

Multi-focus image fusion approach based on CNP systems in NSCT domain

Hong Peng ^{a,*}, Bo Li ^a, Qian Yang ^a, Jun Wang ^b



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

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

ARTICLE INFO

Communicated by Nikos Paragios

Keywords:

Coupled neural p systems
Nonsubsampled contourlet transform
Image fusion
Multi-focus image
Fusion rule

ABSTRACT

Coupled neural P (CNP) systems are recently developed distributed and parallel computing models that are abstracted by the mechanisms of coupled and spiking neurons. CNP systems differ from spiking neural P (SNP) systems in two main ways, namely the utilization of three data units, and a coupled firing and dynamic threshold mechanism for neurons. This paper focuses on the application of CNP systems to solve multi-focus image fusion problems, and proposes a novel image fusion approach based on CNP systems. Based on two CNP systems with local topology, a multi-focus image fusion framework in the non-subsampled contourlet transform (NSCT) domain is developed, where the two CNP systems are utilized to control the fusion of low-frequency coefficients in the NSCT domain. The proposed fusion approach is evaluated on an open data set of 19 multi-focus images based on five fusion quality indices, and compared to 11 state-of-the-art fusion approaches. Quantitative and qualitative experimental results demonstrate the advantages of the proposed fusion approach in terms of visual quality and fusion performance.

1. Introduction

Because optical imaging devices have a limited depth of field, the images captured by such devices are not all in focus. Objects at a specific depth of field are in focus, while other objects are blurred. Some local blurred losses can result in the failure of certain image processing tasks and generate imperfect results. Multi-focus image fusion is an effective technique for resolving these issues. Multi-focus image fusion is a procedure in which two or more images, with different depths of field, are merged to generate a single composite image. The fused image contains detailed information that is more suitable for human visual systems. Image fusion techniques have been successfully implemented in imaging technology, computer vision, remote sensors, and other methods.

Many multi-focus image fusion approaches have been proposed in recent years. These fusion approaches can be divided into two main categories: spatial domain approaches and multi-scale transform (MST) approaches (Li et al., 2017; Meher et al., 2019). In addition, several fusion approaches based on convolutional neural networks (CNN) have received attention in recent years (Liu et al., 2018). However, the existing fusion approaches showed significant difference in fusion performance.

1.1. Related work

The existing fusion approaches have two main categories: spatial domain approaches and transform domain approaches (Stathaki, 2011;

Liu et al., 2020a,b). Spatial domain approaches operate directly on multi-focus source images, and do not have to convert images into other types of expressions. Such approaches can be further classified into two groups: pixel- and region/block-level approaches (Li et al., 2013; Liu et al., 2015b). Pixel-level approaches fuse 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/block-level approaches partition source images into regions or blocks, and then apply various sharpness measures (such as spatial frequency or gradients) to select regions or blocks for merging (De and Chanda, 2013; Li et al., 2006; Li and Yang, 2008). Additionally, pulse coupled neural networks (PCNNs) have been utilized for multi-focus image fusion (Huang and Jing, 2007; Wang et al., 2010; Zhang et al., 2014a).

In recent years, transform domain approaches have received significant attention. They consist of three main steps. First, the source images are converted into a transform domain. Second, coefficients in the transform domain are merged to produce the fused coefficients based on fusion rules. Finally, the fused coefficients are converted back into a spatial domain to form a composite image via inverse transform. Multi-scale transform has been widely used in image fusion based on its excellent locality and multi-resolution features, such as the Laplacian pyramid (LP) (Burt and Adelson, 1983; Toe, 1989; He et al., 2011), gradient pyramid (Petrovic and Xydeas, 2004; Zhou et al., 2014), and wavelet transform (Zeeuw, 1998). The wavelet transform has become a popular fusion method in the transform domain approaches, including

* Corresponding author.

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

the discrete wavelet transform (DWT) (Tian and Chen, 2012; Redondo et al., 2009; Pajares and de la Cruz, 2004) and dual-tree complex wavelet transform (Lewis et al., 2007). 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 the curvelet transform (citenencini-gba07,li-y08-1), surfacelet transform (Zhang et al., 2014b; Li et al., 2020), non-subsampled shearlet transform (Jin et al., 2018), shift-invariant dual-tree complex shearlet transform (Yin et al., 2017), and sparse representations (Liu et al., 2015a; Nejati et al., 2015; Zhang et al., 2018). Currently, image fusion based on the non-subsampled contourlet transform (NSCT) provides some excellent properties (Zhang and Guo, 2009; Yang et al., 2008), such as shift invariance and the ability to capture higher dimensions of singularities (such as direction, edge, and detail information). This study aims to develop a novel image fusion approach in the NSCT domain.

In recent years, several convolutional neural network (CNN)-based fusion approaches have been developed for the fusion of multi-focus images. (Liu et al., 2017) discussed a CNN-based multi-focus image fusion approach where a CNN-based model was considered to solve classification problems. (Tang et al., 2017) investigated a pixel CNN for the fusion of multi-focus images, where a model was trained to learn the probabilities of focused, defocused, and unknown pixels based on their neighborhood pixel information. (Gao et al., 2018) investigated a CNN for the fusion of multi-focus images where a deeper network was utilized for constructing an initial decision map. (Amin-Naji et al., 2019) proposed a CNN-based multi-focus image fusion approach combined with ensemble learning. (Zhang and Liu, 2020) presented a general image fusion framework based on a CNN called IFCNN. Moreover, several generative adversarial networks (GAN)-based fusion approaches have been investigated recently (Guo et al., 2019; Wang et al., 2020; Ma et al., 2020). These CNN-based and GAN-based approaches have demonstrated competitive fusion performance compared to previous fusion approaches. However, their training processes are very time consuming.

1.2. Motivation and contribution

Coupled neural P (CNP) systems were proposed by Peng and Wang (2019) as a novel variant of SNP systems, and they are abstracted as the mechanisms of coupled neurons and spiking neurons. CNP systems are a type of distributed and parallel computing model with the topology of a directed graph, where neurons act as the nodes of the graph and arcs represent the synapses. It is important to point out that CNP systems have an interesting characteristic: cooperative spiking mechanism of neurons in local neighborhood. The NSCT-based approaches are a class of popular image fusion approaches. However, the NSCT-based approaches can suffer from a major challenge: poor fusion performance. To address the challenge, the motivation of this work is to introduce CNP systems and make full use of the cooperative spiking mechanism for greatly improving the fusion performance of NSCT-based approaches. For this end, CNP systems with local topology are introduced to propose a novel CNP-based fusion rule for low-frequency NSCT coefficients of multi-focus images.

The main contributions of this work are summarized as follows.

- (i) CNP systems with local topology are introduced for the fusion of multi-focus images.
- (ii) A new multi-focus image fusion framework in the NSCT domain is developed, where two CNP systems act as the main components.
- (iii) An CNP-based fusion rule is developed, which utilizes the cooperative spiking mechanism of CNP systems.

- (iv) Quantitative and qualitative experimental results on multi-focus images indicate that the proposed approach can approach or even reach state-of-the-art CNN-based fusion approaches in terms of fusion performance. This benefits from the cooperative spiking mechanism of CNP systems.

The remainder of this paper is organized as follows. Section 2 briefly reviews CNP systems with local topology and NSCT. Section 3 details the proposed image fusion framework in the NSCT domain. Section 4 presents experimental results. Conclusions are summarized in Section 5.

2. Methodology

We will briefly review two methods utilized in this paper, namely the CNP systems with local topology and NSCT. Additional details can be found in Peng and Wang (2019) and Zhang and Guo (2009).

2.1. CNP systems with local topology

As a variant of spiking neural P (SNP) systems (Ionescu et al., 2006; Peng et al., 2017, 2019, 2020a,b,c), CNP systems are a neural-like computing model, inspired from the mechanisms of coupled neurons and spiking neurons. The formal definition of CNP systems can be found in Peng and Wang (2019). To deal with image fusion, CNP systems are extended to a matrix of neurons, namely CNP systems with local topology. For easy understanding, the extension can be described informally as follows.

Let $I = \{I_{ij}\}_{m \times n}$ be an image of size $m \times n$, and $C = \{C_{ij}\}_{m \times n}$ a set of the corresponding high-frequency NSCT coefficients. Thus, an CNP system Π with $m \times n$ neurons is designed, shown in Fig. 1.

An CNP system with local topology, Π , is described as follows.

- (1) Π consists of a matrix of $m \times n$ neurons, where each neuron is not only an input neuron but also an output neuron. The set of NSCT coefficients of image I , $C = \{C_{ij}\}_{m \times n}$ is the external input matrix of Π , and its output is an active level matrix $A = \{A_{ij}\}_{m \times n}$, where neuron σ_{ij} has an external input C_{ij} and an external output A_{ij} , shown in Fig. 1. For active level matrix A , element A_{ij} indicates the active level of neuron σ_{ij} .
- (2) A local topology is introduced, that is, each neuron σ_{ij} can communicate only with the neighboring neurons in its r -neighborhood defined by $\delta_r(\sigma_{ij}) = \{\sigma_{kl} \mid |i - k| \leq r, |j - l| \leq r\}$, shown in Fig. 1. In each computing step, neuron σ_{ij} retrieves the weighted spikes from the neurons in $\delta_r(\sigma_{ij})$, and then sends the generated spikes back to the neighboring neurons when it fires. The local topology is associated with a local weigh matrix $W_{r \times r} = \{w_{ij}\}_{r \times r}$.
- (3) All neurons are homogeneous, that is, they have the same data structure $[u, v, \tau]$ and a spiking rule $E/(a^u, a^v, a^\tau) \rightarrow a^p$, where $p \leq u(1 + v)$, $\tau \geq 0$, and $p \geq 0$, for example, neuron σ_{ij} shown in Fig. 1.

As shown in Fig. 1, neuron σ_{ij} has three units: feedback input u , linking input v and threshold τ . Based on the local topology, neuron σ_{ij} has a local connection relation with the neighboring neurons in its r -neighborhood $\delta_r(\sigma_{ij})$.

