

Wavelet transform and texture recognition based on spiking neural network for visual images

Zhenmin Zhang, Qingxiang Wu*, Zhiqiang Zhuo, Xiaowei Wang, Liuping Huang

College of Photonic and Electronic Engineering, Fujian Provincial Key Laboratory for Photonics Technology, Fujian Normal University, Fuzhou 350007, China



ARTICLE INFO

Article history:

Received 22 November 2013

Received in revised form

27 February 2014

Accepted 5 March 2014

Available online 23 October 2014

Keywords:

Spiking neural networks

Human visual system

Fast wavelet transform

Image reconstruction

Texture classification

ABSTRACT

The functionalities of spiking neurons can be applied to deal with biological stimuli and explain complicated intelligent behaviors of the brain. The wavelet transforms are widely used in image feature extraction and image compression. Based on the principles from the visual system and wavelet theory, spiking neural networks with the ON/OFF neuron pathways inspired from the human visual system are proposed to perform the fast wavelet transform and the reconstruction for visual images. By this way we try to simulate how the human brain uses the volition-controlled method to extract useful image information. Furthermore, we decompose each texture sample with the established networks and calculate the normalized energy of the obtained sub-images at different scales. These energy values are used as features for texture classification. The simulation results show that the spiking neural network can extract the main information of images so that the images can be accurately classified using the information.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The human visual system has the ability to selectively attend to certain locations while ignoring others in a typical complexity of the visual environment [1]. The ability is crucial for reducing the amount of visual information to manageable levels in a computation system and for optimizing behavioral performance and response times. In the research of efficient visual processing models, visual attention has been studied by researchers along two main directions [2]. The first one is a bottom-up approach in which a saliency map is obtained by finding salient positions based on a low level representation of an image. The second one is a top-down approach in which an attention map is obtained by volition-controlled signals from a high level in the brain. In this paper wavelet transform networks of spiking neurons are proposed to extract the attention key features which are applied to texture classification.

Since implementation of Hodgkin–Huxley Spiking Neuron Model in 1952 [3] will encounter a very high computational complexity if it is applied to a large scale network, therefore, the simplified conductance-based integrate-and-fire model [4] is used for each neuron in Spiking Neuron Networks (SNN) in this paper because the behavior of this neuron model is very close to the Hodgkin–Huxley model. In the human visual system, there are

various receptive fields from simple cells in the striate cortex to those of the retina and lateral geniculate nucleus [5–7]. Within the retina, information travels from the photoreceptors to the bipolar cells and then on to the ganglion cells [8]. The visual images are transferred among these neurons in the form of spiking trains through the ON or OFF pathways [9]. An ON response increases the amount of neurotransmitters released and an OFF response reduces this amount. Generally, the ON/OFF type neurons can be used to construct specific ON/OFF pathways in the visual system. So, in the integrate-and-fire model, it is assumed that each neuron receives spike trains through excitatory synapse with strength W_{ON} for ON neurons and through inhibitory synapse with strength W_{OFF} for the OFF neurons [9,10]. Different ON/OFF pathways are used to construct the specific networks for wavelet transform in a biological manner.

As wavelet transform can reserve the signal information in time domain and frequency domain simultaneously and efficiently extract the key features of images [11–13], in this paper, two spiking neural networks are proposed. The first network is to mimic behaviors of spiking neurons in the human visual system for wavelet transform and extract the key information of visual images. The second network is used to reconstruct the original image with the key information.

On the other hand, because the textures can provide important characteristics for surface and object identification from aerial or satellite photographs, biomedical images, and many other types of images, much research work has been done on texture analysis and classification in last three decades [14]. Despite the effort, texture analysis is still considered to be an interesting but difficult

* Corresponding author.

E-mail addresses: qxwu@fjnu.edu.cn, min1011@126.com (Q. Wu).

problem. Recent developments in spatial/frequency analysis such as the Gabor transform, Wigner distribution, and wavelet transform provide good multi-resolution analytical tools and should be helpful in overcoming these difficulties [15–17]. Today, it is confirmed that multi-resolution analysis is a useful tool for texture feature extraction during image processing [18].

