in more detail, one area in the fused images was magnified.

Figures 6(a)–6(d) show magnified images of the region marked by the boxes in Figures 5(c)–5(f). The performance of the different fusion algorithms can be observed from these magnified images. The images fused using LPT, DWT and NSCT-simple methods (Figures 6(a)–6(c)) have some deformation leading to bend edges. Figure 6(d) has the better visual quality than others with the best visual effect and smoother and sharper edges. This comparison reveals that the NSCT-SPCNN-based fusion approach effectively determines complementary or redundant information between source images. It can preserve all the important information of the source images while avoiding artifacts. In addition, a clearer comparison is made by examining the differences between the fused and source images, shown in Figure 7. Figures 7(a)–7(d) show the difference images between Figures 5(c)–5(f) and Figure 5(a). Observing Figures 7(a)–7(d), we can see that there is little difference between the fused image by NSCT-SPCNN (Figure 5(d)) and the right-focus clock (Figure 5(a)), while a lot of difference between the fused images in Figures 5(c)–5(e) and the right-focus source image in Figure 5(a) can be seen. This further demonstrates that the NSCT-SPCNN-based method is with higher fusion performance.

Figure 8 shows a group of multimodal medical images and images fused using four different fusion algorithms. A set of spatial out-of-focus images are shown in Figure 9,

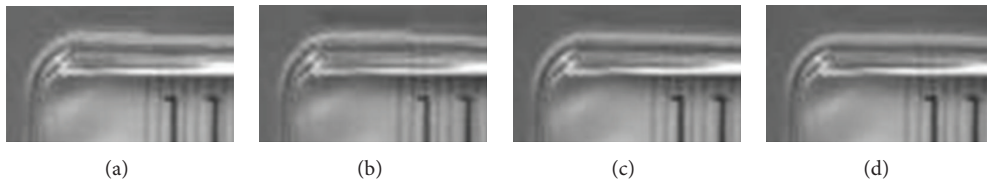


FIGURE 6: Magnified regions from the fused images in Figures 5(c)–5(f) using (a) LPT, (b) DWT, (c) NSCT-simple, and (d) NSCT-SPCNN methods.

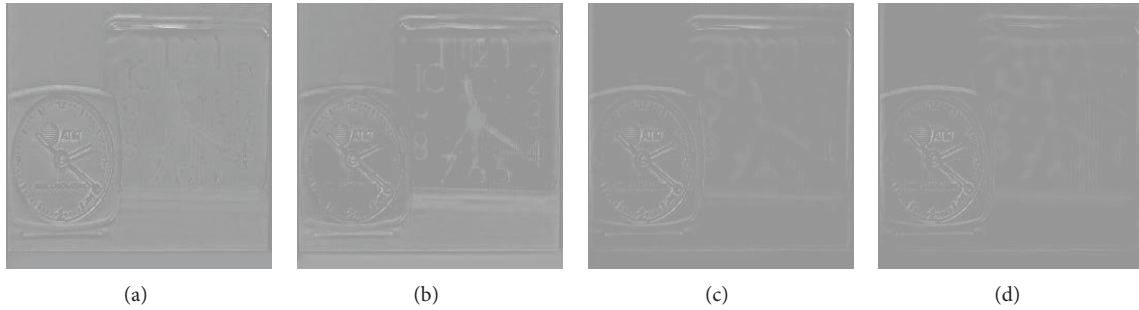


FIGURE 7: (a) Difference image between Figures 5(c) and 5(a); (b) difference image between Figures 5(d) and 5(a); (c) difference image between Figures 5(e) and 5(a); (d) difference image between Figures 5(f) and 5(a).