The firing condition is expressed by $E \equiv (n_{ij}(t) \geq \tau_{ij}(t)) \wedge (u_{ij}(t) \geq u) \wedge (v_{ij}(t) \geq v) \wedge (\tau_{ij}(t) \geq \tau)$, where $n_{ij}(t) = u_{ij}(t)(1 + v_{ij}(t))$ (called the nonlinear coupled modulation mechanism). If satisfying the firing condition, then neuron σ_{ij} fires, indicating spikes with the value u and spikes with the value v , and threshold τ are consumed, and spikes with the value p are generated and sent to the neighboring neurons. Therefore, according to the spiking mechanism, the state for neuron σ_{ij} can be updated by

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

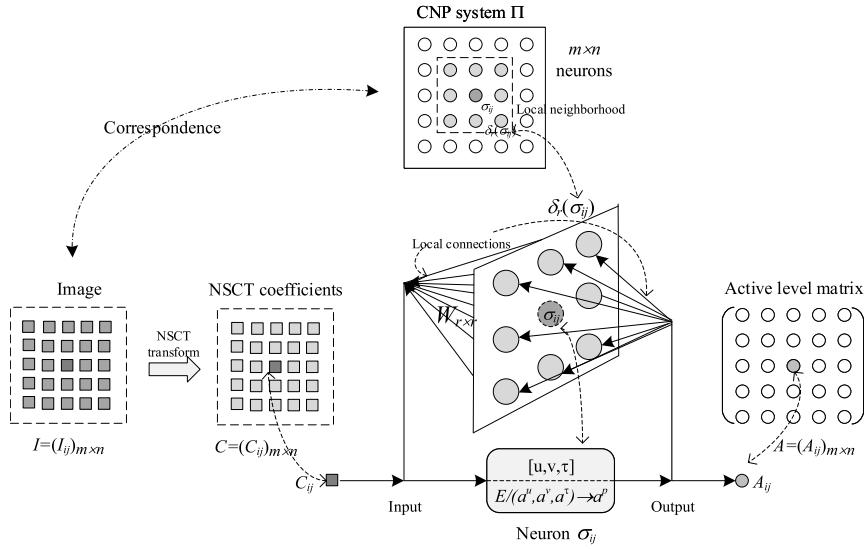


Fig. 1. An CNP system of $m \times n$ neurons, which corresponds to an image $I = \{I_{ij}\}_{m \times n}$ (top side); the bottom shows a neuron σ_{ij} and its local neighborhood structure.

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

where $u_{ij}(t+1)$ and $v_{ij}(t+1)$ are the states in neuron σ_{ij} at time $t+1$, and $u_{ij}(t)$ and $v_{ij}(t)$ are the states in neuron σ_{ij} at time t ; $p_{kl}(t)$ is the value of spikes retrieved from neighboring neuron σ_{kl} at time t ; u and v are the consumed values of spikes; w_{kl} is a local weight, and C_{ij} is an external input.

Moreover, threshold for neuron σ_{ij} can be updated by

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

where $\tau_{ij}(t+1)$ and $\tau_{ij}(t)$ are thresholds in neuron σ_{ij} at time $t+1$ and time t respectively; τ is the consumed threshold; p_{ij} denotes the value of spikes generated by neuron σ_{ij} when it fires, and w is the corresponding threshold weight.

The execution process of CNP system Π can be described as follows. For each neuron σ_{ij} , initial parameters are assigned: $u_{ij}(0) = v_{ij}(0) = 0$ and $\tau_{ij}(0) = \tau_0$, and set $W_{r \times r}$ and w . The CNP system Π begins from the initial state and operates iteratively until reaching iteration number t_{max} . The system then halts. During the computation, the number of times that each neuron σ_{ij} fires is considered as its output and utilized as a control signal for image fusion. For simplicity, a maximum consumption strategy is adopted in the application of firing rules (i.e., set $u = u_{ij}(t)$, $v = v_{ij}(t)$). Based on the working procedure, CNP system Π is implemented in Algorithm 1.

2.2. NSCT

It is established that traditional multi-resolution transform approaches can suffer from a weak capability for capturing direction, edge, and detail information. Such approaches include the LP, DWT, and contourlet transform (CT). The NSCT is a novel two-dimensional image analysis tool. It is a multi-resolution and multi-directional extension for images derived from the CT (Zhang and Guo, 2009; Yang et al., 2008). Therefore, the NSCT can overcome the problem of wavelets being unable to express higher-dimensional singularities effectively, as well as the fact that the CT does not provide shift invariance.

Unlike the CT, the NSCT has two main components, namely non-subsampled pyramid filter bands (NSPFBs) and non-subsampled directional filter bands (NSDFBs). NSPFBs can provide a subband decomposition structure similar to the LP for the CT, which ensures the

Algorithm 1: CNP systems. Call format: $A = \text{CNPS}(F)$

Require: The matrix of NSCT coefficients, $C = (C_{ij})_{m \times n}$;
Ensure: Active level matrix, $A = (A_{ij})_{m \times n}$;

- 1: Assign priori parameters: t_{max} , τ_0 , w , p , r , $W = (w_{ij})_{r \times r}$;
/ Initialization */*
- 2: $U = (u_{ij})_{m \times n} = (0)_{m \times n}$;
- 3: $V = (v_{ij})_{m \times n} = (0)_{m \times n}$;
- 4: $T = (\tau_{ij})_{m \times n} = (\tau_0)_{m \times n}$;
- 5: $A = (A_{ij})_{m \times n} = (0)_{m \times n}$;
/ Execution procedure of CNP systems */*
- 6: **for** each $t \in [1, t_{max}]$ **do**
- 7: **for** each $i \in [1, m]$ **do**
- 8: **for** each $j \in [1, n]$ **do**
- 9: Update u_{ij} in U by using Eq. (1);
- 10: Update v_{ij} in V by using Eq. (2);
- 11: Update τ_{ij} in T by using Eq. (3);
- 12: **if** $u_{ij}(1 + v_{ij}) \geq \tau_{ij}$ **then**
- 13: $A_{ij} = A_{ij} + 1$;
- 14: **end if**
- 15: **end for**
- 16: **end for**
- 17: **end for**
/ Output */*
- 18: **return** A ;

multi-scale property. NSDFBs are generated by removing the down- and up-samplers in the tree structures of directional filter bands, and up-sampling filters directly instead. The NSDFBs also provide directionality.

Fig. 2(a) shows the structure of the NSCT. Based on this structure, a source image is first decomposed by the NSPFBs. A low- and high-frequency image can be generated at each decomposition level. Subsequent decomposition can be performed to decompose the low-frequency components iteratively, which can capture the singularities in an image. Next, the high-frequency subbands at each scale are decomposed by NSDFBs with l stages, which can produce 2^l directional subbands. Based on NSDFB decomposition, the NSCT can extract more precise directional detail information. It should be noted that the subbands produced by the NSPFBs and NSDFBs have the same size as the source image. Fig. 2(b) shows a structure, consisting of a band of filters, which splits the 2D frequency plane into subbands.

In summary, compared to traditional multi-resolution transforms, such as the LP, DWT, and CT, the NSCT provides several advantages,

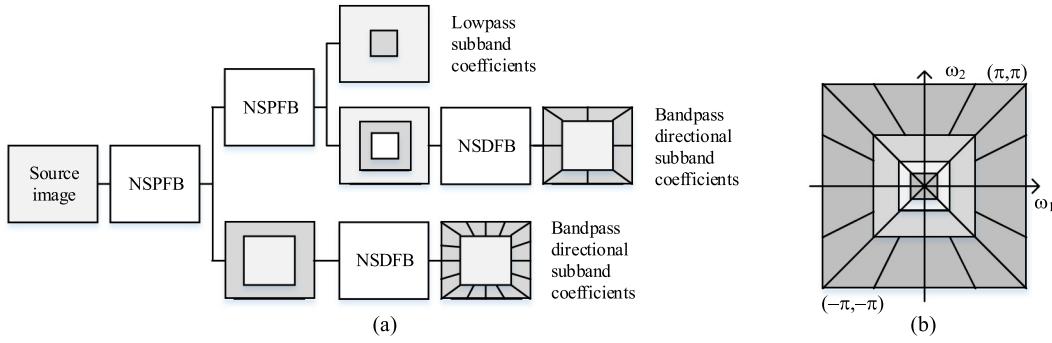


Fig. 2. Structure of the NSCT: (a) decomposition framework and (b) ideal frequency partitioning.

including shift invariance, and the ability to capture direction, edge, and detailed information. These properties are utilized in the proposed fusion approach.

3. The proposed approach

3.1. Overview of image fusion framework

We propose a multi-focus image fusion framework based on CNP systems in the NSCT domain, as shown in Fig. 3. This fusion framework includes five components: (i) the NSCT transform, (ii) INSML-based fusion rules, (iii) CNP-based fusion rules (iv) inverse transform, and (v) consistency verification. In Fig. 3, “ I_A ” and “ I_B ” are two source images, “ I_C ” is the reconstructed or initial fusion image, and “ I_F ” is the final fusion image.

The proposed approach is essentially a NSCT-based fusion approach. Therefore, like most of the existing NSCT-based fusion approaches, the proposed approach includes two basic components: NSCT transform and inverse NSCT transform, shown in Fig. 3.

As mentioned above, the existing NSCT-based fusion approaches encountered a serious challenge: poor fusion performance. The purpose of designing this fusion framework is to introduce CNP system with local topology for greatly improving the fusion performance of NSCT-based fusion approaches. The interesting mechanism can work well with low-frequency NSCT coefficients.

The proposed fusion framework differs from the existing NSCT-based fusion approaches in three core components:

- (i) CNP-based fusion rules. Two CNP systems Π_A and Π_B are designed to realize the fusion rules of low-frequency NSCT coefficients. The low-frequency NSCT coefficients of images I_A and I_B are the external inputs for CNP systems Π_A and Π_B , respectively. The two CNP systems begin from initial states and operate iteratively until reaching iteration number t_{max} . Then, they halt. The outputs of the two CNP systems, namely active level matrices, A_A and A_B , are utilized as control signals for the fusion rules of low-frequency NSCT coefficients (Eq. (4)).
- (ii) INSML-based fusion rules. An INSML feature is introduced for high-frequency NSCT coefficients. The INSML features of high-frequency NSCT coefficients of images I_A and I_B are extracted and utilized to develop the INSML-based fusion rules (Eq. (7)).
- (iii) Consistency verification. The initial fusion image I_C is further optimized utilizing Eqs. (8)–(10).

In the following sections, we detail three core components in the proposed fusion framework.

3.2. Fusion rules for low-frequency coefficients

Low-frequency coefficients express high-level information in an image because they contain most of an image's energy. The weighted

average and absolute maximum are two popular fusion approaches. However, they can reduce contrast in fusion images and lose detail information. To overcome these defects, an CNP-systems-based low-frequency rules are adopted in the proposed fusion framework.