Textures play important roles in many image processing applications, since images of real objects often do not exhibit regions of uniform and smooth intensities, but variations of intensities with certain repeated structures or patterns, referred to as visual texture [19]. A large class of natural textures can be modeled as a quasi-periodic pattern and detected by highly concentrated spatial frequencies and orientations [20]. A recent study of the human vision system indicates that the spatial frequency representation, which preserves both global and local information, is adequate for quasi-periodic signals. This observation has motivated researchers to develop multi-resolution texture models. In this paper, we decompose each texture sample based on the established neuron networks and calculate the normalized energy of the sub-images at different scales. Ultimately the selected quasi-periodic texture images have been classified with these energy values. The simulation results show that the proposed spiking neural network can extract the main information of images. Using the information, the quasi-periodic texture can be classified accurately.

This paper is organized as follows. A brief review of 2-D wavelet transform and invert transform is described in Section 2. Section 3 describes the proposed SNN for wavelet transform in details. Image reconstruction with SNN is described in Section 4. In Section 5, some experimental results of texture classification are presented based on the extracted energy from SNN and KNN (K-Nearest Neighbor) algorithm. Several traditional multi-resolution techniques and our method are compared. Finally, the conclusions are summarized in Section 6.

2. Fast wavelet transform and inverse wavelet transform

By wavelet transform, the decomposition of a signal with a family of real orthonormal bases $\psi_{m,n}(x)$ obtained through translation and dilation of a kernel function $\psi(x)$ known as the mother wavelet, i.e.,

$$\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m}x - n) \quad (1)$$

where m and n are integers. Due to the orthonormal property, the wavelet coefficients of a signal $f(x)$ can be easily computed via

$$c_{m,n} = \int_{-\infty}^{+\infty} f(x) \psi_{m,n}(x) dx \quad (2)$$

and the synthesis formula

$$f(x) = \sum_{m,n} c_{m,n} \psi_{m,n}(x) \quad (3)$$

can be used to recover $f(x)$ from its wavelet coefficients.

Based on the multi-resolution theory and the wavelet analysis theory, Mallat proposed fast wavelet transform (FWT) which was also called pyramid structured wavelet transform and inverse wavelet transform (IWT) in 1987 [21,22]. The flow chart of the two-dimensional FWT is shown in Fig. 1.

As shown in Fig. 1, $f(m,n)$ is the digital input signal, h_0 is the impulse response of the low-pass filter, h_1 is the impulse response of the high-pass filter, h_0 and h_1 form the analytical filter bank and $2\downarrow$ represent the interval down-sampling. Firstly, the input signal is passed through the high-pass and low-pass filters, respectively, and down-sampled. Then the two obtained signals are passed separately through the two filters, and down-sampling is performed again. Finally, the four signals ultimately obtained are

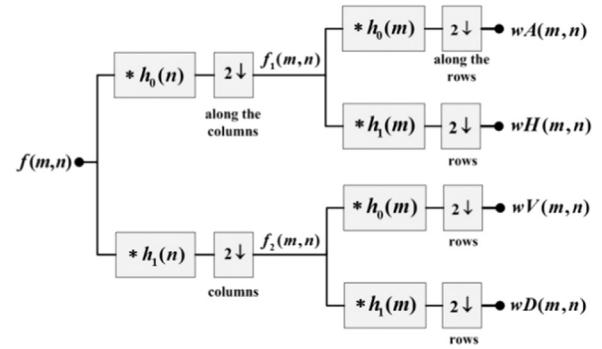


Fig. 1. Achieving 2D-FWT with the analytical filter bank and down-sampling.

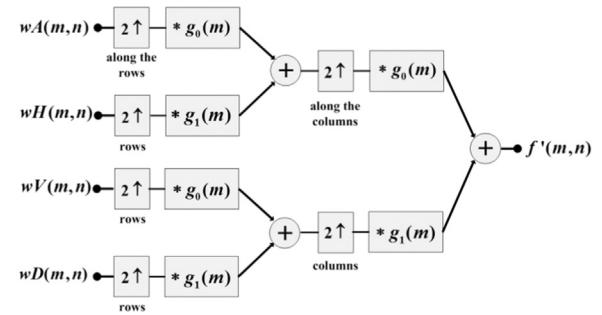


Fig. 2. Achieving 2D-IWT with up-sampling and the integrated filter bank.

approximate coefficients, horizontal details, vertical details and diagonal details of the wavelet transform, which are denoted, respectively, by wA , wH , wV and wD . Meanwhile, the resolution of these final results is a quarter of the original image.