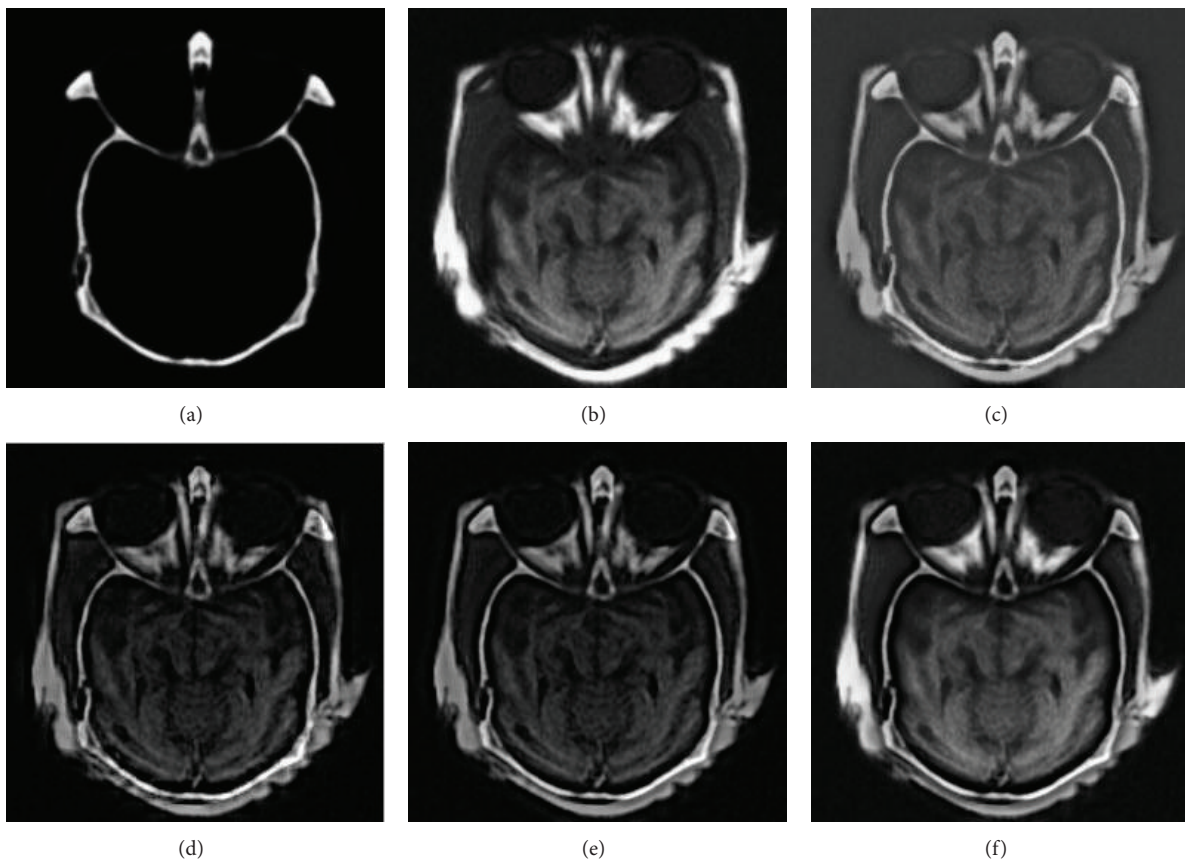


FIGURE 8: Medical source images (256 level, size of  $256 \times 256$ ) and fused images: (a) and (b) source images and fused images using (c) LPT, (d) DWT, (e) NSCT-simple, and (f) NSCT-SPCNN methods.