Suppose that Π_A and Π_B are two CNP systems with local topology, which are associated with source images I_A and I_B , respectively. The low-frequency NSCT coefficients of images I_A and I_B are regarded as the external inputs for Π_A and Π_B , respectively. Beginning at initial state, the two CNP systems operate iteratively until reaching t_{max} , thus, they halt. Let A_A and A_B denote active level matrices associated with Π_A and Π_B , respectively (i.e., $A_A = (A_{ij}^A)_{m \times n}$ and $A_B = (A_{ij}^B)_{m \times n}$, where A_{ij}^A (resp. A_{ij}^B) is active level that neuron σ_{ij} in Π_A (resp. Π_B) fires. Based on the two active level matrices, fusion rules of low-frequency coefficients are given by:

$$C_{l_0}^C(i, j) = \begin{cases} C_{l_0}^A(i, j), & \text{if } A_{ij}^A > A_{ij}^B \\ C_{l_0}^B(i, j), & \text{if } A_{ij}^A < A_{ij}^B \\ (C_{l_0}^A(i, j) + C_{l_0}^B(i, j))/2, & \text{otherwise} \end{cases} \quad (4)$$

where $C_{l_0}^A(i, j)$ and $C_{l_0}^B(i, j)$ are the low-frequency NSCT coefficients of images I_A and I_B , respectively, at position (i, j) ; $C_{l_0}^C(i, j)$ are the low-frequency NSCT coefficients of fusion image I_C at position (i, j) , where $1 \leq i \leq m$ and $1 \leq j \leq n$.

3.3. Fusion rules for high-frequency coefficients

High-frequency coefficients include the edges and contours of an image. Previous studies have indicated that a modified Laplacian operator can efficiently express the significant features and sharp edges of images (Huang and Jing, 2007; Chai et al., 2010), particularly at high-frequencies (Qu et al., 2009). However, the influence of diagonal coefficients and adjacent information is not considered. Therefore, an improved sum-modified Laplacian (ISML) operator is given by

$$\begin{aligned} \text{ISML}_{lr}(i, j) = & |2C_{lr}(i, j) - C_{lr}(i-1, j) - C_{lr}(i+1, j)| \\ & + |2C_{lr}(i, j) - C_{lr}(i, j-1) - C_{lr}(i, j+1)| \\ & + \frac{1}{\sqrt{2}}|2C_{lr}(i, j) - C_{lr}(i-1, j-1) - C_{lr}(i+1, j+1)|, \\ & + \frac{1}{\sqrt{2}}|2C_{lr}(i, j) - C_{lr}(i-1, j+1) - C_{lr}(i+1, j-1)| \end{aligned} \quad (5)$$

where $C_{lr}(i, j)$ is the high-frequency NSCT coefficient in layer l and direction r at position (i, j) . Note that diagonal coefficients and adjacent information are considered in $\text{ISML}_{lr}(i, j)$ and a factor $\frac{1}{\sqrt{2}}$ is assigned for diagonal coefficients. Based on the ISML, the INSML can be computed as follows:

$$\text{INSML}_{lr}(i, j) = \sum_a \sum_b W'(a, b) \text{ISML}_{lr}(i+a, j+b), \quad (6)$$

where $\text{INSML}_{lr}(i, j)$ denotes the INSML feature value of high-frequency coefficients in layer l , direction r , and position (i, j) , and W' is defined

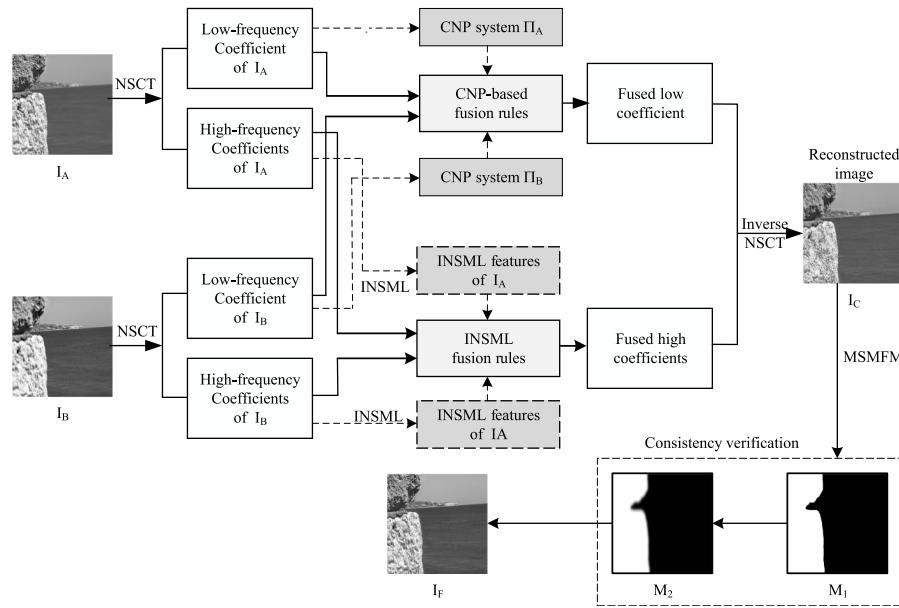


Fig. 3. The presented CNP-systems-based framework in the NSCT domain, in which I_A and I_B are two multi-focus images, I_C is the reconstructed image or initial fusion image, I_F is the final fusion image, and M_1 and M_2 are the initial and final decision matrices, respectively.

as

$$W' = \frac{1}{15} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

Suppose that $\text{INSML}_{lr}^A(i, j)$ and $\text{INSML}_{lr}^B(i, j)$ denote the INSML values of high-frequency NSCT coefficients of images I_A and I_B , respectively, in layer l , direction r , and position (i, j) . Based on the INSML features, fusion rules for high-frequency NSCT coefficients are defined by

$$C_{lr}^C(i, j) = \begin{cases} C_{lr}^A(i, j), & \text{if } \text{INSML}_{lr}^A(i, j) \geq \text{INSML}_{lr}^B(i, j) \\ C_{lr}^B(i, j), & \text{otherwise} \end{cases} \quad (7)$$

where $C_{lr}^A(i, j)$, $C_{lr}^B(i, j)$ and $C_{lr}^C(i, j)$ are the high-frequency NSCT coefficients of original images A and B , and fusion image I_C in layer l , direction r , and position (i, j) .

3.4. Consistency verification

Consistency verification is an optimization process, purpose of which is to address a weakness in the fusion of multifocus images: transform domain approaches may fail to retrieve some more important details from multi-focus images. For this reason, we adapt a consistency verification to optimize the initial fusion image.

Zhang et al. (2017) discussed a multi-scale morphological focus measure (MSMFM) for evaluating fusion images. Let $I_A(i, j)$ (resp. $I_B(i, j)$) is the pixel of source image I_A (resp. I_B) at (i, j) and $I_C(i, j)$ the pixel in the initial fusion image I_C at (i, j) . The details of computing MSMFM can be found in Zhang et al. (2017). Two difference images are defined as follows:

$$\begin{aligned} D_1 &= \{D_1(i, j)\}_{m \times n}, \quad D_1(i, j) = I_C(i, j) - I_A(i, j) \\ D_2 &= \{D_2(i, j)\}_{m \times n}, \quad D_2(i, j) = I_C(i, j) - I_B(i, j) \end{aligned} \quad (8)$$

By utilizing the MSMFM on D_1 and D_2 , a binary decision map $M_1 = \{M_1(i, j)\}$ is computed as

$$M_1(i, j) = \begin{cases} 1, & \text{if } \text{MSMFM}(D_1(i, j)) \leq \text{MSMFM}(D_2(i, j)) \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

where $\text{MSMFM}(D_1(i, j))$ and $\text{MSMFM}(D_2(i, j))$ are MSMFM values on D_1 and D_2 at position (i, j) , respectively.

Because there are some artifacts at the boundaries between focused and unfocused areas, we utilize a filter to improve the initial decision

map M_1 . Specifically, we utilize an effective edge-preserving filter (Liu et al., 2017) 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 as 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 $I_F = \{I_F(i, j)\}_{m \times n}$ can be constructed as follows:

$$I_F(i, j) = M_2(i, j)I_A(i, j) + (1 - M_2(i, j))I_B(i, j), \quad (10)$$

where $I_A(i, j)$, $I_B(i, j)$ and $I_F(i, j)$ is the pixels in original images A and B , and fusion image I_F at position (i, j) ; $M_2(i, j)$ is the element of M_2 at position (i, j) . In this work, assign $r' = 5$ and $\epsilon = 0.1$.

3.5. The implementation of proposed fusion approach

The proposed approach is an CNP-systems-based fusion approach in NSCT domain for multi-focus images. According to image fusion framework in Fig. 3, the proposed fusion approach is implemented in Algorithm 2.

Algorithm 2 has two inputs, multi-focus images I_A and I_B , and an output, the fusion image I_F . In Algorithm 2, NSCT(·) is a NSCT transform, and INSCT(·) is the corresponding inverse NSCT transform. C_0^A (resp. C_0^B) is the matrix of low-frequency NSCT coefficients of image I_A (resp. I_B), and C_l^A (resp. C_l^B) is the matrix of high-frequency NSCT coefficients of image I_A (resp. I_B), $1 \leq l \leq N$, where N is the number of high-frequency decomposition planes. Note that lines 3–15 in Algorithm 2 correspond to fusion rule (4), while lines 16–28 are associated with fusion rule (7).

4. Experimental results and discussion

4.1. Data set

In the experiments, an open data set containing of 19 pairs of multi-focus images was utilized to evaluate the proposed and compared approaches. Fig. 4 shows the test images. For each pair of images, the left side is near-focused image, while the right side is far-focused. Table 1 lists the sizes of these images.



Fig. 4. Image data set utilized in our experiments (19 pairs of multi-focus images).

Table 1
The sizes of images in data set.

Images	Seascape	Temple	Balloon	Clock	Toy
Size	577 × 314	481 × 516	640 × 480	512 × 512	512 × 512
Images	Calendar	Wine	Newspaper	Leopard	Leaf
Size	267 × 175	256 × 256	322 × 234	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	Grass	
Size	636 × 476	640 × 480	640 × 480	261 × 177	

4.2. Compared approaches

In the experiments, the proposed fusion approach is evaluated on the data set, and compared to eight previous approaches and three state-of-the-art CNN-based approaches.

The eight previous fusion approaches are the wavelet (Zeeuw, 1998), CVT (Nencini et al., 2007), MSVD (Naidu, 2011), LP-SR (Liu et al., 2015a), GFF (Li et al., 2013), DSIFT (Liu et al., 2015b), BF (Zhang et al., 2017), and NSCT-PCNN (Qu et al., 2008).

Three state-of-the-art CNN-based fusion approaches are CNN (Liu et al., 2017), ECNN (Amin-Naji et al., 2019), and IFCNN (Zhang and Liu, 2020).

