

Multi-focus image fusion based on dynamic threshold neural P systems and surfacelet transform

Bo Li^a, Hong Peng^{a,*}, Jun Wang^b, Xiangnian Huang^a

^a School of Computer and Software Engineering, Xihua University, Chengdu, Sichuan, 610039, China

^b School of Electrical Engineering and Electronic Information, Xihua University, Chengdu, 610039, China



ARTICLE INFO

Article history:

Received 1 January 2020

Received in revised form 12 March 2020

Accepted 18 March 2020

Available online 23 March 2020

Keywords:

Multi-focus image fusion
Dynamic threshold neural P systems
Local topology
Surfacelet domain

ABSTRACT

Dynamic threshold neural P systems (DTNP systems) are recently proposed distributed and parallel computing models, inspired from the intersecting cortical model. DTNP systems differ from spiking neural P systems (SNP systems) due to the introduction of dynamic threshold mechanism of neurons. DTNP systems have been theoretically proven to be Turing universal computing devices. This paper discusses how to apply DTNP systems to deal with the fusion of multi-focus images, and proposes a novel image fusion method based on DTNP systems in surfacelet domain. Based on four DTNP systems with local topology, a multi-focus image fusion framework in surfacelet domain is developed, where DTNP systems are applied to control the fusion of low- and high-frequency coefficients in surfacelet domain. The proposed fusion method is evaluated on an open dataset of 20 multi-focus images in terms of five fusion quality metrics, and compared with 10 state-of-the-art fusion methods. Quantitative and qualitative experimental results demonstrate the advantages of the proposed fusion method in terms of visual quality, fusion performance and computational efficiency.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Multi-focus image fusion has become an emerging research topic because of its availability and effectiveness in image processing and computer vision [1]. Since each imaging device with optical camera has a limited depth of field, the image captured by the camera cannot all be in focus. As a result, the objects with a specific depth of field are sharp, but other objects are blurred. Multi-focus image fusion is an effective technique to address the above problems. Multi-focus image fusion is such a task that merges two or more source images with the different depths of field to produce a sharper image. Because the fused image has more detail information, it is more suitable for human visual system.

Many multi-focus image fusion methods have been presented in recent years. These methods can be classified into two classes: spatial and transform domain methods [2]. Spatial domain methods directly merge multi-focus source images without converting images into other types of expressions. Such methods are further divided into two subclasses: pixel- and region-level methods [3, 4]. Pixel-level methods merge source images by averaging the corresponding pixels. They are simple in implementation and computationally fast. However, their major drawback is that they

can introduce artifacts, such as ghosting and blurring. Region-level methods divide source images into regions, and then use various sharpness measures (such as spatial frequency or gradients) to choose regions for fusion [5–7]. In addition, pulse-coupled neural networks (PCNNs) were applied for multi-focus image fusion [8–10].

In recent years, transform domain methods have received a lot of attention. These methods contain three steps: source images are first converted to a transform domain; and then coefficients in transform domain are merged to produce the fused coefficients based on fusion rules; finally, the fused coefficients are converted back into spatial domain to form a composite image via inverse transform. Multi-scale transform has been widely applied in image fusion due to its excellent locality and multi-resolution features, for example, Laplacian pyramid (LP) [11–13], gradient pyramid [14,15], and wavelet transform [16]. The wavelet transform has become a popular fusion method in transform domain methods, including discrete wavelet transform (DWT) [17–19] and dual-tree complex wavelet transform [20]. However, these wavelet transforms have drawbacks in terms of non-shift-invariance, poor spatiality, and non-time-invariance. To overcome these drawbacks, some multi-scale transforms have been introduced for image fusion, including curvelet transform [21,22], non-subsampled contourlet transform [23,24], non-subsampled shearlet transform [25], shift-invariant dual-tree complex shearlet transform [26], and sparse representations [27–29]. Currently, image fusion based on surfacelet transform (ST) provides some

* Corresponding author.

E-mail address: ph.xhu@hotmail.com (H. Peng).

excellent properties, such as shift invariance and the ability to capture high-dimensional singularities [30]. This work aims to develop a novel image fusion method in ST domain.

In recent years, several convolutional neural network (CNN)-based fusion methods have been developed for multi-focus image fusion. Liu et al. [31] discussed a CNN-based multi-focus image fusion method where an CNN-based model was considered to solve classification problems. Tang et al. [32] investigated a pixel CNN for multi-focus image fusion, where a model was trained to learn the probabilities of focused, defocused, and unknown pixels based on their neighborhood pixel information. Gao et al. [33] investigated an CNN for multi-focus image fusion where a deeper network was used for constructing an initial decision map. Amin-Naji et al. [34] presented an CNN-based multi-focus image fusion method combined with ensemble learning. Zhang et al. [35] presented a general image fusion framework based on an CNN called IFCNN. These CNN-based methods have demonstrated competitive fusion performance compared with previous fusion methods. However, their training processes are very time-consuming.

Dynamic threshold neural P systems (DTNP systems) are a recently developed distributed and parallel computing model [36], combining the spiking mechanism and dynamic threshold mechanism of neurons. Our previous work has proved that DTNP systems are Turing-universal number generating/accepting devices and function computing devices. This paper focuses on application of DTNP systems in the fusion of multi-focus images, and proposes a novel DTNP-systems-based fusion method in surfacelet transform (ST) domain for multi-focus images. For this goal, four DTNP systems with local topology are designed to develop a fusion framework for multi-focus images. The feature matrixes of low-frequency and high-frequency ST coefficients of multi-focus images are considered as the external inputs of four DTNP systems and the corresponding outputs are used as control condition of fusion rules. The contribution of this paper can be summarized as follows.

- (i) DTNP systems with local topology are designed;
- (ii) A novel fusion framework in ST domain for multi-focus images is developed, where four DTNP systems are its key component.
- (iii) A low-frequency fusion rule based on DTNP system is developed, where SML feature matrix of low-frequency ST coefficients of multi-focus images is the external input of DTNP systems and the corresponding outputs are used to control the low-frequency fusion rule.
- (iv) A high-frequency fusion rule based on DTNP system is developed, where SF feature matrix of high-frequency ST coefficients of multi-focus images is the external input of DTNP systems and the corresponding outputs are used to control the high-frequency fusion rule.

The rest of this paper is organized as follows. Section 2 first introduces DTNP systems with local topology, and then briefly reviews surfacelet transform. Section 3 describes in detail the proposed fusion framework in ST domain for multi-focus images. Section 4 gives the experimental results. Conclusions and discussion are drawn in Section 5.

2. Methodology

In this section, we first review the literature related to variants of SNP systems, and then introduce dynamic threshold neural P systems (DTNP systems) with local topology and briefly review surfacelet transform.

2.1. Literature review

Spiking neural P systems (SNP systems) are a kind of distributed and parallel computing models, abstracted by the mechanism of spiking neurons [37]. Due to the abstraction of spiking neurons, SNP systems are viewed as 3rd-generation neural network models, like spiking neural networks. An SNP system consists of a directed graph of neurons, where every neuron contains a data unit, which stores one or more spikes, and a set of spiking/firing rules of the form $E/a^c \rightarrow a^p$. Data units are used to express the states of neurons, and are evolved by spiking/firing rules. Therefore, SNP systems can be used to characterize dynamic systems. Moreover, SNP systems can work in three modes: generating, accepting and computing modes.

Abstracted by different biological mechanisms and/or introduced the methods in mathematics or computer sciences, a lot of variants of SNP systems have been proposed in the past years. With this biological inspiration that astrocytes have excitatory and inhibitory influence on synapses, SNP systems with astrocytes have been proposed in Păun [38]. Considering a pair of anti-spikes (a, \bar{a}), Pan et al. [39] discussed SNP system with anti-spikes. Considering a new communication strategy among neurons, Pan et al. [40] investigated an SNP systems with communication on request. Abstracted by the biological fact that the synapse has one or more chemical channels, SNP systems with multiple channels were discussed in Peng et al. [41]. By moving the rules from neurons to synapses, SNP systems with rules on synapses have been investigated in Song et al. [42]. Inspired from the biological phenomena that every neuron has a negative or positive charge, SNP systems with polarizations were investigated in Wu et al. [43]. Abstracted by the structural dynamism of biological synapses, Cabarle et al. [44] presented an SNP system with structural plasticity. Considering coupled modulation mechanism in SNP systems, coupled neural P systems were investigated by Peng et al. [45]. Moreover, spiking neural P systems with inhibitory rules and nonlinear spiking neural P systems were discussed in Peng et al. [46,47].