Utilizing the four wavelet coefficients, inverse wavelet transform can be implemented, that is, to reconstruct the original visual image. The flow chart of the two-dimensional IWT is shown in Fig. 2.

In Fig. 2, input signal is the four wavelet coefficients, $2\uparrow$ represent the interval up-sampling, g_0 and g_1 form the integrated filter bank, which can be easily computed via

$$g_0(n) = (-1)^{n+1} h_1(n) \quad (4)$$

$$g_1(n) = (-1)^n h_0(n). \quad (5)$$

After the calculation as shown in Fig. 2, the original image is reconstructed.

3. Spiking neural network model for fast wavelet transform (SNN-FWT)

The human visual system has a superior performance in feature extraction. There are various receptive fields from simple cells in the striate cortex to those of the retina and lateral geniculate nucleus [6]. Based on the Mallat fast wavelet transform algorithm and mechanism inspired by the ON/OFF pathways in the visual system [10,23,24], a spiking neural network model is proposed to perform the fast wavelet transform of the visual images as shown in Fig. 3. Because the calculation of the Hodgkin-Huxley spiking neuron model is too complicated, the simplified conductance-based integrate-and-fire neuron model is used to simulate the network.

In the above spiking neural networks, each neuron in the input neuron array generates spikes induced by a synapse current of photonic receptor in terms of the corresponding pixel brightness in a visual image. The dimension of the input neuron array is

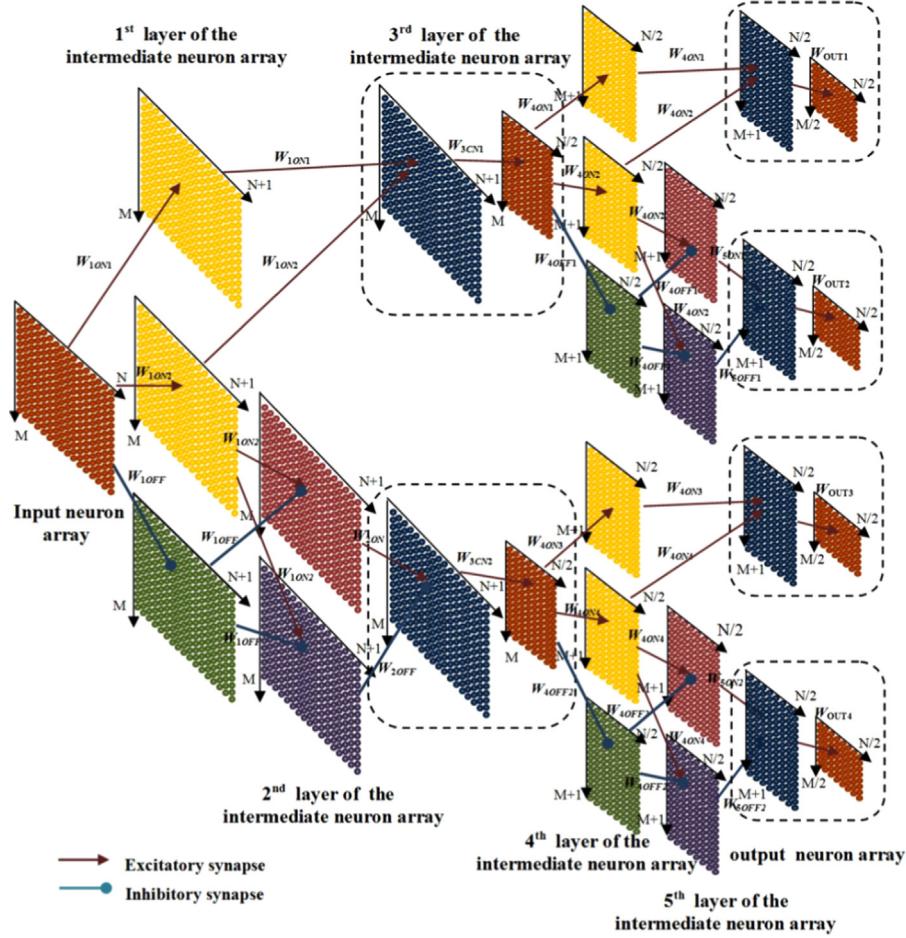


Fig. 3. Spiking neural network for fast wavelet transform.