4.3. Parameter set

The proposed fusion approach is an CNP-systems-based multi-focus image fusion approach in the NSCT domain, and it was implemented in Matlab 2017b and on an Intel Core i7-6700 CPU running at 3.4 GHz with 16 GB of RAM. In our experiments, $m \times n$ was determined automatically based on the size of each test source image and we set $r = 3$,

$\tau_0 = 1$, $p = 20$, $t_{max} = 200$, $r' = 5$, $\epsilon = 0.1$, and local weight matrix is

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

The parameters above are a group of the optimal parameters selected from many experiments. In the NSCT transform, the directional and pyramid filter utilize parameters of "pkva", "9/7", and "[1 1 1]", with a layer number of four.

The eight previous fusion approaches were realized on the same platform according to open-source code, and utilized the same parameters as the original studies.

The three CNN-based fusion approaches were implemented on the Ubuntu 16.04 OS with an Intel Xeno Silver 4110 CPU running at 2.10 GHz and an NVIDIA Tesla P100 with 16 GB of RAM based on the open source code provided by the original authors.

4.4. Evaluation metrics

To evaluate the proposed and compared approaches, we present subjective and objective comparison results. The subjective comparisons are visual comparison of fusion images. Objective evaluations are

Algorithm 2: The presented fusion algorithm for multi-focus images.**Require:**Multi-focus images, I_A and I_B ;**Ensure:**

```

The fused image  $I_F$ ;
/* NSCT decomposition of image  $I_A$  */
1:  $[C_0^A, C_1^A, \dots, C_N^A] = \text{NSCT}(I_A)$ ;
/* NSCT decomposition of image  $I_B$  */
2:  $[C_0^B, C_1^B, \dots, C_N^B] = \text{NSCT}(I_B)$ ;
/* Fusion of low-frequency coefficients */
3:  $A_A = \text{CNPS}(C_0^A)$ ;
4:  $A_B = \text{CNPS}(C_0^B)$ ;
5: for each  $i \in [1, m]$  do
6:   for each  $j \in [1, n]$  do
7:     if  $A_A(i, j) > A_B(i, j)$  then
8:        $C_0^C(i, j) = C_0^A(i, j)$ ;
9:     else if  $A_A(i, j) < A_B(i, j)$  then
10:       $C_0^C(i, j) = C_0^B(i, j)$ ;
11:    else
12:       $C_0^C(i, j) = (C_0^A(i, j) + C_0^B(i, j))/2$ ;
13:    end if
14:   end for
15: end for
/* Fusion of high-frequency coefficients */
16: for each  $l \in [1, N]$  do
17:   for each  $i \in [1, m]$  do
18:     for each  $j \in [1, n]$  do
19:       Compute  $\text{INSML}_l^A(i, j)$  according to Eq. (6);
20:       Compute  $\text{INSML}_l^B(i, j)$  according to Eq. (6);
21:       if  $\text{INSML}_l^A(i, j) \geq \text{INSML}_l^B(i, j)$  then
22:          $C_l^C(i, j) = C_l^A(i, j)$ ;
23:       else
24:          $C_l^C(i, j) = C_l^B(i, j)$ ;
25:       end if
26:     end for
27:   end for
28: end for
/* Inverse NSST transform */
29:  $I_C = \text{INSST}(C_0^C, C_1^C, \dots, C_N^C)$ ;
/* Consistency verification of initial fusion image  $I_C$  */
30: Compute final fusion image  $I_F$  according to Eq. (10);
31: return  $I_F$ ;

```

also performed on fusion images. The objective evaluation indices utilized in our experiments are total information transferred from source images to the fusion images (Q_G) (Xydeas and Petrovic, 2009), feature mutual information (Q_{FMI}) (Haghigheh et al., 2011), phase congruency (Q_P) (Zhao et al., 2006), structural similarity (Q_Y) (Wang et al., 2004), and human perception (Q_{CB}) (Chen and Blum, 2009). Additionally, we present comparisons of the proposed fusion approach and compared approaches for each test image in terms of execution time for evaluating computational efficiency.

4.5. Subjective evaluation

Subjective evaluation was conducted by comparing the visual quality of the images fused by the proposed and compared approaches. Figs. 5–8 present the images fused by the nine approaches on four typical multi-focus images, namely "Balloon", "Leaf", "Girl" and "Grass", respectively.

Fig. 5 compares the fusion images generated by the proposed approach and eight other approaches for multi-focus image "Balloon". Figs. 5 (a) and (b) present two source images with different focus settings. In Fig. 5, one can see that the image fused by the wavelet approach has some missing details. Additionally, blurring and noise exist in the image fused by the MSVD approach.

Fig. 6 presents the fusion images generated by the proposed approach and eight other approaches for multi-focus image "Leaf". Figs. 6

(a) and (b) present source images with different focus settings. To identify the visual differences in the fusion images, we utilize rectangular boxes to mark local observation areas, and present magnified views in the bottom-right corners of the figures. One can see that the fusion images in Figs. 6 (e) and (i) have significant blurring, and the fusion images in Figs. 6 (c), (d), and (j) have low brightness. This indicates that the MSVD, BF, wavelet, CVT, and NSCT-PCNN approaches do not provide the desired fusion effect.

Fig. 7 presents the fusion images generated by the proposed approach and eight other approaches for multi-focus image "Girl". Figs. 7 (a) and (b) present two source images with different focus settings. In the large boxes, one can see that there are some blurry artifacts in the fusion images in Figs. 7 (c), (d), (e), (f), and (j). This indicates that the wavelet, LP-SR, MSVD, CVT, and NSCT-PCNN approaches fail to distinguish boundaries in image "Girl".

Fig. 8 presents the fusion images generated by the proposed approach and eight other approaches for multi-focus image "Grass". Figs. 8 (a) and (b) present two source images with different focus settings. In the large boxes, one can see that there are some blurry artifacts in the fusion images in Figs. 8 (c) and (e), indicating that the wavelet and MSVD approaches do not provide the desired fusion effect.

Overall, we easily judge that the proposed approach outperforms the wavelet, CV, MSVD, and LP-SR approaches in terms of fusion quality. However, it is difficult to rank the fusion qualities of the images generated by the proposed approach and GFF, DSIFT, BF, and NSCT-PCNN approaches based on subjective visual evaluations.

Additionally, we present images fused by the proposed approach for the other fifteen images in the data set in Fig. 9. From the perspective of visual quality, the proposed approach provides a good fusion effect.

Several CNN-based multi-focus image fusion approaches have been developed, including the basic CNN (Liu et al., 2017), ECNN (Amin-Naji et al., 2019) and IFCNN (Zhang and Liu, 2020). Previous studies have shown that CNN-based fusion approaches provide excellent fusion quality and performance for multi-focus image fusion. Fig. 10 presents fusion images generated by the proposed approach and three CNN-based approaches for six multi-focus images, namely "Seascape", "Temple", "Toy", "Calendar", "Wine" and "Corner". All four approaches provide good visual quality for the fusion of all six multi-focus images. However, it is difficult to distinguish which approaches performed the best or worst visually. Therefore, an objective assessment was performed.

4.6. Quantitative evaluation

To compare the proposed approach and other approaches, five fusion quality metrics were utilized to evaluate fusion performance, namely Q_G , Q_{FMI} , Q_P , Q_Y , and Q_{CB} . Experiments were conducted for all 19 multi-focus images in the data set. The experimental results are provided in Table 2–Table 4. For each metric, the greater its value, the better the performance of the corresponding approach. In each table, bold font for each metric indicates that the corresponding approach provides the best index value.

Table 2 lists comparison results for the proposed approach and eight other approaches on four multi-focus source images ("Balloon", "Leaf", "Girl" and "Grass") in terms of the five metrics mentioned above.

For the "Balloon" image, the proposed approach achieves the best values of Q_{FMI} , Q_P , Q_Y and Q_{CB} and obtains the next-best value of Q_G .

For the "Leaf" image, the proposed approach obtains the best values for Q_G , Q_{FMI} and Q_Y and obtains the next-best value of Q_P , and it still has a large metric value of Q_{CB} .

For the "Girl" image, the proposed approach obtains the best values for Q_G , Q_{FMI} and Q_P , and it obtains the next-best values of Q_Y and Q_{CB} .

For the "Grass" image, the proposed approach achieves the best values of all five metrics.

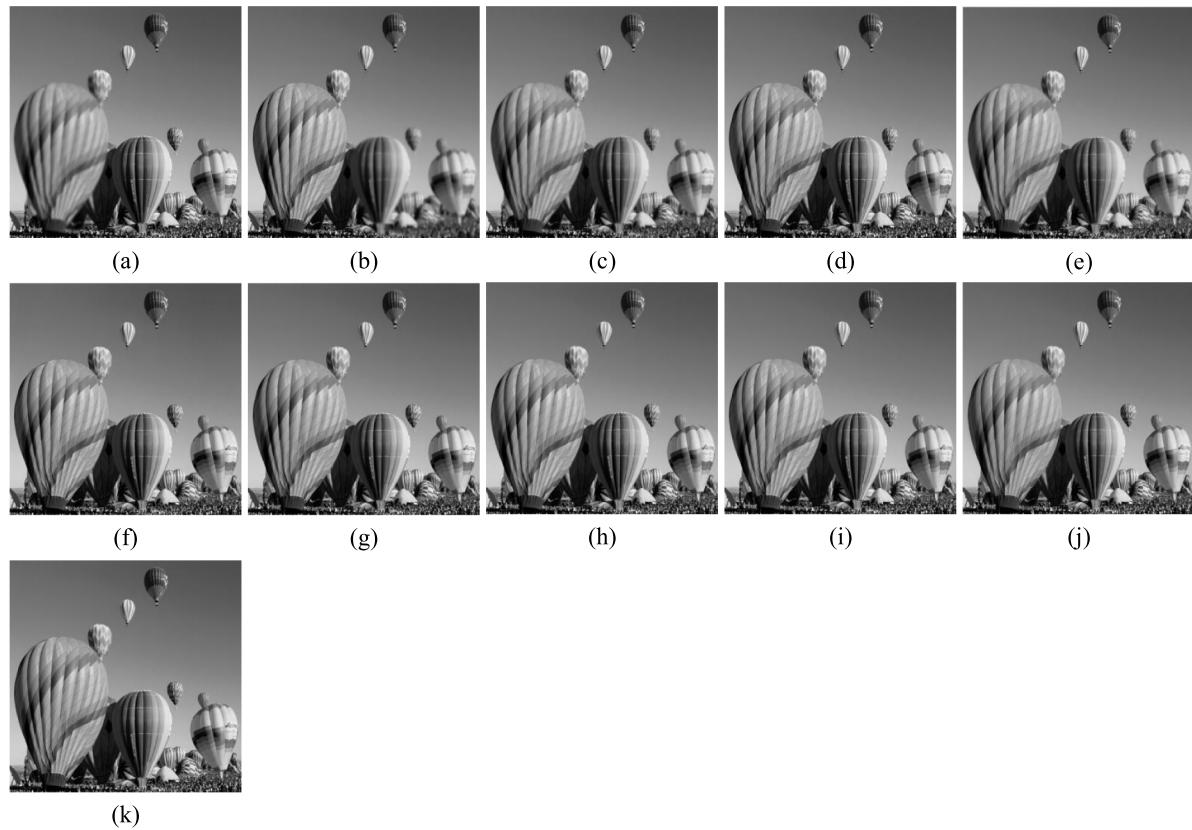


Fig. 5. Source images and fusion results of all approaches on the “Balloon” image: (a) and (b) are two multi-focus source images and (c)–(k) are the results fused by wavelet, CTV, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN, and presented approaches, respectively.