With the restriction that at each step one neuron works at most, several sequential SNP systems were discussed in Ibarra et al. [48]. Usually, a global clock is assumed in SNP system, so they are synchronized. However, asynchronous SNP systems was investigated in Cavaliere et al. [49]. By integrating fuzzy logics in SNP systems, several fuzzy SNP systems were developed in the recent years, for example, fuzzy reasoning SNP systems [50], weighted fuzzy SNP systems [51] and Interval-valued fuzzy SNP systems [52] and intuitionistic fuzzy spiking neural P systems [53]. Computational power of SNP systems has been investigated: most variants of SNP systems were proven to be Turing universal as number generating/accepting devices [54], language generating devices [55] and function computing devices [56]. SNP systems were used to (theoretically) solve computationally hard problems in a feasible (polynomial or linear) time [57,58].

2.2. DTNP systems with local topology

Dynamic threshold neural P systems (DTNP systems) are a variant of SNP systems. Our previous theoretical work has proved that DTNP systems are Turing-universal number generating/accepting and function computing devices [36]. To deal with multi-focus images, DTNP systems are designed as an array of neurons with local topology, i.e., DTNP systems with local topology.

Suppose that I is a multi-focus image with size $m \times n$, and $(SML_{ij})_{m \times n}$ (resp. $(SF_{ij})_{m \times n}$) is a matrix of low-frequency (resp. high-frequency) ST coefficients of the image. Therefore, an DTNP

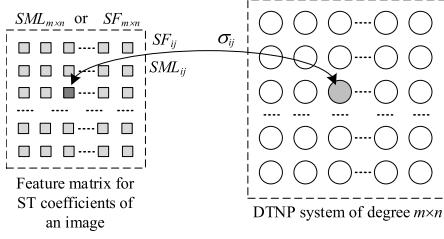


Fig. 1. An DTNP system Π and the corresponding feature matrix of ST coefficients, $SF_{m \times n}$ or $SML_{m \times n}$.

system Π with local topology, of degree $m \times n$, is designed, which consists of an array of $m \times n$ neurons. Fig. 1 shows the corresponding relationship between DTNP system Π and feature matrix, $SML_{m \times n}$ or $SF_{m \times n}$, where the feature matrix has $m \times n$ elements, Π contains $m \times n$ neurons, and neuron σ_{ij} is associated with feature value SML_{ij} or SF_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$.

Definition 1. An DTNP system with local topology, of degree $m \times n$, is defined by

$$\Pi = (O, \sigma_{11}, \sigma_{12}, \dots, \sigma_{1n}, \dots, \sigma_{m1}, \sigma_{m2}, \dots, \sigma_{mn}, syn)$$

where

- (1) $O = \{a\}$ is an alphabet (the object a is known as the spike);
- (2) $\sigma_{11}, \sigma_{12}, \dots, \sigma_{mn}$ are an array of $m \times n$ neurons of the form

$$\sigma_{ij} = (u_{ij}, \tau_{ij}, R_{ij}), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n$$

where

- (a) $u_{ij} \in \mathbb{R}$ is the value of spikes in neuron σ_{ij} ;
- (b) $\tau_{ij} \in \mathbb{R}$ is the (dynamic) threshold in neuron σ_{ij} ;
- (c) R_{ij} denotes the finite set of firing rules, of the form $E/(a^u, a^\tau) \rightarrow a^p$, where E is the firing condition, $E \equiv (u_i(t) \geq \tau_i(t)) \wedge (u_i(t) \geq u) \wedge (\tau_i(t) \geq \tau)$. If the firing condition is satisfied, then neuron σ_{ij} fires, meaning that the spikes with the value u in data unit and the spikes with the value τ in dynamic threshold unit are consumed, and then the spikes with the value p are produced and sent to the neighboring neurons. According to the firing mechanism, the state equation for neuron σ_{ij} can be defined by

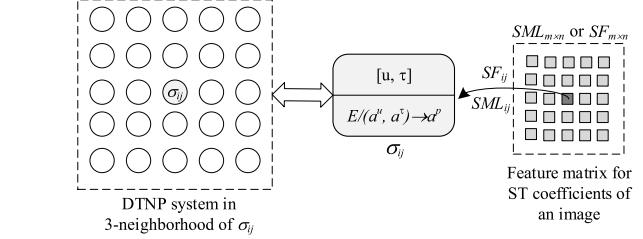


Fig. 2. The DTNP system in 3-neighborhood δ_3 of neuron σ_{ij} .

neurons and (2) an external stimulus, i.e., feature value SML_{ij} or SF_{ij} .

Neuron σ_{ij} has a firing rule of the form of $E/(a^u, a^\tau) \rightarrow a^p$, where E is a firing condition defined by $E \equiv (u_i(t) \geq \tau_i(t)) \wedge (u_i(t) \geq u) \wedge (\tau_i(t) \geq \tau)$. If the firing condition is satisfied, then neuron σ_{ij} fires, meaning that the spikes with the value u in data unit and the spikes with the value τ in dynamic threshold unit are consumed, and then the spikes with the value p are produced and sent to the neighboring neurons. According to the firing mechanism, the state equation for neuron σ_{ij} can be defined by

$$u_{ij}(t+1) = \begin{cases} u_{ij}(t) - u + SML_{ij}(\text{or } SF_{ij}) & \text{if } \sigma_{ij} \text{ fires} \\ + \sum_{\sigma_{kl} \in \delta_r} w_{kl} p_{kl}(t), \\ u_{ij}(t) + SML_{ij}(\text{or } SF_{ij}) & \text{otherwise} \\ + \sum_{\sigma_{kl} \in \delta_r} w_{kl} p_{kl}(t), \end{cases} \quad (2)$$

$$\tau_{ij}(t+1) = \begin{cases} \tau_{ij}(t) - \tau + p, & \text{if } \sigma_{ij} \text{ fires} \\ \tau_{ij}(t), & \text{otherwise} \end{cases} \quad (3)$$

where $p_{kl}(t)$ is the value of spikes received by neuron σ_{ij} from neighboring neuron σ_{kl} and $w_{kl}(t)$ is the corresponding local weight, and SML_{ij} or SF_{ij} is an external stimulus; p is the value of spikes produced by neuron σ_{ij} when it fires.

The working procedure for DTNP system Π can be described as follows. DTNP system Π has five priori parameters: (i) iteration number t_{max} , (ii) initial threshold τ_0 for each neuron, (iii) value of generated spikes p in firing rule for each neuron, (iv) neighborhood radius r , and (v) local weight matrix $W_{r \times r}$. Initially, for each neuron σ_{ij} , we set $u_{ij}(0) = v_{ij}(0) = 0$ and $\tau_{ij}(0) = \tau_0$, and assign $W_{r \times r}$. DTNP system Π starts from the initial state and then it is executed iteratively until iteration number t_{max} is reached. Thus, the system halts. During the computation, the sum of spikes that each neuron σ_{ij} sends out is considered as its output and used as a control signal for the fusion of multi-focus images. For simplicity, a maximum consumption strategy is adopted in the application of firing rules (i.e., set $u = u_{ij}(t)$ and $\tau = \tau_{ij}(t)$).

2.3. Surfacelet transform

We briefly review surfacelet transform used in this paper. More details of surfacelet transform can be found in Refs. [30,59,60].

Yu et al. [60] extended the directional filter bank (DFB) [59] to propose a new filter banks for N-dimensional ($N \geq 2$) signals, N-dimensional directional filter bank (NDFB). Based on the NDFB with a new multiscale pyramid, Yu et al. [60] proposed the surfacelet transform, which can be used to efficiently capture and express surface-like singularities with different sizes in multi-dimensional data. The idea is analogous to the contourlet construction, however, an important distinction is that surfacelet transform employs a new multiscale pyramid structure instead

The DTNP system consists of $m \times n$ neurons, and has a topology of directed graph, where $m \times n$ neurons are the nodes and the arcs describe synaptic connections between these neurons.

In DTNP system Π , each neuron only communicates with its neighboring neurons. For convenience, we define a r -neighborhood of neurons σ_{ij} as follows:

$$\delta_r(\sigma_{ij}) = \{\sigma_{kl} \mid |k - i| \leq r, |l - j| \leq r\}. \quad (1)$$

Fig. 2 shows an example of a 3-neighborhood of neuron σ_{ij} . The neuron σ_{ij} first receives the spikes from the neighboring neurons in δ_3 , and then sends the generated spikes back to each of its neighboring neurons after the use of firing rule. Therefore, two-way arrows are used to describe the synaptic connections between neuron σ_{ij} and its neighboring neurons.

Each neuron σ_{ij} has a data unit u_{ij} and a dynamic threshold unit τ_{ij} . The middle part in Fig. 2 shows the detail of neuron σ_{ij} , including its data unit and dynamic threshold unit, $[u_{ij}, \tau_{ij}]$, and firing rule $E/(a^u, a^\tau) \rightarrow a^p$. In addition, neuron σ_{ij} has an external input, SML_{ij} or SF_{ij} , i.e., feature value of ST coefficient at position (i, j) , shown in right side in Fig. 2.