$M \times N$. Each neuron in the array is labeled with $IN(m, n)$, where $m=1, \dots, M$ and $n=1, \dots, N$. Each pixel of the image corresponds to a receptor. Assume that $G_{m,n}(t)$ represent the gray scale of an image pixel and each photonic receptor transfers the pixel brightness to a synapse current $I_{m,n}(t)$ [25–27]. The synapse current $I_{m,n}(t)$ of the integrate-and-fire model can be represented as follows:

$$\frac{dI_{m,n}(t)}{dt} = -\frac{1}{\tau} I_{m,n}(t) + \alpha G_{m,n}(t) \quad (6)$$

where α is a constant for transformation from gray scale to current and τ is a time constant for the decay of the synapse current. The neuron potential $v_{m,n}(t)$ is governed by the following equation:

$$c \frac{dv_{m,n}(t)}{dt} = g_l(E_l - v_{m,n}(t)) + I_{m,n}(t) + I_0 \quad (7)$$

where g_l is the membrane conductance, E_l is the reverse potential, c represents the membrane capacitance and I_0 is simulated by an average current produced by background noise. If the membrane potential passes threshold v_{th} , then the neuron generates a spike. Let $S_{m,n}(t)$ represent the spike train generated by the neuron such as that

$$S_{m,n}(t) = \begin{cases} 1 & \text{if neuron } (m, n) \text{ fires at time } t \\ 0 & \text{if neuron } (m, n) \text{ does not fire at time } t. \end{cases} \quad (8)$$

The first layer of the intermediate neuron array is composed of three neuron arrays as shown in Fig. 2: the first two arrays are the ON neuron arrays and the third is the OFF neuron array. The ON/OFF neuron arrays have the same dimension $M \times (N+1)$. Neurons in the ON/OFF neuron arrays are labeled with $1ON1(p, q)$, $1ON2(p, q)$ and

$1OFF(p, q)$, where $p=1, \dots, M$ and $q=1, \dots, N+1$. The convolutions of wavelet transform between the input signal and the impulse responses are converted to the accumulation of different neural arrays. According to the above equations, gray scale has been transferred to spike trains. Assume that these spike trains are transferred to the ON/OFF neuron arrays through excitatory synapses $W_{1ON1(p,q)}$ and $W_{1ON2(p,q)}$ and inhibitory synapse $W_{1OFF(p,q)}$. Based on the principle of the FWT, synapse strength distribution can be set as follows:

$$W_{1ON1(p,q)} = a_{ON}f(p, q) \quad (9)$$

$$W_{1ON2(p,q+1)} = a_{ON}f(p, q) \quad (10)$$

$$W_{1OFF(p,q)} = -a_{OFF}f(p, q) \quad (11)$$

where $1 \leq p \leq M$, $1 \leq q \leq N$. According to the Haar filter coefficients, the weight coefficient is set as $a_{ON}=1/\sqrt{2}$, $a_{OFF}=-1/\sqrt{2}$. The total synapse currents $I_{1ON1(p,q)}(t)$, $I_{1ON2(p,q)}(t)$ and $I_{1OFF(p,q)}(t)$ are set as follows:

$$\frac{dI_{1\sigma(p,q)}(t)}{dt} = -\frac{1}{\tau} I_{1\sigma(p,q)}(t) + \sum_{p=1}^M \sum_{q=1}^{N+1} W_{1\sigma(p,q)} \beta_1 S_{p,q}(t) \quad (12)$$

where $\sigma \in \{ON, OFF\}$. Then the neuron potential in the ON/OFF array is governed by the following equation:

$$c \frac{dv_{1\sigma(p,q)}(t)}{dt} = g_l(E_l - v_{1\sigma(p,q)}(t)) + I_{1\sigma(p,q)}(t) + I_0 \quad (13)$$