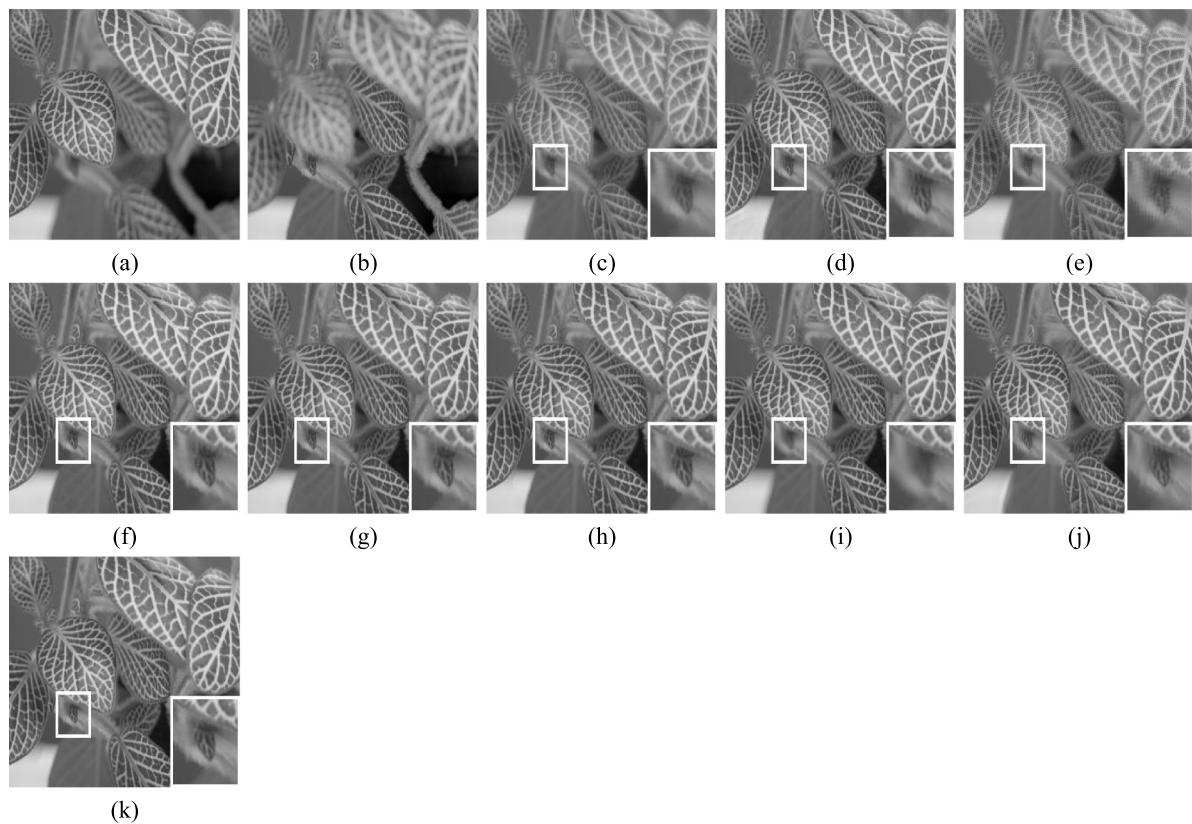


Fig. 6. Source images and fusion results of all approaches on the “Leaf” image: (a) and (b) are two multi-focus source images and (c)–(k) are the results fused by the wavelet, CTV, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN, and presented approaches, respectively.

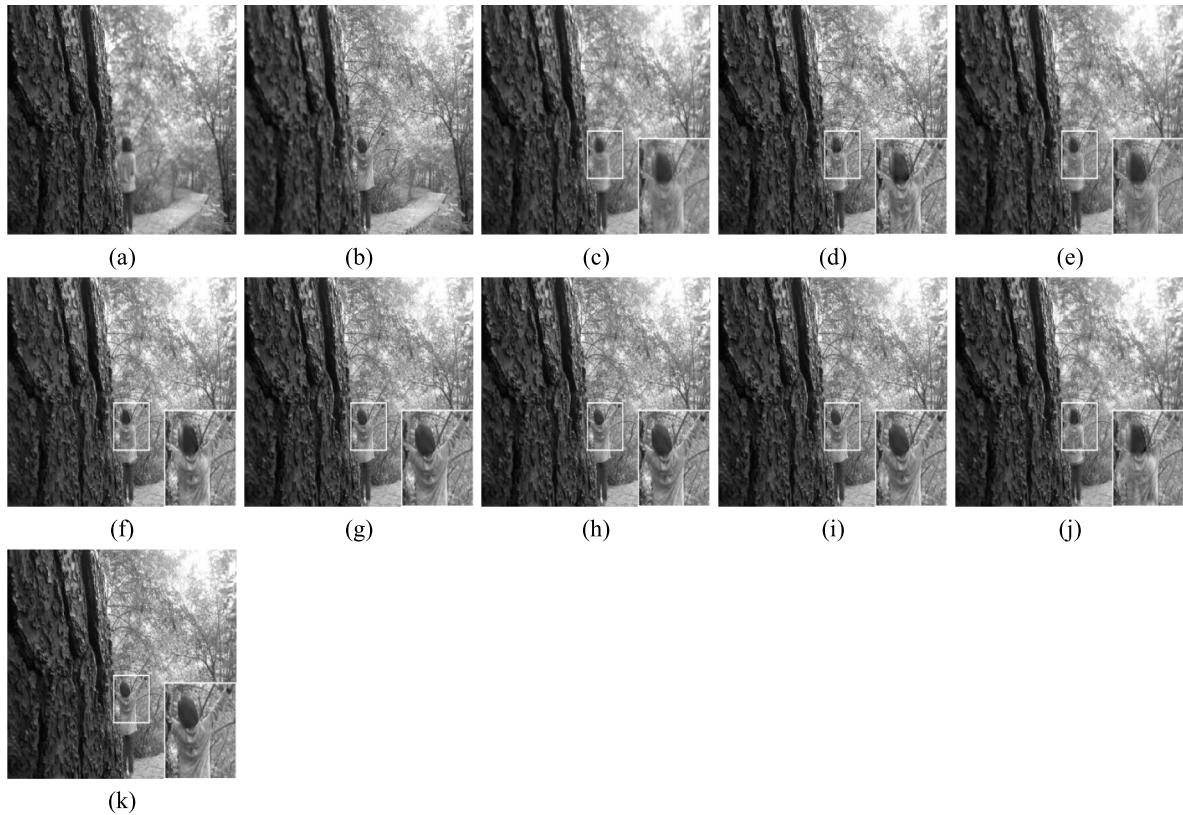


Fig. 7. Source images and fusion results of all approaches on the “Girl” image: (a) and (b) are two multi-focus source images and (c)–(k) are the results fused by wavelet, CVT, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN, and presented approaches, respectively.

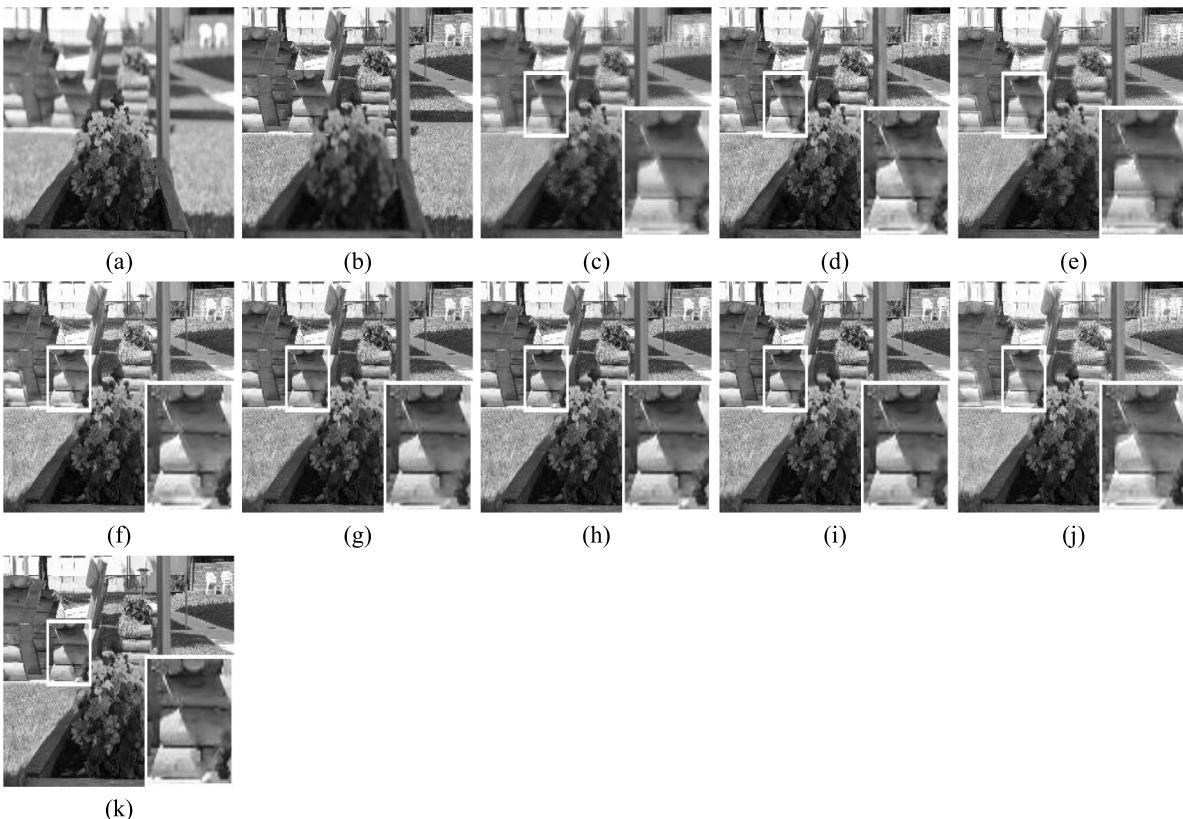


Fig. 8. Source images and fusion results of all approaches on the “Grass” image: (a) and (b) are two multi-focus source images and (c)–(k) are the results fused by wavelet, CVT, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN, and presented approaches, respectively.