We now explain the working mechanism of DTNP system Π . Suppose that in DTNP system Π , neuron σ_{ij} has local connections to neighboring neurons in r -neighborhood δ_r . Neuron σ_{ij} has inputs of two types: (1) the spikes received from its neighboring

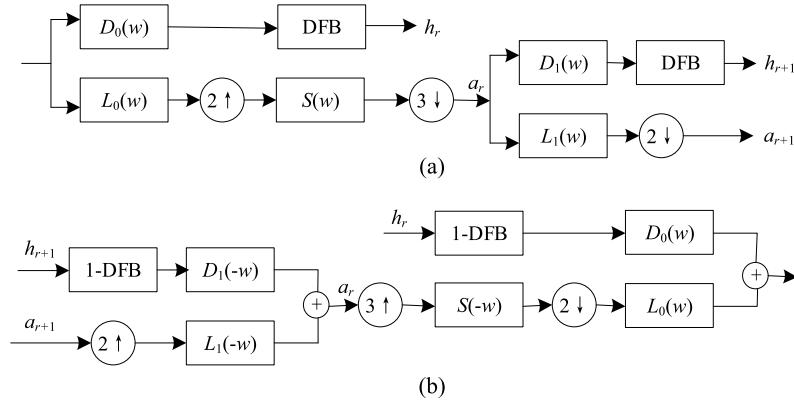


Fig. 3. The structure of surfacelet transform: (a) decomposition and (b) synthesis.

of the Laplacian pyramid as in contourlet. Although this method contains much redundancy information than Laplacian pyramid, it can avoid the mixing phenomenon that NDFB produces in the frequency domain. This specific sampling method makes the decomposing coefficient of surfacelet well express in the space and frequency domain. Therefore, main advantage of designing this filter in the space domain is that the frequency response of the image can be strictly controlled. Fig. 3 shows the surfacelet transform structure.

3. The proposed fusion method for multi-focus images

We propose an DTNP-systems-based image fusion framework in ST domain for multi-focus images, shown in Fig. 4. The fusion framework includes four parts: (i) ST transform; (ii) fusion rules; (iii) inverse ST transform; (iv) optimization. In Fig. 4, source images A and B are two multi-focus images.

The multi-focus images are first decomposed to the ST coefficients by using ST transform. Then, the ST coefficients are fused to generate the fused ST coefficients. However, low-frequency coefficients and high-frequency coefficients are fused separately by two groups of fusion rules based on DTNP systems, each using two DTNP systems. DTNP systems are a kernel part in the proposed fusion framework, and they are designed to control the fusion of low-frequency and high-frequency ST coefficients, respectively. Two DTNP systems, Π_{AL} and Π_{BL} , are considered to merge low-frequency coefficients of multi-focus images, whereas other two DTNP systems, Π_{AH} and Π_{BH} , are applied to merge high-frequency coefficients of multi-focus images. The merged ST coefficients are converted back to reconstruct a fusion image by using inverse ST transform. Finally, the obtained fusion image is further optimized.

According to the fusion framework, the proposed fusion method for multi-focus images is implemented by the following steps:

- ST transform: multi-focus images, A and B , are decomposed to the ST coefficients by using ST transform, respectively; thus, we obtain low-frequency ST coefficients and high-frequency ST coefficients for each of multi-focus images.
- The fusion of low-frequency ST coefficients: based on the low-frequency ST coefficients of multi-focus images A and B , we compute two feature matrixes, SML^{AL} and SML^{BL} , which are regarded as the external inputs of DTNP systems Π_{AL} and Π_{BL} respectively; the two DTNP systems start from initial states, and then they are executed constantly until iteration number t_{max} reaches; hence, they halt; the outputs of the two DTNP systems are used as the control signal of low-frequency fusion rules (Eq. (see (6))).

- The fusion of high-frequency ST coefficients: based on the high-frequency ST coefficients of multi-focus images A and B , we compute two feature matrixes, SF^{AH} and SF^{BH} , which are regarded as the external inputs of DTNP systems Π_{AH} and Π_{BH} respectively; the two DTNP systems start from initial states, and then they are executed constantly until iteration number t_{max} reaches; hence, they halt; the outputs of the two DTNP systems are used as the control signal of high-frequency fusion rules (Eq. (see (10))).
- Inverse ST transform: the combined ST coefficients are converted back to reconstruct a fusion image by using inverse transform.
- Consistent verification: the initial fusion image C is further optimized by using Eqs. (11)–(13).

In the following, we describe in detail the used two fusion rules for low-frequency and high-frequency ST coefficients and a consistent verification procedure for optimization.

3.1. Fusion rules for low-frequency ST coefficients

The low-frequency coefficients express approximate information of an image because they contain the most of the energy. The sum-modified Laplacian (SML)-based method can be used to choose the coefficients in transform domain containing more detail information. To better fit the human visual system, SML is used to express the edge features and is regarded as the external inputs of DTNP systems. The SML is defined by

$$SML_{l_0}(i, j) = \sum_a \sum_b ML_{l_0}(i + a, j + b) \quad (4)$$

where sliding window size is 3×3 and ML is given by

$$ML_{l_0}(i, j) = |2C_{l_0}(i, j) - C_{l_0}(i - 1, j) - C_{l_0}(i + 1, j)| + |2C_{l_0}(i, j) - C_{l_0}(i, j - 1) - C_{l_0}(i, j + 1)| \quad (5)$$

where $C_{l_0}(i, j)$ is low-frequency ST coefficient at position (i, j) .

Suppose that Π_{AL} and Π_{BL} are two DTNP systems with local topology, which are associated with low-frequency coefficients of two multi-focus source images, A and B . Based on Eq. (4), we compute the SML values of low-frequency coefficients of source images A and B , denoted by $SML_{l_0}^{AL}(i, j)$ and $SML_{l_0}^{BL}(i, j)$, which are used as the external inputs of Π_{AL} and Π_{BL} respectively. Starting from initial states, the two DTNP systems work constantly until iteration number t_{max} reaches. Then, they halt. Denote by P_{AL} and P_{BL} the spike matrixes associated with Π_{AL} and Π_{BL} , i.e., $P_{AL} = (p_{l_0}^{AL}(i, j))_{m \times n}$ and $P_{BL} = (p_{l_0}^{BL}(i, j))_{m \times n}$, where $p_{l_0}^{AL}(i, j)$ (resp. $p_{l_0}^{BL}(i, j)$) is the cumulative value of spikes generated by neuron σ_{ij} in Π_{AL}

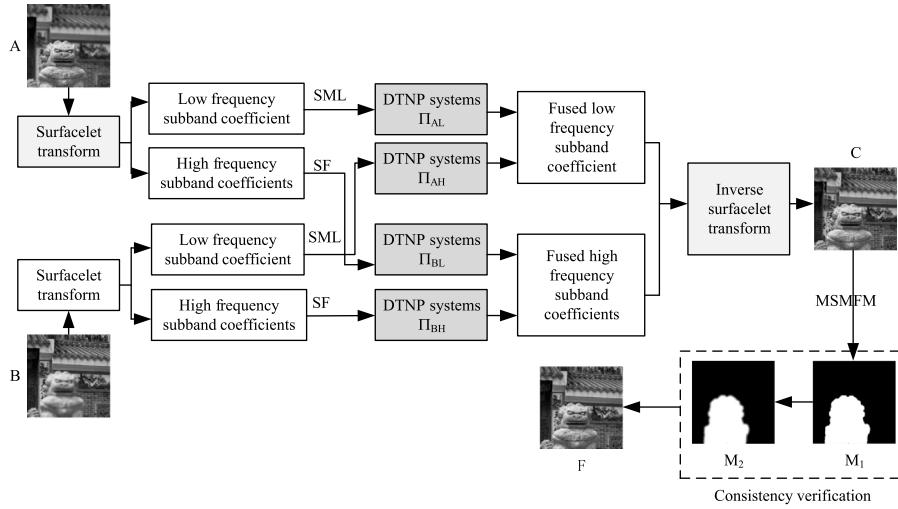


Fig. 4. The proposed fusion framework based on DTNP systems in ST domain, where A and B are two source images, C is initial fusion image, and F is final fusion image.

(resp. Π_{BL}). Based on the two spike matrixes, fusion rules of low-frequency ST coefficients are defined by

$$C_{l_0}^C(i, j) = \begin{cases} C_{l_0}^{AL}(i, j), & \text{if } p_{l_0}^{AL}(i, j) \geq p_{l_0}^{BL}(i, j) \\ C_{l_0}^{BL}(i, j), & \text{if } p_{l_0}^{AL}(i, j) < p_{l_0}^{BL}(i, j) \end{cases} \quad (6)$$

where $C_{l_0}^{AL}(i, j)$ and $C_{l_0}^{BL}(i, j)$ are low-frequency ST coefficients of multi-focus images at position (i, j) respectively, and $C_{l_0}^C(i, j)$ are low-frequency ST coefficients of the fused image C at position (i, j) , $1 \leq i \leq m$ and $1 \leq j \leq n$.