Let $S_{\sigma(p,q)}(t)$ represent a spike train which is generated by the ON/OFF neurons such that

$$S_{\sigma(p,q)}(t) = \begin{cases} 1 & \text{if ON/OFF neuron } (p,q) \text{ fires at time } t. \\ 0 & \text{if ON/OFF neuron } (p,q) \text{ does not fire at time } t. \end{cases} \quad (14)$$

The second layer of the intermediate neuron array is composed of two $M \times (N+1)$ neuron arrays. Neurons in these arrays are labeled with $2ON(p,q)$ and $2OFF(p,q)$. Each neuron receives spike trains through excitatory synapse with strength $W_{2ON(p,q)}$ for ON neurons and through inhibitory synapse with strength $W_{2OFF(p,q)}$ for the OFF neurons, they are calculated by the following expression:

$$W_{2ON(p,q)} = \begin{cases} W_{1ON2(p,q)} - W_{1OFF(p,q)}, & \text{if } W_{1ON2(p,q)} - W_{1OFF(p,q)} > 0 \\ 0, & \text{if } W_{1ON2(p,q)} - W_{1OFF(p,q)} \leq 0 \end{cases} \quad (15)$$

$$W_{2OFF(p,q)} = \begin{cases} -(W_{1ON2(p,q)} - W_{1OFF(p,q)}), & \text{if } W_{1ON2(p,q)} - W_{1OFF(p,q)} < 0 \\ 0, & \text{if } W_{1ON2(p,q)} - W_{1OFF(p,q)} \geq 0 \end{cases} \quad (16)$$

where $1 \leq p \leq M$, $1 \leq q \leq N+1$. The synapse current is represented as follows:

$$\frac{dI_{2\sigma(p,q)}(t)}{dt} = -\frac{1}{\tau} I_{2\sigma(p,q)}(t) + \sum_{p=1}^M \sum_{q=1}^{N+1} W_{2\sigma(p,q)} \beta_2 S_{p,q}(t) \quad (17)$$

The neuron potential in the ON/OFF array is governed by the following equation:

$$C \frac{dv_{2\sigma(p,q)}(t)}{dt} = g_l(E_l - v_{2\sigma(p,q)}(t)) + I_{2\sigma(p,q)}(t) + I_0 \quad (18)$$

The third layer of the intermediate array is still composed of two $M \times (N+1)$ neuron arrays. Neurons in these arrays are labeled with $3CN1^*(p, q)$ and $3CN2^*(p, q)$. The synapses strength distribution can be calculated by the following expression:

$$W_{3CN1^*(p,q)} = W_{1ON1(p,q)} + W_{1ON2(p,q)} \quad (19)$$

$$W_{3CN2^*(p,q)} = W_{2ON(p,q)} - W_{2OFF(p,q)} \quad (20)$$

The synapse current is set as the following equations:

$$\begin{aligned} \frac{dI_{3CN1^*(p,q)}(t)}{dt} &= -\frac{1}{\tau} I_{3CN1^*(p,q)}(t) + \sum_{p=1}^M \sum_{q=1}^{N+1} W_{1ON1(p,q)} \beta_3 S_{p,q}(t) \\ &\quad + \sum_{p=1}^M \sum_{q=1}^{N+1} W_{1ON2(p,q)} \beta_3 S_{p,q}(t) \end{aligned} \quad (21)$$

$$\begin{aligned} \frac{dI_{3CN2^*(p,q)}(t)}{dt} &= -\frac{1}{\tau} I_{3CN2^*(p,q)}(t) + \sum_{p=1}^M \sum_{q=1}^{N+1} W_{2ON(p,q)} \beta_3 S_{p,q}(t) \\ &\quad - \sum_{p=1}^M \sum_{q=1}^{N+1} W_{2OFF(p,q)} \beta_3 S_{p,q}(t) \end{aligned} \quad (22)$$

where $\beta_1, \beta_2, \beta_3$ is a constant, $S_{p,q}(t)$ represent a spike train.