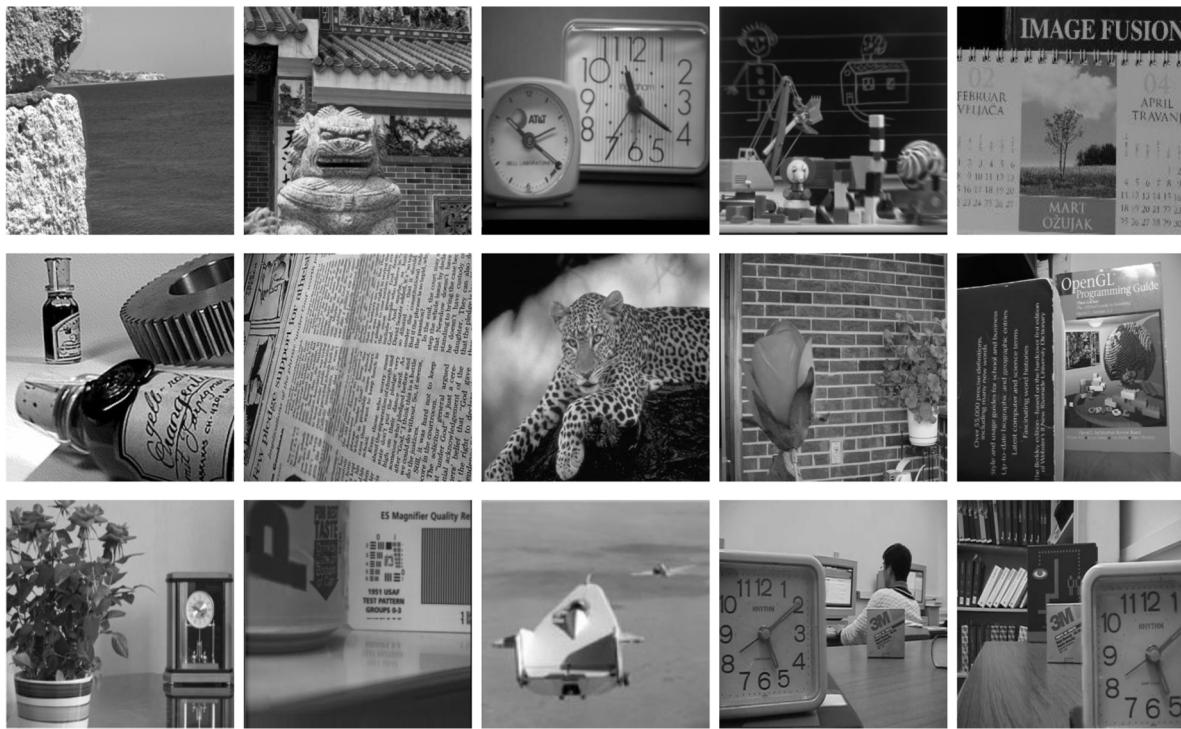


Fig. 9. Fusion results of the presented approach on the other fifteen multi-focus images.

Table 2

The comparison results of the proposed approach with previous approaches on four multi-focus source images in terms of five metrics.

Images	Methods	Q_G	Q_{FMI}	Q_P	Q_Y	Q_{CB}
Balloon	Wavelet	0.7676	0.9209	0.8874	0.9392	0.8348
	CVT	0.8099	0.9223	0.9419	0.9837	0.8492
	MSVD	0.7672	0.9150	0.8394	0.9549	0.8385
	LP-SR	0.8120	0.9229	0.9351	0.9808	0.7208
	GFF	0.8195	0.9233	0.9488	0.9937	0.9056
	DSIFT	0.8194	0.9233	0.9488	0.9938	0.9059
	BF	0.8194	0.9233	0.9486	0.9937	0.9057
	NSCT-PCNN	0.8038	0.9230	0.9227	0.9774	0.8701
	Proposed	0.8194	0.9233	0.9488	0.9938	0.9060
Leaf	Wavelet	0.5093	0.7561	0.6769	0.8347	0.6195
	CVT	0.6760	0.7660	0.7590	0.9162	0.6961
	MSVD	0.4384	0.7438	0.5320	0.7645	0.6023
	LP-SR	0.7134	0.7647	0.8171	0.9566	0.7569
	GFF	0.7192	0.7656	0.8388	0.9779	0.7854
	DSIFT	0.7189	0.7648	0.8349	0.9829	0.7941
	BF	0.7177	0.7655	0.8506	0.9857	0.7927
	NSCT-PCNN	0.6791	0.7654	0.7627	0.9342	0.7110
	Proposed	0.7197	0.7664	0.8435	0.9865	0.7912
Girl	Wavelet	0.5102	0.8086	0.6812	0.8532	0.6888
	CVT	0.6265	0.8161	0.7176	0.9544	0.7004
	MSVD	0.5468	0.8048	0.6147	0.8959	0.6869
	LP-SR	0.6574	0.8179	0.7542	0.9665	0.7223
	GFF	0.6802	0.8215	0.8409	0.9893	0.7798
	DSIFT	0.6819	0.8216	0.8435	0.9926	0.7934
	BF	0.6813	0.8217	0.8440	0.9926	0.7931
	NSCT-PCNN	0.5780	0.8120	0.6461	0.9379	0.7253
	Proposed	0.6819	0.8218	0.8446	0.9924	0.7932
Grass	Wavelet	0.4748	0.8278	0.5657	0.8170	0.6313
	CVT	0.5575	0.8363	0.5881	0.9220	0.6372
	MSVD	0.5239	0.8337	0.5252	0.8773	0.6357
	LP-SR	0.6182	0.8477	0.6416	0.9517	0.7025
	GFF	0.6549	0.8554	0.7161	0.9878	0.7597
	DSIFT	0.6551	0.8545	0.7140	0.9872	0.7623
	BF	0.6531	0.8546	0.7187	0.9888	0.7624
	NSCT-PCNN	0.5047	0.8212	0.4982	0.8920	0.6476
	Proposed	0.6557	0.8557	0.7223	0.9892	0.7646

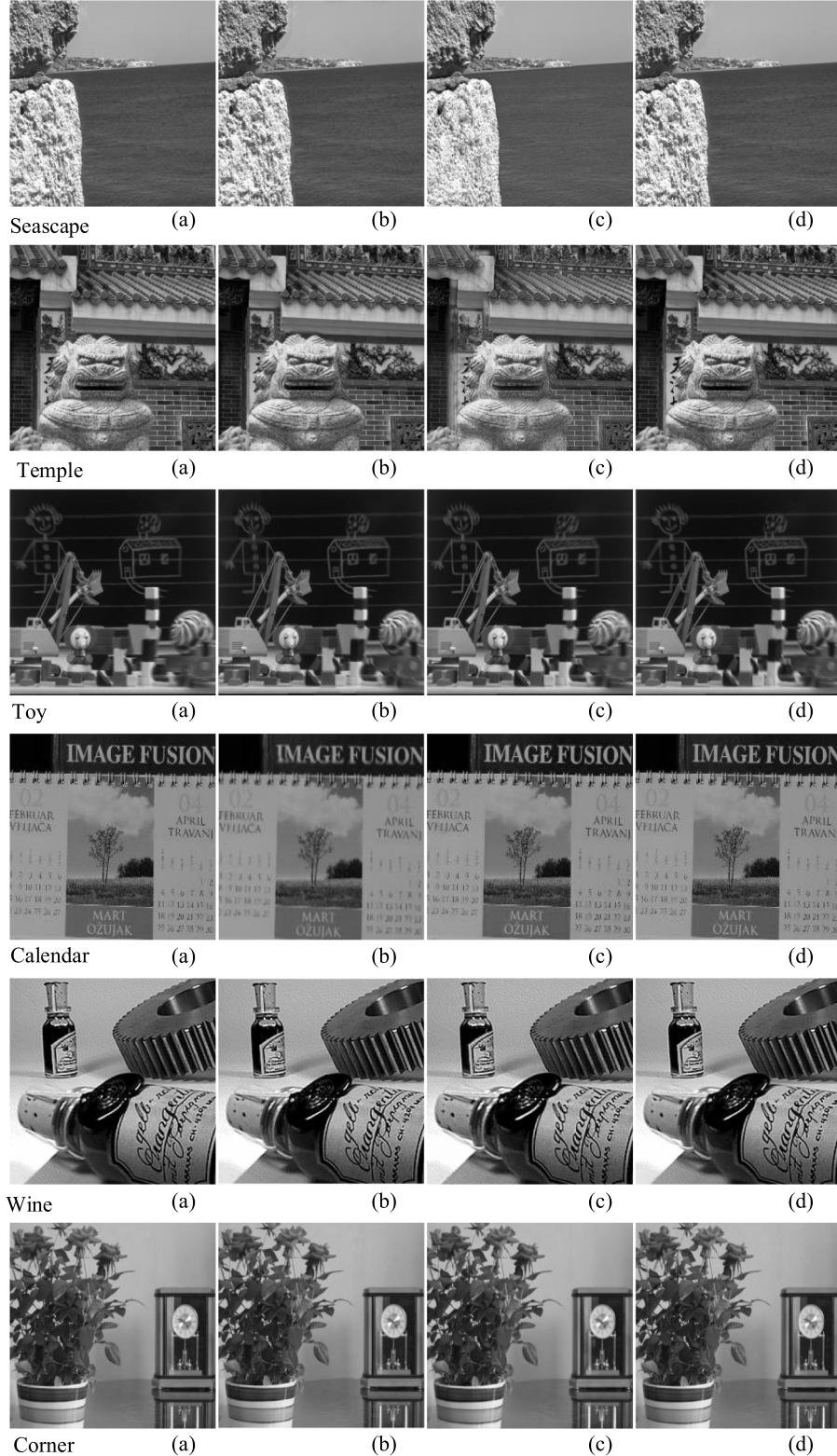


Fig. 10. Images generated by the presented approach and (d) three CNN-based approaches ((a) CNN, (b) ECNN, and (c) IFCNN-max (c)).