3.2. Fusion rules for high-frequency ST coefficients

Usually, the high-frequency coefficient with a large absolute value is associated with some extreme values, for example, edges and textures of an image. Spatial frequency can be used to express the spatial distribution of images. Therefore, we use spatial frequency to express the edge and texture features of images, which is regarded as the external inputs of DTNP systems. The spatial frequency (SF) is defined as follows

$$SF_{lr}(i, j) = \sqrt{RF_{lr}(i, j)^2 + CF_{lr}(i, j)^2} \quad (7)$$

where row-frequency (RF) and column-frequency (CF) are given by

$$RF_{lr}(i, j) = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [C_{lr}(i, j) - C_{lr}(i, j-1)]^2} \quad (8)$$

$$CF_{lr}(i, j) = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [C_{lr}(i, j) - C_{lr}(i-1, j)]^2} \quad (9)$$

where $C_{lr}(i, j)$ is the high-frequency coefficient in layer l and direction r at position (i, j) ; sliding window size is 3×3 ($M = N = 3$).

Suppose that Π_{AH} and Π_{BH} are two DTNP systems with local topology, which are associated with high-frequency coefficients of two multi-focus source images, A and B , respectively. Based on Eq. (7), we compute the SF values of high-frequency coefficients of source images A and B , denoted by $SF_{lr}^{AH}(i, j)$ and $SF_{lr}^{BH}(i, j)$, which are used as the external inputs of Π_{AH} and Π_{BH} respectively. Starting from initial states, the two DTNP systems work constantly until iteration number t_{max} reaches. Then, they halt. Denote by P_{AH} and P_{BH} the spike matrixes associated with Π_{AH}

and Π_{BH} , i.e., $P_{AH} = (p_{lr}^{AH}(i, j))_{m \times n}$ and $P_{BH} = (p_{lr}^{BH}(i, j))_{m \times n}$, where $p_{lr}^{AH}(i, j)$ (resp. $p_{lr}^{BH}(i, j)$) is the cumulative value of spikes generated by neuron σ_{ij} in Π_{AH} (resp. Π_{BH}). Based on the two spike matrixes, fusion rules of high-frequency ST coefficients are defined by

$$C_{lr}^C(i, j) = \begin{cases} C_{lr}^{AH}(i, j), & \text{if } p_{lr}^{AH}(i, j) \geq p_{lr}^{BH}(i, j) \\ C_{lr}^{BH}(i, j), & \text{if } p_{lr}^{AH}(i, j) < p_{lr}^{BH}(i, j) \end{cases} \quad (10)$$

where $C_{lr}^{AH}(i, j)$ and $C_{lr}^{BH}(i, j)$ denote high-frequency ST coefficients of multi-focus images in layer l and direction r at position (i, j) , and $C_{lr}^C(i, j)$ are high-frequency ST coefficient of the fused image C in layer l and direction r at position (i, j) .

3.2.1. Consistent verification

A common shortcoming in transform domain methods is that they may not be able to extract certain important details from multi-focus source images. To overcome this shortcoming, we adopt an optimization model.

Zhang et al. [61] proposed a multi-scale morphological focus measure (MSMFM) to evaluate fusion images. Suppose that $A(i, j)$ and $B(i, j)$ are the pixels of multi-focus source images A and B , at position (i, j) and $C(i, j)$ is a pixel in the initial fusion image C at position (i, j) . The computation details of the MSMFM can be found in Ref. [61]. Two difference images are given by

$$D_1 = \{D_1(i, j)\}_{m \times n}, \quad D_1(i, j) = C(i, j) - A(i, j) \quad (11)$$

$$D_2 = \{D_2(i, j)\}_{m \times n}, \quad D_2(i, j) = C(i, j) - B(i, j)$$

By using the MSMFM on D_1 and D_2 , a binary decision map $M_1 = \{M_1(i, j)\}_{m \times n}$ is given by

$$M_1(i, j) = \begin{cases} 1, & \text{if } \text{MSMFM}(D_1(i, j)) \leq \text{MSMFM}(D_2(i, j)) \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

Because there are some artifacts at the boundaries between focused and unfocused areas, we use a filter to improve the initial decision map M_1 . Specifically, we use an effective edge-preserving filter [48] that converts the structural information in a guided image into the filtering result for an input image. The initial fusion image is regarded as a guiding image to guide the filtering of the initial decision map M_1 . The filtering result is the final decision map, denoted by M_2 . The guided filter has two parameters: local window radius r' and regularization parameter ϵ . According to the decision map M_2 , the final fusion image $F = \{F(i, j)\}_{m \times n}$ can be constructed as follows:

$$F(i, j) = M_2(i, j)A(i, j) + (1 - M_2(i, j))B(i, j), \quad (13)$$



Fig. 5. The multi-focus image dataset used in our experiments.

Table 1
The sizes of images in dataset.

Images	Seascape	Temple	Balloon	Clock	Toy
Size	580 × 320	480 × 516	640 × 480	512 × 512	512 × 512
Images	Calendar	Wine	Newspaper	Leopard	Leaf
Size	260 × 180	256 × 256	320 × 240	480 × 360	268 × 204
Images	Flower	OpenGL	Corner	Pepsi	Craft
Size	512 × 384	512 × 384	320 × 240	512 × 512	160 × 160
Images	Girl	Lab	Desk	Hoed	Cameraman
Size	636 × 476	640 × 480	640 × 480	256 × 256	256 × 256

where $F(i, j)$ is the pixel in fusion image F at position (i, j) , and $A(i, j)$ and $B(i, j)$ are the pixels of source images A and B at position (i, j) , respectively. In this paper, we set $r' = 5$ and $\epsilon = 0.1$.

4. Experimental results

In our experiments, an open image dataset consisting of 20 pairs of multi-focus images was used as a set of test source images for evaluating the proposed and compared fusion methods. Fig. 5 shows the multi-focus source images, where for every pair of images, the left side is a near-focused image and the right side is a far-focused image. The source images all are gray images, and their sizes are listed in Table 1.

The proposed fusion method has been evaluated on the image dataset and compared with seven previous fusion methods and three deep-learning-based fusion methods. The compared fusion methods include curvelet transform (CVT) [21], multi-resolution singular value decomposition (MSVD) [62], Laplacian pyramid with sparse representation (LP-SR) [27], guided filtering based fusion (GFF) [3], dense scale invariant feature transform (DSIFT) [4], boundary finding (BF) [61], spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain (NSCT-PCNN) [63] and three deep-learning-based fusion methods (CNN [31], ECNN [34] and IFCNN [35]).

To compare the proposed fusion method with other fusion methods, we provide subjective and objective comparison results. The subjective comparison is visually on the images fused by these methods. Moreover, the objective evaluation is necessary

for the fused images. The objective evaluation metrics used in the experiments include Qabf [64], FMI [65], Qw [66], Qe [66] and CE [67].

The proposed fusion method is an DTNP-systems-based method in ST domain and it has six prior parameters: (i) size of the source images, $m \times n$; (ii) number of iterations, t_{max} ; (iii) initial threshold, τ_0 ; (iv) neighborhood radius, r ; (v) value of spikes generated by the firing rules at each neuron, p ; (vi) local weight matrix, $W_{r \times r}$. The proposed fusion method has been implemented in Matlab 2017b and on an Intel Core i7-6700 CPU running at 3.4 GHz with 16 GB of RAM. For the proposed method, $m \times n$ was determined according to the size of each test source image and we experimentally set $r = 3$, $\tau_0 = 1$, $\tau = 1$, $p = 1.1$ and $t_{max} = 200$ after repeated experiments, and local weight matrix is

$$W_{3 \times 3} = \begin{pmatrix} 0.7071 & 1 & 0.7071 \\ 1 & 0 & 1 \\ 0.7071 & 1 & 0.7071 \end{pmatrix}.$$

Parameters of surfacelet transform (ST) are given experimentally: three layers; $pyr_mode = 1$; $lev_array = [-1 0; 0 -1]$; $HGname = 'ntf'$ and $Varargin = 'bo'$. The compared methods were implemented on the same platform based on open codes and used the parameters in references. Three deep-learning-based fusion methods were implemented on Ubuntu 16.04 OS and on Inter Xeno Silver 4110 CPU 2.10 GHz and NVIDIA Tesla P100 with 16GB RAM by the open source codes provided by their authors.

4.1. Compared with seven fusion methods

Subjective evaluation was conducted by comparing the visual quality of the images merged by the presented and compared methods. Figs. 6–7 show the images fused by the proposed method and other seven methods on two pairs of multi-focus images respectively, namely “Clock” and “Pepsi”.