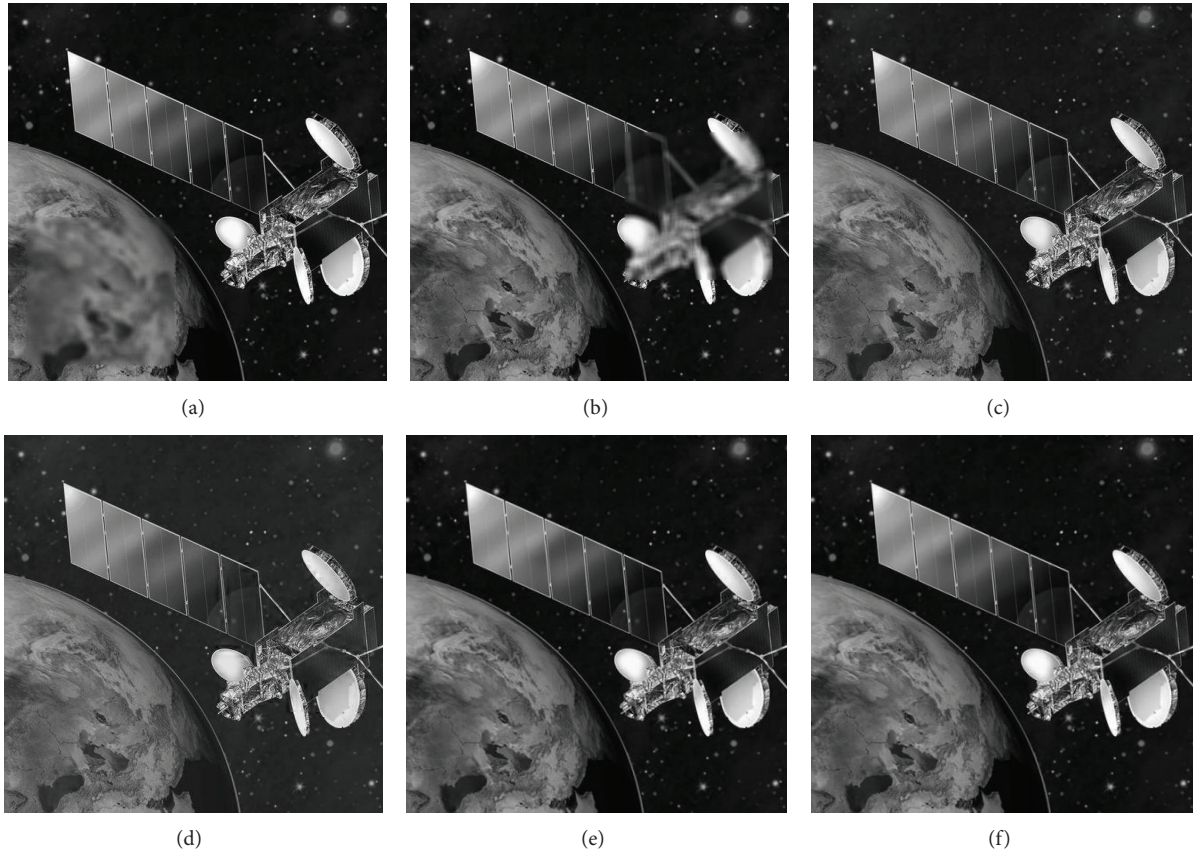


FIGURE 9: Spatial source images (256 level, size of  $512 \times 512$ ) and fused images: (a) focus on the right; (b) focus on the left and fused images using (c) LPT, (d) DWT, (e) NSCT-simple, and (f) NSCT-SPCNN methods.

which is obtained by artificial blurring different regions of the ground truth image using a Gaussian filter. Experimental results demonstrate the visual effects of two sets of images in Figures 8 and 9 coinciding with Figure 5. Thus, the proposed algorithm NSCT-SPCNN-based can effectively improve the quality of the fused image both multifocus, and multimodal images.

**4.2. Objective Analysis.** In previous discussion, the fusion results of different algorithms have been analyzed by visual aspect. However, the performance of fusion algorithms needs to be further evaluated using objective metric tools. A successful fusion technique has to satisfy many conditions, such as preserving important features of source images, enhancing contrast, and avoiding artifacts. Mutual information (MI) [25] and an objective image fusion performance measure ( $Q_{AB/F}$ ) [26] are employed to evaluate the fusion performance of different fusion methods quantitatively. MI indicates how much of the input information the fused image contains.  $Q_{AB/F}$  reflects the preservation of input edge information in the fused image. For the two metrics, the higher the values are, the better are the fusion results.

Figure 10 shows the quality measurement results for fused images in Figure 5 and Figures 8 and 9. Observing Figure 10, we can see that the LPT and DWT methods are the worst. This

is consistent with the subjective visual analysis. Compared with other fusion algorithms, the NSCT-SPCNN yields the optimal performance. Experimental results demonstrate that the proposed NSCT-SPCNN algorithm can preserve the saliency regions of source images and improve the quality of the fused image.