Table 3 lists comparison results for the proposed approach and three CNN-based approaches on six multi-focus images ("Seascape", "Temple", "Toy", "Calendar", "Wine" and "Corner") in terms of the five metrics. In **Table 3**, one can see that the proposed approach and three CNN-based approaches achieve similar values for each multi-focus image for each metric. However, one can observe the following results:

- (i) for the "Seascape" image, the proposed approach obtains the best values for Q_G , Q_{FMI} , Q_P and Q_{CB} , and it obtains the next-best value of Q_Y ;
- (ii) for the "Temple" image, the proposed approach obtains the best values for all five metrics;

Table 3

The comparison results of the proposed approach with three CNN-based approaches on six multi-focus source images in terms of five metrics.

Images	Methods	Q_G	Q_{FMI}	Q_P	Q_Y	Q_{CB}
Seacape	CNN	0.6886	0.8568	0.6390	0.9934	0.6143
	ECNN	0.6886	0.8567	0.6394	0.9886	0.6134
	IFCNN	0.5390	0.8102	0.6291	0.9767	0.5901
	Proposed	0.6888	0.8569	0.6396	0.9915	0.6159
Temple	CNN	0.7556	0.8238	0.7920	0.9928	0.8159
	ECNN	0.7448	0.8226	0.7761	0.9928	0.8080
	IFCNN	0.6768	0.8194	0.6861	0.9155	0.6069
	Proposed	0.7571	0.8243	0.7934	0.9946	0.8198
Toy	CNN	0.6966	0.9161	0.8315	0.9821	0.8041
	ECNN	0.6803	0.9147	0.7989	0.9764	0.7843
	IFCNN	0.6675	0.9048	0.7708	0.8644	0.6795
	Proposed	0.6870	0.9151	0.8366	0.9834	0.8108
Calendar	CNN	0.7033	0.8848	0.7528	0.9888	0.8057
	ECNN	0.6866	0.8864	0.7459	0.9890	0.8056
	IFCNN	0.6659	0.8791	0.7419	0.8976	0.7135
	Proposed	0.7035	0.8856	0.7542	0.9899	0.8081
Wine	CNN	0.6966	0.8473	0.8912	0.9858	0.7918
	ECNN	0.6960	0.8480	0.8879	0.9860	0.7954
	IFCNN	0.6422	0.8355	0.8295	0.9576	0.7085
	Proposed	0.6986	0.8474	0.8914	0.9865	0.8050
Corner	CNN	0.7068	0.8751	0.9239	0.9728	0.7597
	ECNN	0.7079	0.8751	0.9243	0.9786	0.7711
	IFCNN	0.6733	0.8675	0.8968	0.9225	0.6939
	Proposed	0.7076	0.8751	0.9248	0.9799	0.7743

Table 4

The comparison of average results of the proposed approach and other approaches on data set in terms of five metrics.

Methods	Q_G	Q_{FMI}	Q_P	Q_Y	Q_{CB}
Wavelet	0.5651	0.8570	0.6870	0.8485	0.6447
CVT	0.6746	0.8667	0.7396	0.9265	0.7014
MSVD	0.5419	0.8487	0.5873	0.8488	0.6384
LP-SR	0.6985	0.8691	0.7652	0.9366	0.7202
GFF	0.7177	0.8715	0.8078	0.9676	0.7648
DSIFT	0.7195	0.8720	0.8144	0.9759	0.7868
BF	0.7188	0.8721	0.8193	0.9822	0.7943
NSCT-PCNN	0.6643	0.8669	0.7114	0.9239	0.7150
CNN	0.7195	0.8722	0.8196	0.9823	0.7887
ECNN	0.7154	0.8720	0.8062	0.9817	0.7850
IFCNN	0.6627	0.8609	0.7308	0.9162	0.6742
Proposed	0.7198	0.8723	0.8203	0.9847	0.7901

Table 5

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

	Wavelet	CVT	MSVD	LP-SR	GFF
Times/s	0.2223	0.9209	0.2025	0.0319	0.2439
	DSIFT	BF	NSCT-PCNN	Proposed	
Times/s	5.0435	1.1748	142.7365	8.5459	

- (iii) for the “Toy” image, the proposed approach obtains the best values for Q_P , Q_Y and Q_{CB} , and it has the next-best value of Q_G and Q_{FMI} ;
- (iv) for the “Calendar” image, the proposed approach obtains the best values for Q_G , Q_P , Q_Y and Q_{CB} , and it obtains the next-best value of Q_{FMI} ;
- (v) for the “Wine” image, the proposed approach obtains the best values for Q_G , Q_P , Q_Y and Q_{CB} , and it has the next-best value of Q_{FMI} ;
- (vi) For the “Corner” image, the proposed approach obtains the best values for Q_{FMI} , Q_P , Q_Y and Q_{CB} , and it obtains the next-best value of Q_G .

Table 4 lists the results for the average fusion performance of the proposed approach and all eleven compared approaches in terms of the five metrics. The result for a given approach and metric represents the

average value for all 19 multi-focus images in the data set. In **Table 4**, the presented approach obtains the best values of Q_G , Q_P and Q_Y , and the presented approach obtains the next-best value of Q_{CB} .

To evaluate the computational efficiency, we computed the average execution times of the proposed and compared approaches for all 19 multi-focus images in the data set. Because CNN-based approaches require a time-consuming training process, the CNN, ECNN, and IFCNN approaches were not included in this comparison. **Table 5** lists comparison results for the proposed approach with eight other approaches in terms of average execution time. In **Table 5**, one can observe the following results: (i) Wavelet, CVT, MSVD, LP-SR, and GFF have low execution time, (ii) NSCT-PCNN has high execution time, (iii) the proposed approach and DSIFT have relatively high execution time. The comparative results indicate that the proposed approach has a high computational cost because it utilizes the NSCT transform and two CNP systems. However, the proposed approach still has high computational efficiency compared to CNN-based approaches.

In summary, subjective and objective comparison results demonstrate that the proposed method achieves good quality and fusion performance for the fusion of multi-focus images, with suitable computational efficiency.

4.7. Discussion

The proposed approach belongs to a class of MST-based fusion approach that use NSCT as a MST tool. Usually, the MST-based fusion approaches include MST transform and inverse transform, for example, NSCT and inverse NSCT, and data fusion is performed in the corresponding frequency domain. However, innovation work of this paper is mainly reflected in introducing CNP systems to develop a novel NSCT-based fusion approach for multi-focus images. The innovation is based on the following reasons:

- (1) The proposed fusion framework uses an important component: CNP systems. CNP systems are a theoretical model we developed previously and tried to propose a fusion approach for multi-focus images. CNP systems have two characteristics: spike propagation and cooperative firing in a local region, which are used to design a CNP-based fusion rule for low-frequency coefficients. Since low-frequency coefficients contain most of energy of image, the

- low-frequency coefficients can effectively stimulate the firing of neurons. This is a significant difference from the previous MST-based fusion approaches.
- (2) It is well-known that, previous MST-based fusion approaches almost have a poor fusion performance, including NSCT-based fusion approach. However, surprisingly, the proposed fusion approach (a NSCT-based fusion approach improved by CNP) can reach or approach the fusion performance of state-of-the-art CNN-based fusion approaches. This powerful improvement in fusion performance shows the capabilities of CNP model for the fusion of multi-focus images.

CNP systems are a computing model inspired by the mechanism of spiking neurons. CNP and PCNN have a similar point: spike (pulse) propagation. However, they have the following differences:

- (1) CNP and PCNN have the different state equations since CNP is inspired by the mechanism of spiking neurons.
- (2) CNP has a small number of priori parameters, while PCNN has a lot of priori parameters. Generally, to obtain a group of optimal priori parameters of PCNN, some optimization techniques (such as PSO and DE) are often utilized.

In our comparison experiments, NSCT-PCNN is a PCNN-based fusion approach in NSCT domain. The comparison results in Tables 2 and 4 indicate that CNP is superior to PCNN in terms of fusion performance.

Multi-focus images and other multimodal images (such as medical images) have the different imaging principles: (i) multi-focus images are retrieved by the same imaging device under the different focal lengths; (ii) multimodal medical images are retrieved by the different imaging devices, for example, CT vs. MRI. In the existing references of multi-focus images, there are many multi-focus image fusion approaches that often used a post-processing process (an optimization process) for further improving fusion quality. Similarly, we also adopted an existing method as a post-processing method, MSMFM. It is worth noting that, DL (deep learning)-based fusion approaches usually do not use this post-processing for the fusion of multi-focus images, for example, CNN-based and GAN-based approaches. One potential reason is that CNN and GAN can automatically extract good features of the image. DL-based fusion approaches are essentially a supervised fusion method, which needs a training procedure and depends on a huge collection of training images. The proposed fusion approach is an unsupervised approach, which does not require a training process and a huge set of training images. From the principle, CNP is not to extract the features of the image, but to use its own unique mechanism to implement a fusion rule. Compared to DL-based approaches, this is a cons of the proposed fusion approach because of the use of post-processing method.

5. Conclusions

Recently, CNP systems, as novel distributed and parallel computing models, have been developed based on the mechanism of coupled neurons and spiking neurons. CNP systems have been proven to be Turing universal number generating/accepting and functional computing devices.

This paper investigated how to apply CNP systems to deal with multi-focus fusion problems, and proposed a novel CNP-systems-based multi-focus fusion approach in the NSCT domain. The cooperative spiking mechanism of CNP systems was developed for the fusion of multi-focus images. Based on two CNP systems with local topology, a multi-focus fusion framework in the NSCT domain was designed. The role of the CNP systems is to control the fusion rules of low-frequency coefficients in the NSCT domain, where the two CNP systems are associated with the low-frequency NSCT coefficients of two multi-focus source images. Additionally, to overcome common shortcomings of transform domain approaches, a consistent verification approach is

utilized to improve the quality of initial fusion images to obtain a final fusion image. The proposed fusion approach was compared to eight previous approaches and three state-of-the-art CNN-based fusion approaches on 19 multi-focus images in terms of five fusion quality indices. Subjective and objective comparison results demonstrated a significant advantage of the proposed fusion approach that CNP can greatly improve the fusion performance of NSCT-based fusion approaches for multi-focus image fusion.

Compared to CNN-based and GAN-based approaches, however, the proposed fusion approach has a cons: it uses a post-processing method, MSMFM, for further improving the fusion performance. In future work, we will further improve the proposed CNP-based approach to overcome the cons. In addition, the proposed approach could be extended to other image fusion applications, such as multi-exposure fusion, visual and infrared image fusion, and color images. Our future work will focus on the fusion of more than two images, images with moving objects, and color images.