Fig. 6 compares the fusion images generated by the proposed method and other seven methods for “Clock” images with size 512×512 , where (a) and (b) are source multi-focus images. It can be observed from Fig. 6(c)–(i) that some blurs exist in the images

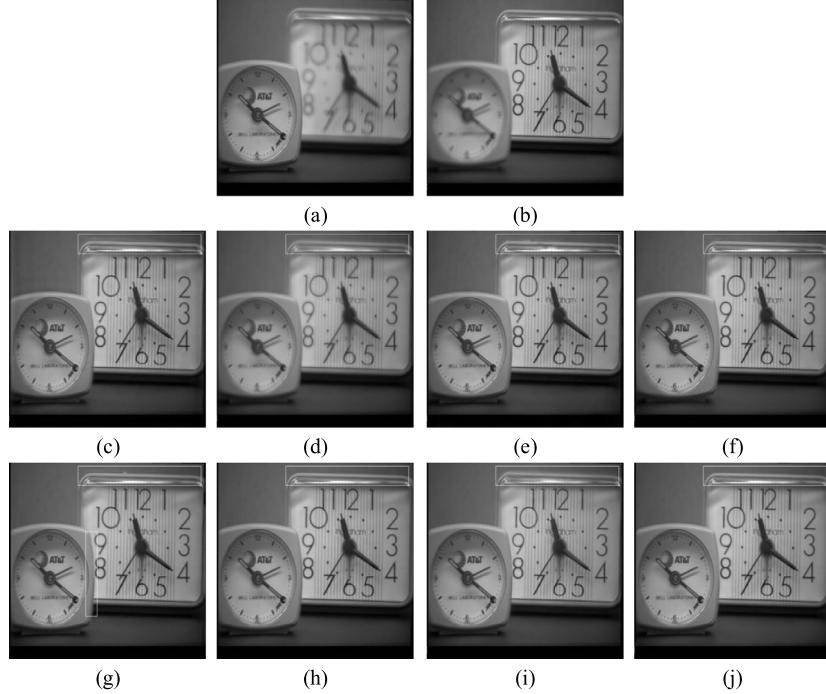


Fig. 6. Source images and fusion results of the proposed and compared methods on the “Clock” image: (a) and (b) are source images, and (c)–(j) are the fusion results of CVT, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN and the proposed method, respectively.

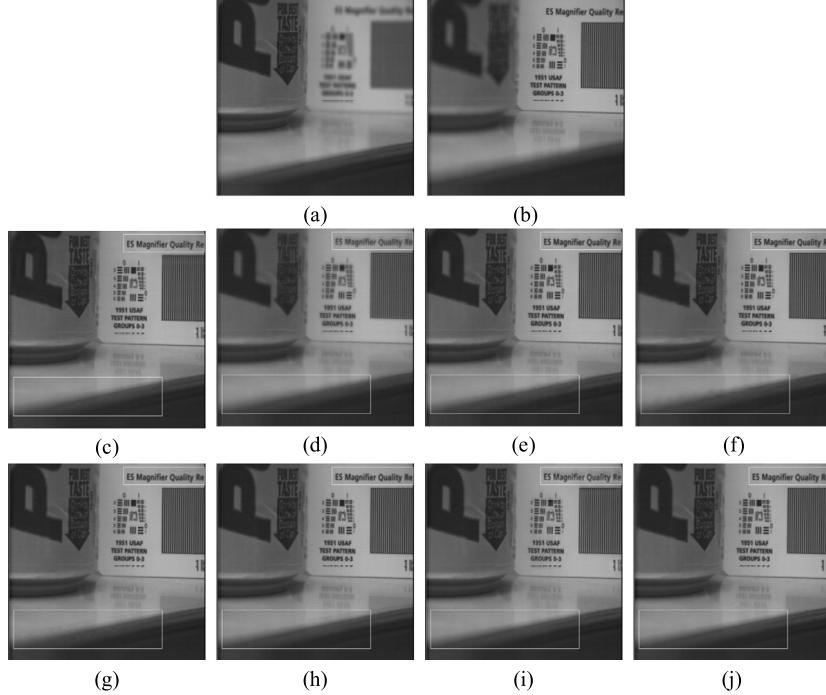


Fig. 7. Source images and fusion results of the proposed and compared methods on the “Pepsi” image: : (a) and (b) are source infrared and visible images, and (c)–(j) are the fusion results of CVT, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN and the proposed method, respectively.

fused by CVT, MSVD, LP-SR, GFF, DSIFT, NSCT-PCNN (see white rectangle box). We can also find that the image fused by MSVD has obvious overall blur and low contrast; some information is missing at the junction of the big clock and the small clock in the image merged by DSIFT. The proposed method and BF have good visual effect, including high clarity and contrast.

Fig. 7 compares the fusion images generated by the proposed method and other seven methods for “Pepsi” images with size

512×512 , where (a) and (b) are multi-focus source images. From **Fig. 7**, one can observe that the image fused by CVT has low contrast; the image fused by MSVD has severe blurry ghosting; the images fused by GFF and DSIFT have a lot of noise in large box, and the image fused by DSIFT has some noise in small box; the image fused by NSCT-PCNN has some ghosting in small box. The proposed method, LP-SR and BF have good visual effect. However, it is difficult to distinguish which methods performed

Table 2

The comparison results of the proposed method with seven previous fusion methods on source images "Clock" and "Pepsi" in terms of five metrics.

Images	Methods	Qabf	FMI	Qw	Qe	CE
Clock	CVT	0.6674	0.9184	0.8848	0.5386	0.0849
	MSVD	0.5695	0.9048	0.8005	0.3436	0.0309
	LP-SR	0.6942	0.9232	0.8175	0.5564	0.1053
	GFF	0.7169	0.9225	0.9068	0.5619	0.0217
	DSIFT	0.7178	0.9249	0.8929	0.5681	0.0168
	BF	0.7174	0.9249	0.8620	0.5611	0.0177
	NSCT-PCNN	0.6685	0.9174	0.8877	0.4949	0.0342
Pepsi	Proposed	0.7178	0.9257	0.8873	0.5782	0.0177
	CVT	0.7657	0.9221	0.9636	0.6020	0.0286
	MSVD	0.6733	0.9089	0.8966	0.4824	0.0273
	LP-SR	0.7695	0.9238	0.9640	0.5585	0.0327
	GFF	0.7833	0.9245	0.9644	0.5318	0.0244
	DSIFT	0.7819	0.9246	0.9627	0.5981	0.0266
	BF	0.7890	0.9247	0.9633	0.6002	0.0280
Dataset	NSCT-PCNN	0.7637	0.9194	0.9605	0.5626	0.0287
	Proposed	0.7899	0.9247	0.9634	0.6066	0.0274

Table 3

The comparison results of the proposed method with seven previous fusion methods on dataset in terms of five metrics.

Images	Methods	Qabf	FMI	Qw	Qe	CE
Dataset	CVT	0.6848	0.8656	0.9096	0.5416	0.0960
	MSVD	0.5406	0.8475	0.8142	0.3918	0.0970
	LP-SR	0.7097	0.8675	0.9097	0.5618	0.0856
	GFF	0.7266	0.8697	0.9132	0.5833	0.0761
	DSIFT	0.7274	0.8701	0.9101	0.5935	0.0710
	BF	0.7273	0.8703	0.9052	0.5897	0.0705
	NSCT-PCNN	0.6843	0.8651	0.9107	0.5266	0.0906
Dataset	Proposed	0.7277	0.8704	0.9096	0.5933	0.0701

the best or worst visually. Therefore, an objective assessment was performed.

Table 2 shows quantitative comparison of eight fusion methods in terms of five metrics for "Clock" and "Pepsi" images, respectively. In **Table 2**, bold font marks that the corresponding method reaches the best metric value. For "Clock" image, it can be observed from **Table 2** that:

- (1) the proposed fusion method obtains the best value in terms of four metrics, i.e., Qabf = 0.7178, FMI = 0.9257, Qe = 0.5782 and CE = 0.0177.
- (2) GFF gets the best value in terms of Qw metric, Qw = 0.9068, and the proposed fusion method achieves the better value, Qw = 0.8873.

For "Pepsi" image, one can see that:

- (1) the proposed fusion method obtains the best value in terms of three metrics, i.e., Qabf = 0.7899, FMI = 0.9247 and Qe = 0.6066.
- (2) BF achieves the best value in terms of FMI metric, FMI = 0.9247.
- (3) GFF obtains the best value in terms of Qw metric, Qw = 0.9644, and the proposed fusion method achieves the better value, Qw = 0.9634.
- (4) GFF obtains the best value in terms of CE metric, CE = 0.0244, and the proposed fusion method achieves the better value, CE = 0.0274.

Table 3 lists the results for the average fusion performance of the proposed method and seven compared methods in terms of the five metrics. The result for a given method and index represents the average value for 20 pairs of multi-focus images in the dataset. In **Table 3**, one can observe the following results:

Table 4

The comparison results of the proposed method with three deep-learning-based fusion methods on source images "Street" and "Tank" in terms of five metrics.