Finally, the computational performance of the proposed NSCT-SPCNN algorithm is tested on three sets of images (Figures 5, 8, and 9). The hardware setup is an Intel Core i5-3479 PC with 4 GB RAMs. Our Matlab implementation takes about 16 seconds for Figures 5 and 8 and 73 seconds for Figure 9. Meanwhile, the NSCT-simple-based fusion method takes about 17 seconds for Figures 5 and 8 and 72 seconds for Figure 9. The LPT- and DWT-based fusion methods take less than 1 second. From the comparison, we can observe that the computational bottleneck lies in the NSCT transform. Therefore, a more efficient MST tool needs to be applied in the future.

## 5. Conclusion

The paper proposes a novel image fusion algorithm based on NSCT and saliency-motivated PCNN. In fusion for high-pass subbands, a saliency-motivated PCNN model is proposed. The key idea is that depending on the human visual attention



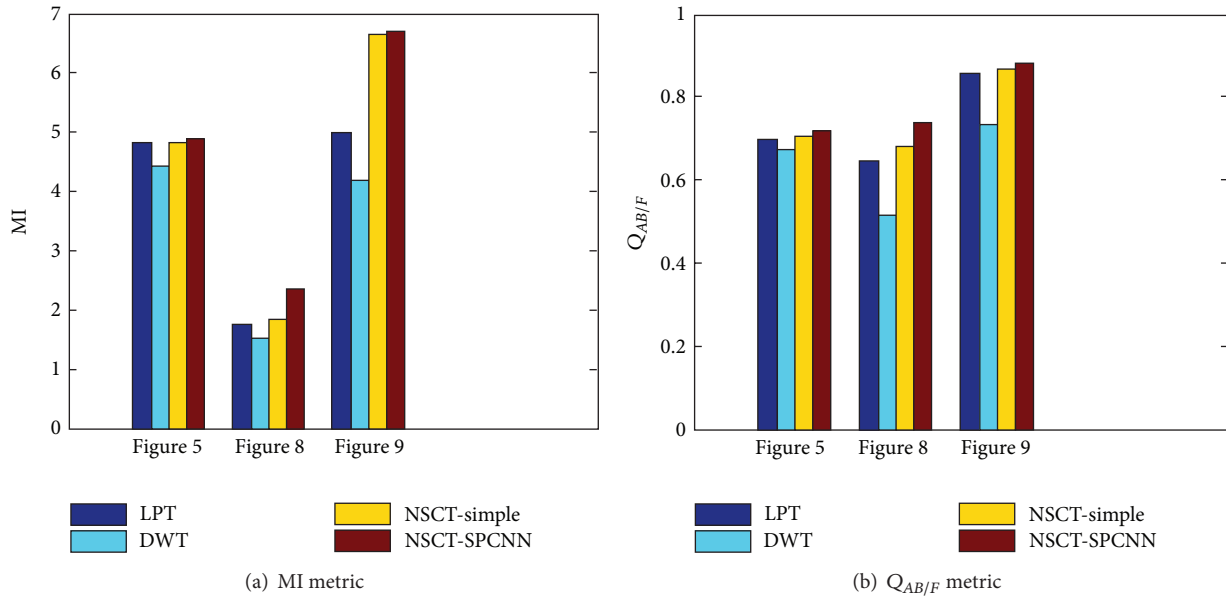


FIGURE 10: Quality metrics for the different fusion methods.

model, the visual saliency map is first built on high-pass sub-band coefficients of NSCT, and then the algorithm combines the visual saliency map with the coefficients of NSCT as input to motivate PCNN. Coefficients with large firing times are employed as the fused high-pass subband coefficients. Low-pass subband coefficients are merged to develop a weighted fusion rule based on firing times of PCNN. The algorithm can preserve the completeness and the sharpness of object regions. The fused image is more natural and can satisfy the requirement of HVS. Experiments illustrate that the proposed fusion algorithm improves greatly the quality of the fused images.

## Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program nos. 2012CB821200 2012CB821206), the National Natural Science Foundation of China (nos. 91024001, 61070142), and the Beijing Natural Science Foundation (no. 4111002).

## References

- [1] A. Leykin, Y. Ran, and R. Hammoud, "Thermal-visible video fusion for moving target tracking and pedestrian classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, June 2007.
- [2] Y. F. Niu, S. T. Xu, L. Z. Wu, and W. D. Hu, "Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform," *Mathematical Problems in Engineering*, vol. 2012, Article ID 275138, 10 pages, 2012.
- [3] L. Yang, B. L. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing*, vol. 72, no. 1-3, pp. 203-211, 2008.
- [4] J. R. Raol, *Multi-Sensor Data Fusion with Matlab*, CRC Press Taylor and Francis Group, 2010.
- [5] R. Blum, Z. Xue, and Z. Zhang, "Chapter 1: an overview of image fusion," in *Multi-Sensor Image Fusion and Its Applications*, pp. 1-36, CRC Press Taylor and Francis Group, 2006.
- [6] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532-540, 1983.
- [7] V. S. Petrović and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 228-237, 2004.
- [8] G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognition*, vol. 37, no. 9, pp. 1855-1872, 2004.
- [9] X.-C. Yu, F. Ni, S.-L. Long, and W.-J. Pei, "Remote sensing image fusion based on integer wavelet transformation and ordered nonnegative independent component analysis," *GIScience and Remote Sensing*, vol. 49, no. 3, pp. 364-377, 2012.
- [10] M. N. Do and M. Vetterli, "The finite ridgelet transform for image representation," *IEEE Transactions on Image Processing*, vol. 12, no. 1, pp. 16-28, 2003.
- [11] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, 2003.
- [12] K. Liu, L. Guo, and J. Chen, "Contourlet transform for image fusion using cycle spinning," *Journal of Systems Engineering and Electronics*, vol. 22, no. 2, pp. 353-357, 2011.
- [13] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861-899, 2006.
- [14] S. Ren, J. Cheng, and M. Li, "Multiresolution fusion of Pan and MS images based on the Curvelet transform," in *Proceedings of the 30th IEEE International Geoscience and Remote Sensing Symposium (IGARSS '10)*, pp. 472-475, July 2010.
- [15] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampling contourlet transform: theory, design, and applications," *IEEE*

- Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.
- [16] G. Piella, “Image fusion for enhanced visualization: a variational approach,” *International Journal of Computer Vision*, vol. 83, no. 1, pp. 1–11, 2009.
- [17] C. Ludusan and O. Laviolle, “Multifocus image fusion and denoising: a variational approach,” *Pattern Recognition Letters*, vol. 33, no. 10, pp. 1388–1396, 2012.
- [18] Z. Zhang and R. S. Blum, “A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application,” *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315–1326, 1999.
- [19] G. Piella, “A general framework for multiresolution image fusion: from pixels to regions,” *Information Fusion*, vol. 4, no. 4, pp. 259–280, 2003.
- [20] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, “Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex,” *Neural Computation*, vol. 2, no. 3, pp. 293–307, 1990.
- [21] J. L. Johnson and M. L. Padgett, “PCNN models and applications,” *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 480–498, 1999.
- [22] R. P. Broussard, S. K. Rogers, M. E. Oxley, and G. L. Tarr, “Physiologically motivated image fusion for object detection using a pulse coupled neural network,” *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 554–563, 1999.
- [23] X.-B. Qu, J.-W. Yan, H.-Z. Xiao, and Z.-Q. Zhu, “Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain,” *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1508–1514, 2008.
- [24] C. Guo, Q. Ma, and L. Zhang, “Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform,” in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, June 2008.
- [25] G. Qu, D. Zhang, and P. Yan, “Information measure for performance of image fusion,” *Electronics Letters*, vol. 38, no. 7, pp. 313–315, 2002.
- [26] C. S. Xydeas and V. Petrović, “Objective image fusion performance measure,” *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