CRediT authorship contribution statement

Hong Peng: Conceptualization, Software and experiments, Writing. **Bo Li:** Conceptualization, Software and experiments, Writing. **Qian Yang:** Software and experiments. **Jun Wang:** Conceptualization and Writing.

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.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China (No. 62076206), China.

References

- Amin-Naji, M., Aghagolzadeh, A., Ezoji, M., 2019. Ensemble of CNN for multi-focus image fusion. *Inf. Fusion* 51, 201–214.
- Burt, P., Adelson, E., 1983. The laplacian pyramid as a compact image code. *IEEE Trans. Comput.* 31 (4), 532–540.
- Chai, Y., Li, H., Qu, J., 2010. Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain. *Opt. Commun.* 283 (19), 3591–3602.
- Chen, Y., Blum, R., 2009. A new automated quality assessment algorithm for image fusion. *Image Vis. Comput.* 27 (10), 1421–1432.
- De, I., Chanda, B., 2013. Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure. *Inf. Fusion* 14 (2), 136–146.
- Gao, X., Nie, R., Cao, J., Zhou, D., Qian, W., 2018. Fully convolutional network-based multi-focus image fusion. *Neural Comput.* 30 (7), 1775–1800.
- Guo, X., Nie, R., Cao, J., Zhou, D., Mei, L., He, K., 2019. FuseGAN: learning to fuse multi-focus image via conditional generative adversarial network. *IEEE Trans. Multimed.* 21 (8), 1982–1996.
- Haghhighat, M., Aghagolzadeh, A., Seyedarabi, H., 2011. A non-reference image fusion metric based on mutual information of image features. *Comput. Electr. Eng.* 37 (5), 744–756.
- He, D., Meng, Y., Wang, C., 2011. Contrast pyramid based image fusion scheme for infrared image and visible image. In: Proceedings of 2011 IEEE International Conference on Geoscience & Remote Sensing Symposium, pp. 597–600.
- Huang, W., Jing, Z., 2007. Multi-focus image fusion using pulse coupled neural network. *Pattern Recognit. Lett.* 28, 1123–1132.
- Ionescu, M., Păun, G., Yokomori, T., 2006. Spiking neural P systems. *Fund. Inform.* 71, 279–308.
- Jin, X., Chen, G., Hou, J., Jiang, Q., Zhou, D., Yao, S., 2018. Multimodal sensor medical image fusion based on nonsubsampled shearlet transform and S-PCNNs in HSV space. *Signal Process.* 153, 379–395.
- Lewis, J., O'Callaghan, R., Nikolov, S., Bull, D., Canagarajah, N., 2007. Pixel- and region-based image fusion with complex wavelets. *Inf. Fusion* 8 (2), 119–130.
- Li, M., Cai, W., Tan, Z., 2006. A region-based multi-sensor image fusion scheme using pulse-coupled neural network. *Pattern Recognit. Lett.* 27 (16), 1948–1956.
- Li, S., Kang, X., Fang, L., Hu, J., Yin, H., 2017. Pixel-level image fusion: a survey of the state of the art. *Inf. Fusion* 33, 100–112.

- Li, S., Kang, X., Hu, J., 2013. Image fusion with guided filtering. *IEEE Trans. Image Process.* 22 (7), 2864–2875.
- Li, B., Peng, H., Wang, J., Huang, X., 2020. Multi-focus image fusion based on dynamic threshold neural P systems and surfacelet transform. *Knowl.-Based Syst.* 196, 105794, 1–12.
- Li, S., Yang, B., 2008. Multifocus image fusion using region segmentation and spatial frequency. *Image Vis. Comput.* 26 (7), 971–979.
- Liu, Y., Chen, X., Peng, H., Wang, Z., 2017. Multi-focus image fusion with a deep convolutional neural network. *Inf. Fusion* 36, 191–207.
- Liu, Y., Chen, X., Wang, Z., Wang, Z., Ward, R., Wang, X., 2018. Deep learning for pixel-level image fusion: recent advances and future prospects. *Inf. Fusion* 42, 158–173.
- Liu, Y., Liu, S., Wang, Z., 2015a. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf. Fusion* 24, 147–164.
- Liu, Y., Liu, S., Wang, Z., 2015b. Multi-focus image fusion with dense SIFT. *Inf. Fusion* 23, 139–155.
- Liu, S., Lu, Y., Wang, J., Hu, S., J., Z., Zhu, Z., 2020a. A new focus evaluation operator based on max?min filter and its application in high quality multi-focus image fusion. *Multidimens. Syst. Signal Process.* 31, 569–590.
- Liu, S., Ma, J., Yin, L., Li, H., Hu, S., 2020b. Multi-focus color image fusion algorithm based on super-resolution reconstruction and focused area detection. *IEEE Access* 8, 90760–90778.
- Ma, J., Xu, H., Jiang, J., Mei, X., Zhang, X., 2020. DDCGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion. *IEEE Trans. Image Process.* 29, 4980–4995.
- Meher, B., Agrawal, S., Panda, R., Abraham, A., 2019. A survey on region based image fusion methods. *Inf. Fusion* 48, 119–132.
- Naidu, V., 2011. Image fusion technique using multi-resolution singular value decomposition. *Def. Sci. J.* 61, 479–484.
- Nejati, M., Samavi, S., Shirani, S., 2015. Multi-focus image fusion using dictionary-based sparse representation. *Inf. Fusion* 25, 72–84.
- Nencini, F., Garzelli, A., Baronti, S., Alparone, L., 2007. Remote sensing image fusion using the curvelet transform. *Inf. Fusion* 8, 143–156.
- Pajares, G., de la Cruz, J., 2004. A wavelet-based image fusion tutorial. *Pattern Recognit.* 37 (9), 1855–1872.
- Peng, H., Bao, T., Luo, X., Wang, J., Song, X., Riscos-Núñez, A., Pérez-Jiménez, M., 2020a. Dendrite P systems. *Neural Netw.* 127, 110–120.
- Peng, H., Li, B., Wang, J., Song, X., Wang, T., Valencia-Cabrera, L., Pérez-Hurtado, I., Riscos-Núñez, A., Pérez-Jiménez, M., 2020b. Spiking neural P systems with inhibitory rules. *Knowl.-Based Syst.* 188, 105064, 1–10.
- Peng, H., Lv, Z., Li, B., Luo, X., Wang, J., Song, X., Wang, T., Pérez-Jiménez, M., Riscos-Núñez, A., 2020c. Nonlinear spiking neural P systems. *Int. J. Neural Syst.* 30 (10), 2050008, 1–17.
- Peng, H., Wang, J., 2019. Coupled neural P systems. *IEEE Trans. Neural Netw. Learn. Syst.* 30 (6), 1672–1682.
- Peng, H., Wang, J., Pérez-Jiménez, M., Riscos-Núñez, A., 2019. Dynamic threshold neural P systems. *Knowl.-Based Syst.* 163, 875–884.
- Peng, H., Yang, J., Wang, J., Wang, T., Sun, Z., Song, X., Luo, X., Huang, X., 2017. Spiking neural P systems with multiple channels. *Neural Netw.* 95, 66–71.
- Petrovic, V., Xydeas, C., 2004. Gradient-based multiresolution image fusion. *IEEE Trans. Image Process.* 13 (2), 228–237.
- Qu, X., Yan, J., Xiao, H., Zhu, Z., 2008. Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain. *Acta Automat. Sinica* 34 (12), 1508–1514.
- Qu, X., Yan, J., Yang, G.-D., 2009. Multifocus image fusion method of sharp frequency localized contourlet transform domain based on sum-modified-laplacian. *Opt. Precis. Eng.* 17 (5), 1203–1212.
- Redondo, R., Šroubek, F., Fischer, S., Cristóbal, G., 2009. Multifocus image fusion using the log-gabor transform and a multisize windows technique. *Inf. Fusion* 10 (2), 163–171.
- Stathaki, T., 2011. *Image Fusion: Algorithms and Applications*. Academic Press Elsevier.
- Tang, H., Xiao, B., Li, W., Wang, G., 2017. Pixel convolutional neural network for multi-focus image fusion. *Inf. Sci.* 433, 125–141.
- Tian, J., Chen, L., 2012. Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure. *Signal Process.* 92 (9), 2137–2146.
- Toe, A., 1989. Image fusion by a ratio of low-pass pyramid. *Pattern Recognit. Lett.* 9, 245–253.
- Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wang, Z., Ma, Y., Gu, J., 2010. Multi-focus image fusion using PCNN. *Pattern Recognit.* 43 (6), 2003–2016.
- Wang, Y., Xu, S., Liu, J., Zhao, Z., Zhang, C., Zhang, J., 2020. MFIF-GAN: A new generative adversarial network for multi-focus image fusion. *arXiv:2009.09718v3*.
- Xydeas, C., Petrovic, V., 2009. Objective image fusion performance measure. *Electron. Lett.* 36 (4), 308–309.
- Yang, L., Guo, B., Ni, W., 2008. Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. *Neurocomputing* 72 (1–3), 203–211.
- Yin, M., Duan, P., Liu, W., Liang, X., 2017. A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation. *Neurocomputing* 226, 182–191.
- Zeeuw, P., 1998. Wavelets and Image Fusion. CWI Amsterdam (March).
- Zhang, Y., Bai, X., Wang, T., 2017. Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure. *Inf. Fusion* 35, 81–101.
- Zhang, Y., Chen, L., Zhao, Z., Jia, J., Liu, J., 2014a. Multi-focus image fusion based on robust principal component analysis and pulse-coupled neural network. *Optik* 125, 5002–5006.
- Zhang, Q., Guo, B., 2009. Multifocus image fusion using the nonsubsampled contourlet transform. *Signal Process.* 89 (7), 1334–1346.
- Zhang, Y., Liu, Y., 2020. IFCNN: a general image fusion framework based on convolutional neural network. *Inf. Fusion* 54, 99–118.
- Zhang, Q., Liu, Y., Blum, R., Han, J., Tao, D., 2018. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review. *Inf. Fusion* 40, 57–75.
- Zhang, B., Zhang, C., Liu, Y., Wu, J., Liu, H., 2014b. Multi-focus image fusion algorithm based on compound PCNN in surfacelet domain. *Optik* 125, 296–300.
- Zhao, J., Laganière, R., Liu, Z., 2006. Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement. *Int. J. Innovative Comput. Inf. Control* 3 (6), 1433–1447.
- Zhou, Z., Li, S., Wang, B., 2014. Multi-scale weighted gradient-based fusion for multi-focus images. *Inf. Fusion* 20 (1), 60–72.