Images	Methods	Qabf	FMI	Qw	Qe	CE
Wine	CNN	0.6966	0.8473	0.8840	0.5365	0.0301
	ECNN	0.6960	0.8480	0.8807	0.5356	0.0287
	IFCNN-Max	0.6444	0.8371	0.8995	0.5050	0.2230
	Proposed	0.6979	0.8478	0.8847	0.5381	0.0280
Girl	CNN	0.6810	0.8219	0.8687	0.5464	0.0131
	ECNN	0.6806	0.8214	0.8662	0.5462	0.0131
	IFCNN-Max	0.6154	0.8099	0.8691	0.4074	0.0187
	Proposed	0.6812	0.8219	0.8696	0.5461	0.0132

- (i) the proposed fusion method achieves the best value for three metrics, i.e., Qabf = 0.7277, FMI = 0.8704 and CE = 0.0701.
- (ii) for Qw metric, GFF achieves the best value of Qw = 0.9132, and the proposed method achieves the better value, Qw = 0.9096.
- (iii) for Qe metric, DSIFT achieves the best value of Qe = 0.5935, and the proposed method achieves the next best value, Qe = 0.5933.

4.2. Compared with deep-learning-based fusion methods

Recently, deep-learning-based fusion methods have been developed and showed a good fusion effect. The proposed method was compared with three deep-learning-based methods, CNN, ECNN and IFCNN-Max. The proposed and compared methods were evaluated on 20 pairs of multi-focus images in the dataset. **Figs. 8–9** show the images fused by the proposed method and three deep-learning-based methods on two pairs of multi-focus images respectively, namely "Wine" and "Girl".

From **Figs. 8(c)–(f)** and **9(c)–(f)**, we can observe that the proposed method and three deep-learning-based methods can achieve visually good fusion effect, and it is difficult to distinguish which methods performed the best or worst visually. Thus, we compute the difference images between each source image and the images fused by the proposed method and three deep-learning-based methods, shown in **Figs. 8(g)–(j)** and **9(g)–(j)**, respectively. It can be found from the difference images that the images fused by IFCNN-Max, shown in **Figs. 8(i)** and **9(i)**, contain much redundant information. This indicates that IFCNN-Max does not merge more information in source images into the fusion result. However, we could not visually determine which method performed the best or worst among other three methods. Therefore, an objective assessment was performed.

Table 4 shows quantitative comparison of four fusion methods in terms of five metrics for "Wine" and "Girl" images, respectively. In **Table 4**, bold font marks that the corresponding method reaches the best metric value. For "Wine" image, it can be observed from **Table 4** that:

- (1) the proposed fusion method obtains the best value in terms of three metrics, i.e., Qabf = 0.6979, Qe = 0.5381 and CE = 0.0280.
- (2) for FMI metric, ECNN obtains the best value of FMI = 0.8480, and the proposed method achieves the next best value of FMI = 0.8478.
- (3) for Qw metric, IFCNN-Max obtains the best value of Qw = 0.8995, and the proposed method achieves the next best value of Qw = 0.8847.

For "Girl" image, one can see that:

- (1) the proposed fusion method obtains the best value in terms of three metrics, i.e., Qabf = 0.6812, FMI = 0.8219 and Qw = 0.8696.

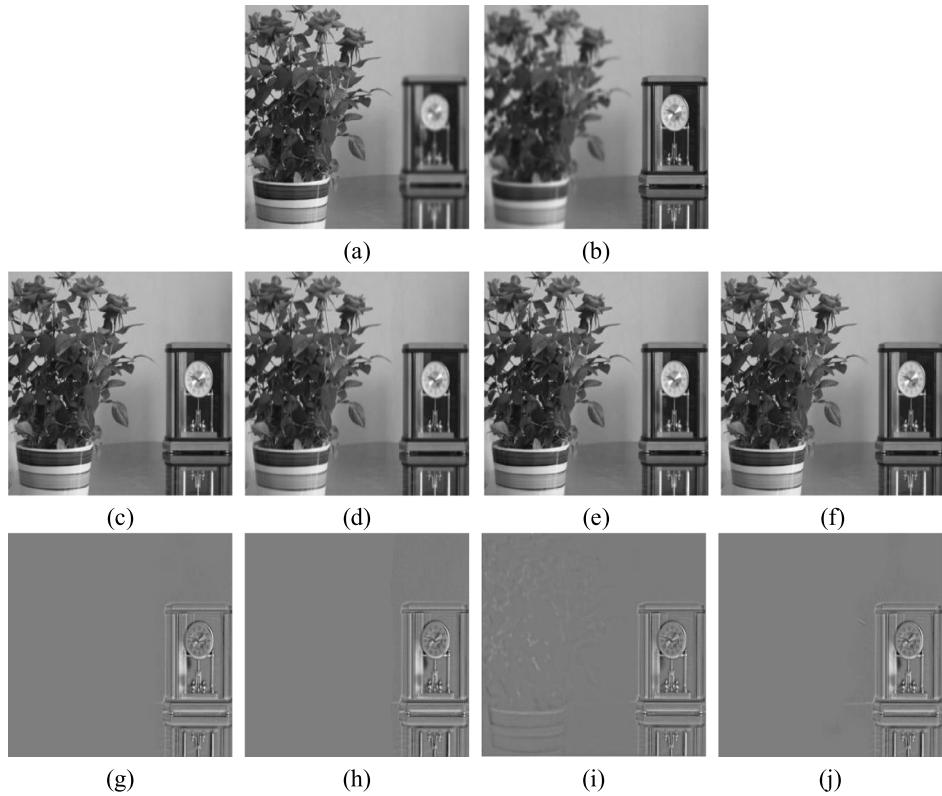


Fig. 8. Source images and fusion results of the proposed method and deep-learning-based methods on the “Wine” image: (a) and (b) are multi-focus source images; (c)–(f) are the fusion results of CNN, ECNN, IFCNN-Max and the proposed method respectively; (g)–(j) are the difference images of source image and CNN, ECNN, IFCNN-Max and the proposed method respectively.

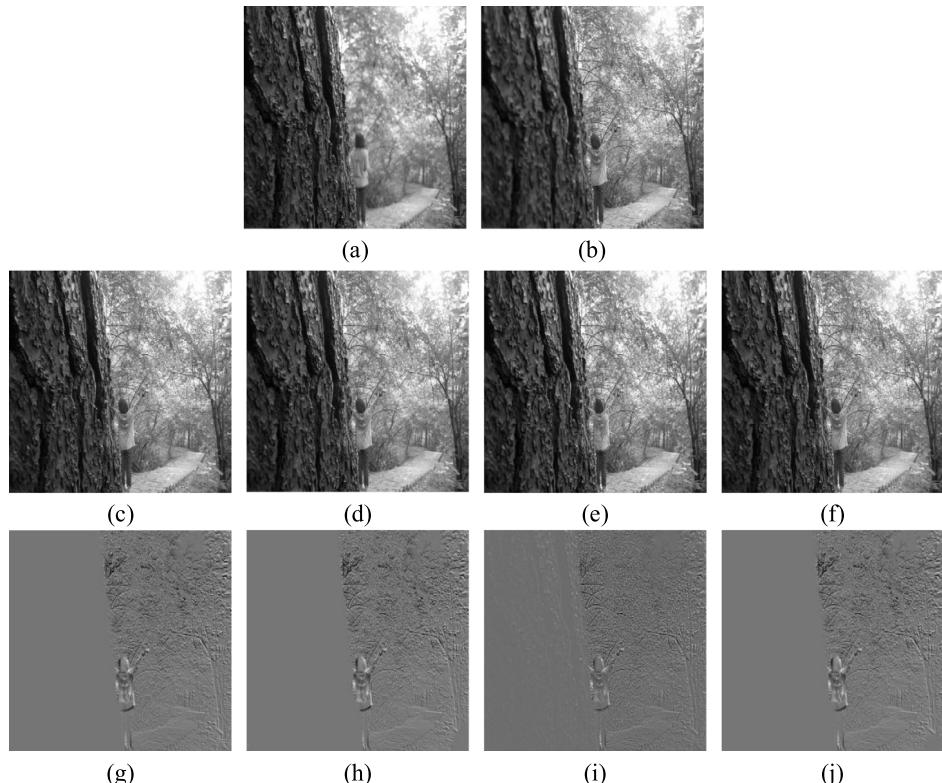


Fig. 9. Source images and fusion results of the proposed method and deep-learning-based methods on the “Girl” image: (a) and (b) are multi-focus source images; (c)–(f) are the fusion results of CNN, ECNN, IFCNN-Max and the proposed method respectively; (g)–(j) are the difference images of source image and CNN, ECNN, IFCNN-Max and the proposed method respectively.



Fig. 10. The fusion results of the proposed method on other multi-focus images, including “Seascape”, “Temple”, “Balloon”, “Toy”, “Calendar”, “Newspaper”, “Leopard”, “Leaf”, “Flower”, “OpenGL”, “Corner”, “Craft”, “Lab”, “Desk”, “Hoed” and “Cameraman”.

Table 5

The comparison results of the proposed method with three deep-learning-based fusion methods on dataset in terms of five metrics.

Images	Methods	Qabf	FMI	Qw	Qe	CE
Database	CNN	0.7273	0.8703	0.9099	0.5892	0.0724
	ECNN	0.7195	0.8685	0.8996	0.5793	0.0708
	IFCNN-Max	0.6725	0.8595	0.9089	0.5076	0.1085
	Proposed	0.7277	0.8704	0.9096	0.5933	0.0701

- (2) CNN obtains the best value in terms of FMI metric, $FMI = 0.8219$.
- (3) for Qe metric, CNN obtains the best value of $Qe = 0.5464$, and the proposed method achieves the better value, $Qe = 0.5461$.
- (4) for CE metric, CNN and ECNN obtain the best value of $CE = 0.0131$, and the proposed method achieves the next best value of $CE = 0.0132$.

Table 5 lists the results for the average fusion performance of the proposed method and three compared methods in terms of five metrics. The result for a given method and index represents the average value for all 20 pairs of multi-focus images in the dataset. In **Table 5**, one can see that the proposed fusion method achieves the best value for four metrics, i.e., $Qabf = 0.7277$, $FMI = 0.8704$, $Qe = 0.5933$ and $CE = 0.0701$; for Qw metric, CNN obtains the best value of $Qw = 0.9099$, and the proposed method achieves the next best value of $Qw = 0.9096$.

In addition, we provide the fusion results of the proposed method on other sixteen multi-focus images, including “Seascape”, “Temple”, “Balloon”, “Toy”, “Calendar”, “Newspaper”, “Leopard”, “Leaf”, “Flower”, “OpenGL”, “Corner”, “Craft”, “Lab”, “Desk”, “Hoed” and “Cameraman” in **Fig. 10**. These fusion results

indicate that the proposed method has good fusion effect from a visual perspective.

4.3. Computational efficiency

To evaluate the computational efficiency, we computed the average execution times of the proposed and compared methods for 20 pairs of multi-focus images in the dataset. Since deep-learning-based methods require a time-consuming training process, CNN, ECNN and IFCNN-Max were not included in this comparison. **Table 6** lists comparison results for the proposed method with other seven methods in terms of average execution time. In **Table 6**, one can observe the following results: (i) CTV, MSVD, LP-SR, GFF, DSIFT and BF have low execution time, (ii) NSCT-PCNN has high execution time, (iii) the proposed method has relatively high execution time. The comparative results illustrate that the proposed method has a relatively high computational cost. However, the proposed method has very high computational efficiency compared to deep-learning-based methods.

5. Conclusions and discussion

Dynamic threshold neural P systems (DTNP systems) are a kind of distributed and parallel computing models and Turing-universal computing devices. This paper developed a novel DTNP-systems-based fusion method in ST domain for multi-focus images. DTNP systems with local topology were designed to propose a fusion framework for multi-focus images. In the fusion framework, fusion rules of low-frequency and high-frequency ST coefficients are controlled by four DTNP systems that are associated with low-frequency or high-frequency ST coefficients in two multi-focus images respectively.

Table 6

The comparison of average execution time of the proposed approach and other approaches on dataset.

Methods	CVT	MSVD	LP-SR	GFF
Times (s)	0.8918	0.1941	0.0294	0.1874
Methods	DSIFT	BF	NSCT-PCNN	Proposed
Times (s)	4.5603	1.3201	128.5321	23.1298

In experiments, three groups of comparisons were designed to contribute the availability and effectiveness of the proposed fusion method: (i) the comparison to seven previous fusion methods; (ii) the comparison to three deep-learning-based fusion methods; (iii) the comparison on execution time. For the first two groups of comparisons, subjective and quantitative evaluations were considered, while the third group of comparisons was used to evaluate the computational efficiency. The experimental results have demonstrated that the proposed fusion method is a competitive method for the fusion of multi-focus images. For the reasons, this is due to the following two aspects:

- (1) ST transform can capture more details of the image and provide rich complementary information, which is beneficial for image fusion.
- (2) DTNP systems are inspired from the spike firing mechanism and dynamic threshold mechanism of neurons. In ST domain, the rich complementary information can trigger more neurons in local region to fire. As a result, the corresponding region in the fused image will become clearer.

Therefore, the combination of these two effects makes the proposed fusion method show good fusion performance on multi-focus images.

This work mainly focuses on the fusion of multi-focus images, which are optical images. In addition to optical images, there are other kinds of images, for example, infrared images, medical images (such as CT images and PET images), and spectral images. These images have different characteristics because they are captured by different types of imaging devices. Experimental results in this paper only demonstrate the effectiveness of DTNP systems on multi-focus images. Are DTNP systems also effective for other kinds of images with different characteristics? However, the answer is unknown. Our future work will attempt to apply DTNP systems to deal with the fusion of these kinds of images.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Bo Li: Conceptualization, Software, Writing - original draft. **Hong Peng:** Conceptualization, Software, Writing - original draft. **Jun Wang:** Conceptualization, Writing - original draft. **Xiangnian Huang:** Software.

Acknowledgments

This work was partially supported by the Research Fund of Sichuan Science and Technology Project, China (No. 2018JY0083), Research Foundation of the Education Department of Sichuan province, China (No. 17TD0034), and the Innovation Fund of Postgraduate, Xihua University, China (Nos. YCJJ2019019 and YCJJ2019020), China.

References

- [1] B. Meher, S. Agrawal, R. Panda, A. Abraham, A survey on region based image fusion methods, *Inf. Fusion* 48 (2019) 119–132.
- [2] T. Stathaki, *Image Fusion: Algorithms and Applications*, Academic Press Elsevier, 2011.
- [3] S. Li, X. Kang, J. Hu, Image fusion with guided filtering, *IEEE Trans. Image Process.* 22 (7) (2013) 2864–2875.
- [4] Y. Liu, S. Liu, Z. Wang, Multi-focus image fusion with dense SIFT, *Inf. Fusion* 23 (2015) 139–155.
- [5] I. De, B. Chanda, Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure, *Inf. Fusion* 14 (2) (2013) 136–146.
- [6] M. Li, W. Cai, Z. Tan, A region-based multi-sensor image fusion scheme using pulse coupled neural network, *Pattern Recognit. Lett.* 27 (16) (2006) 1948–1956.
- [7] S. Li, B. Yang, Multifocus image fusion using region segmentation and spatial frequency, *Image Vis. Comput.* 26 (7) (2008) 971–979.
- [8] W. Huang, Z. Jing, Multi-focus image fusion using pulse coupled neural network, *Pattern Recognit. Lett.* 28 (2007) 1123–1132.
- [9] Z. Wang, Y. Ma, J. Gu, Multi-focus image fusion using PCNN, *Pattern Recognit.* 43 (6) (2010) 2003–2016.
- [10] Y. Zhang, L. Chen, Z. Zhao, J. Jia, J. Liu, Multi-focus image fusion based on robust principal component analysis and pulse-coupled neural network, *Optik* 125 (2014) 5002–5006.
- [11] P. Burt, E. Adelson, The laplacian pyramid as a compact image code, *IEEE Trans. Comput.* 31 (4) (1983) 532–540.
- [12] A. Toe, Image fusion by a ratio of low-pass pyramid, *Pattern Recognit. Lett.* 9 (1989) 245–253.
- [13] D.X. He, Y. Meng, C.Y. Wang, Contrast pyramid based image fusion scheme for infrared image and visible image, in: Proceedings of 2011 IEEE International Conference on Geoscience & Remote Sensing Symposium, 2011, pp. 597–600.
- [14] V.S. Petrovic, C.S. Xydeas, Gradient-based multiresolution image fusion, *IEEE Trans. Image Process.* 13 (2) (2004) 228–237.
- [15] Z.Q. Zhou, S. Li, B. Wang, Multi-scale weighted gradient-based fusion for multi-focus images, *Inf. Fusion* 20 (1) (2014) 60–72.
- [16] P. Zeeuw, *Wavelets and Image Fusion*, CWI, Amsterdam, 1998.
- [17] J. Tian, L. Chen, Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure, *Signal Process.* 92 (9) (2012) 2137–2146.
- [18] R. Redondo, F. Sroubek, S. Fischer, G. Cristóbal, Multifocus image fusion using the log-gabor transform and a multisize windows technique, *Inf. Fusion* 10 (2) (2009) 163–171.
- [19] G. Pajares, J.M. de la Cruz, A wavelet-based image fusion tutorial, *Pattern Recognit.* 37 (9) (2004) 1855–1872.
- [20] J.J. Lewis, R.J. O'Callaghan, S.G. Nikolov, D.R. Bull, N. Canagarajah, Pixel- and region-based image fusion with complex wavelets, *Inf. Fusion* 8 (2) (2007) 119–130.
- [21] F. Nencini, A. Garzelli, S. Baronti, L. Alparone, Remote sensing image fusion using the curvelet transform, *Inf. Fusion* 8 (2007) 143–156.
- [22] S. Li, B. Yang, Multifocus image fusion by combining curvelet and wavelet transform, *Pattern Recognit. Lett.* 29 (9) (2008) 1295–1301.
- [23] Q. Zhang, B. Guo, Multifocus image fusion using the nonsubsampled contourlet transform, *Signal Process.* 89 (7) (2009) 1334–1346.
- [24] L. Yang, B. Guo, W. Ni, Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform, *Neurocomputing* 72 (1–3) (2008) 203–211.
- [25] X. Jin, G. Chen, J. Hou, Q. Jiang, D. Zhou, S. Yao, Multimodal sensor medical image fusion based on nonsubsampled shearlet transform and S-PCNNs in HSV space, *Signal Process.* 153 (2018) 379–395.
- [26] M. Yin, P. Duan, W. Liu, X. Liang, A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation, *Neurocomputing* 226 (2017) 182–191.
- [27] Y. Liu, S. Liu, Z. Wang, A general framework for image fusion based on multi-scale transform and sparse representation, *Inf. Fusion* 24 (2015) 147–164.
- [28] M. Nejati, S. Samavi, S. Shirani, Multi-focus image fusion using dictionary-based sparse representation, *Inf. Fusion* 25 (2015) 72–84.
- [29] Q. Zhang, Y. Liu, R.S. Blum, J. Han, D. Tao, Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review, *Inf. Fusion* 40 (2018) 57–75.
- [30] B. Zhang, C. Zhang, Y. Liu, J. Wu, H. Liu, Multi-focus image fusion algorithm based on compound PCNN in surfacelet domain, *Optik* 125 (2014) 296–300.
- [31] Y. Liu, X. Chen, H. Peng, Z. Wang, Multi-focus image fusion with a deep convolutional neural network, *Inf. Fusion* 36 (2017) 191–207.
- [32] H. Tang, B. Xiao, W. Li, G. Wang, Pixel convolutional neural network for multi-focus image fusion, *Inform. Sci.* 433 (2017) 125–141.
- [33] X. Gao, R. Nie, J. Cao, D. Zhou, W. Qian, Fully convolutional network-based multi-focus image fusion, *Neural Comput.* 30 (7) (2018) 1775–1800.

- [34] M. Amin-Naji, A. Aghagolzadeh, M. Ezoji, Ensemble of CNN for multi-focus image fusion, *Inf. Fusion* 51 (2019) 201–214.
- [35] Y. Zhang, Y. Liu, IFCNN: a general image fusion framework based on convolutional neural network, *Inf. Fusion* 54 (2020) 99–118.
- [36] H. Peng, J. Wang, M.J. Pérez-Jiménez, A. Riscos-Núñez, Dynamic threshold neural P systems, *Knowl.-Based Syst.* 163 (2019) 875–884.
- [37] M. Ionescu, Gh. Păun, T. Yokomori, Spiking neural P systems, *Fund. Inform.* 71 (2006) 279–308.
- [38] Gh. Păun, Spiking neural P systems with astrocyte-like control, *J. UCS* 13 (11) (2007) 1707–1721.
- [39] L. Pan, Gh. Păun, Spiking neural p systems with anti-spikes, *Int. J. Comput. Commun. Control IV* (3) (2009) 273–282.
- [40] L. Pan, Gh. Păun, G. Zhang, F. Neri, Spiking neural P systems with communication on request, *Int. J. Neural Syst.* 28 (8) (2017) 1750042, 1–13.
- [41] H. Peng, J. Yang, J. Wang, T. Wang, Z. Sun, X. Song, X. Luo, X. Huang, Spiking neural P systems with multiple channels, *Neural Netw.* 95 (2017) 66–71.
- [42] T. Song, J. Xu, L. Pan, On the universality and non-universality of spiking neural P systems with rules on synapses, *IEEE Trans. Nanobiosci.* 14 (8) (2015) 960–966.
- [43] T. Wu, A. Păun, Z. Zhang, L. Pan, Spiking neural P systems with polarizations, *IEEE Trans. Neural Netw. Learn. Syst.* 29 (8) (2018) 3349–3360.
- [44] F.G.C. Cabarle, H.N. Adorna, M.J. Pérez-Jiménez, T. Song, Spiking neural P systems with structural plasticity, *Neural Comput. Appl.* 26 (8) (2015) 1905–1917.
- [45] H. Peng, J. Wang, Coupled neural P systems, *IEEE Trans. Neural Netw. Learn. Syst.* 30 (6) (2019) 1672–1682.
- [46] H. Peng, B. Li, J. Wang, X. Song, T. Wang, L. Valencia-Cabrera, I. Pérez-Hurtado, A. Riscos-Núñez, M.J. Pérez-Jiménez, Spiking neural P systems with inhibitory rules, *Knowl.-Based Syst.* 30 (2) (2020) 2050008, 1–17.
- [47] H. Peng, Z. Lv, B. Li, X. Luo, J. Wang, X. Song, T. Wang, M.J. Pérez-Jiménez, A. Riscos-Núñez, Nonlinear spiking neural P systems, *Int. J. Neural Syst.* (2020) Available at <https://doi.org/10.1142/S0129065720500082>.
- [48] O.H. Ibarra, A. Păun, A. Rodríguez-Páton, Sequential SNP systems based on min/max spike number, *Theoret. Comput. Sci.* 410 (30) (2009) 2982–2991.
- [49] M. Cavaliere, O.H. Ibarra, Gh. Păun, O. Egecioglu, M. Ionescu, S. Woodworth, Asynchronous spiking neural P systems, *Theoret. Comput. Sci.* 410 (24) (2009) 2352–2364.
- [50] H. Peng, J. Wang, M.J. Pérez-Jiménez, H. Wang, J. Shao, T. Wang, Fuzzy reasoning spiking neural P system for fault diagnosis, *Inform. Sci.* 235 (20) (2013) 106–116.
- [51] J. Wang, P. Shi, H. Peng, M.J. Pérez-Jiménez, T. Wang, Weighted fuzzy spiking neural P system, *IEEE Trans. Fuzzy Syst.* 21 (2) (2013) 209–220.
- [52] J. Wang, H. Peng, W. Yu, J. Ming, M.J. Pérez-Jiménez, C. Tao, X. Huang, Interval-valued fuzzy spiking neural P systems for fault diagnosis of power transmission networks, *Eng. Appl. Artif. Intell.* 82 (2019) 102–109.
- [53] H. Peng, J. Wang, J. Ming, P. Shi, M.J. Pérez-Jiménez, W. Yu, C. Tao, Fault diagnosis of power systems using intuitionistic fuzzy spiking neural P systems, *IEEE Trans. Smart Grid* 9 (5) (2018) 4777–4784.
- [54] T. Wu, F.D. Bîlbiie, A. Păun, L. Pan, F. Neri, Simplified and yet tuning universal spiking neural p systems with communication on request, *Int. J. Neural Syst.* 28 (8) (2018) 1850013, 1–19.
- [55] H. Chen, R. Freund, M. Ionescu, Gh. Păun, M.J. Pérez-Jiménez, On string languages generated by spiking neural P systems, *Fund. Inform.* 75 (1) (2007) 141–162.
- [56] A. Păun, Gh. Păun, Small universal spiking neural p systems, *BioSystems* 90 (1) (2007) 48–60.
- [57] L. Pan, Gh. Păun, M.J. Pérez-Jiménez, Spiking neural p systems with neuron division and budding, *Sci. China Inf. Sci.* 54 (8) (2011) 1596–1607.
- [58] A. Leporati, G. Mauri, C. Zandron, Gh. Păun, M.J. Pérez-Jiménez, Uniform solutions to SAT and subset sum by spiking neural P systems, *Nat. Comput.* 8 (4) (2009) 681–702.
- [59] R.H. Bamberger, M.J. Smith, A filter bank for the directional decomposition of images: theory and design, *IEEE Trans. Signal Process.* 40 (1992) 882–893.
- [60] Y.M. Yu, M.N. Do, Multidimensional directional filter banks and surfacelets, *IEEE Trans. Image Process.* 16 (2007) 918–931.
- [61] Y. Zhang, X. Bai, T. Wang, Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure, *Inf. Fusion* 35 (2017) 81–101.
- [62] V. Naidu, Image fusion technique using multi-resolution singular value decomposition, *Def. Sci. J.* 61 (2011) 479–484.
- [63] X. Qu, J. Yan, H. Xiao, Z. Zhu, Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain, *Acta Automat. Sinica* 34 (12) (2008) 1508–1514.
- [64] C. Xydeas, V. Petrovic, Objective image fusion performance measure, *Electron. Lett.* 36 (4) (2009) 308–309.
- [65] M.B.A. Haghighat, A. Aghagolzadeh, H. Seyedarabi, A non-reference image fusion metric based on mutual information of image features, *Comput. Electr. Eng.* 37 (5) (2011) 744–756.
- [66] G. Piella, H. Heijmans, A new quality metric for image fusion, in: Proceedings of the International Conference on Image Processing, 2003, pp. 173–176.
- [67] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, second ed., John Wiley & Sons, 2006.