After the accumulation of signals, only the neurons of the even-numbered columns of the $3CN1$ and $3CN2$ neuron layer generate spikes, while the neurons of the odd-numbered columns do not fire. Then two new neuron arrays are obtained, which are composed of the neurons of the even-numbered columns of the $3CN1^*$ and $3CN2^*$ neuron layer. Neurons in these arrays are labeled with $3CN1(p, q)$ and $3CN2(p, q)$, both of them have the same dimension $M \times (N/2)$. Synapse strength distribution can be set as follows:

$$W_{3CN1(p,q)} = W_{3CN1^*(p,2k)} \quad (23)$$

$$W_{3CN2(p,q)} = W_{3CN2^*(p,2k)} \quad (24)$$

where $k = 1, 2, \dots, N/2$, $1 \leq p \leq M$, $1 \leq q \leq N/2$.

Thereafter, the remaining synapse strength distribution of the network can be set in a similar iteration and down-sampling manner, and eventually we will obtain four outputs from neuron arrays $OUT1$, $OUT2$, $OUT3$ and $OUT4$ as the bottom layer, and the firing rate for these layers is calculated by the following expression:

$$r_{OUT(i)(m,n)}(t) = \frac{1}{T} \sum_t^{t+T} S_{OUT(i)(m,n)}(t) \quad (25)$$

where $i=\{1,2,3,4\}$ and $S_{OUT(i)(m,n)}(t)$ represent the spike train generated by the four output neuron arrays. A firing rate map can be obtained by plotting $r_{OUT(i)(m,n)}(t)$. Let $r_{max\{i\}}$ represent the maximum firing rate in the four neuron arrays. The final results with 255 Gy scale levels are obtained using the following equation:

$$R_{i|m,n}(t) = \frac{255}{r_{max\{i\}}} r_{OUT(i)(m,n)}(t) \quad (26)$$

The obtained four groups of neuron arrays correspond to the four coefficients of the wavelet transform.

The above network model is simulated by using the Euler method with a time step of 0.1 ms by Matlab. The following parameters were used in the experiments corresponding to biological neurons: $v_{th} = -60$ mv, $E_l = -70$ mv, $g_l = 1.0 \mu\text{s/mm}^2$, $c = 8 \text{nF/mm}^2$, $\tau = 16$ ms, $T = 400$ ms, $\alpha = 0.02$, $\beta_1 = 4.3$, $\beta_2 = 5.1$, $\beta_3 = 7.9$ and $I_0 = 7 \mu\text{A}$.

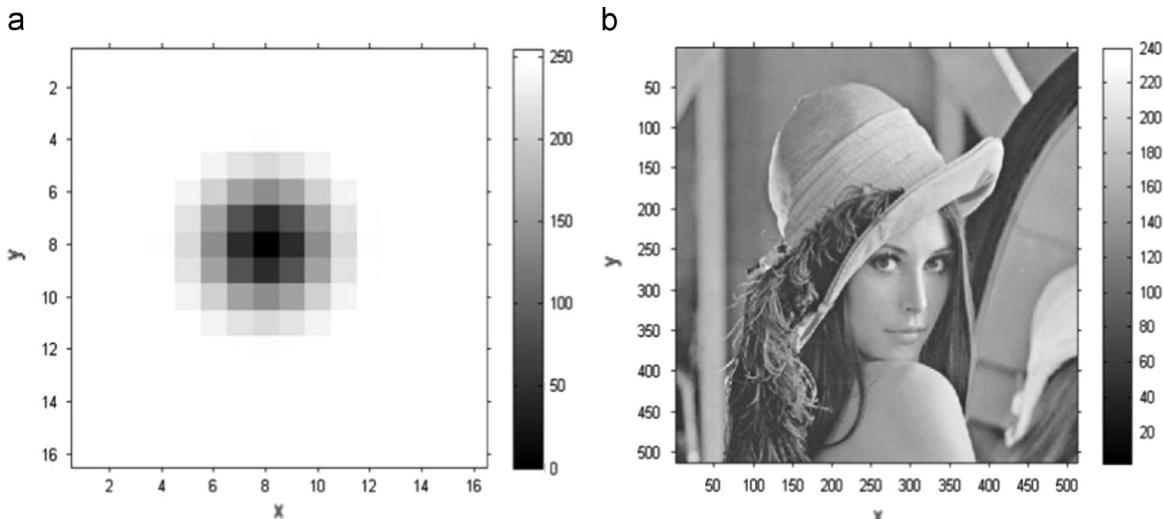


Fig. 4. Test images: (a) Gaussian gray scale distribution and (b) Lena image.

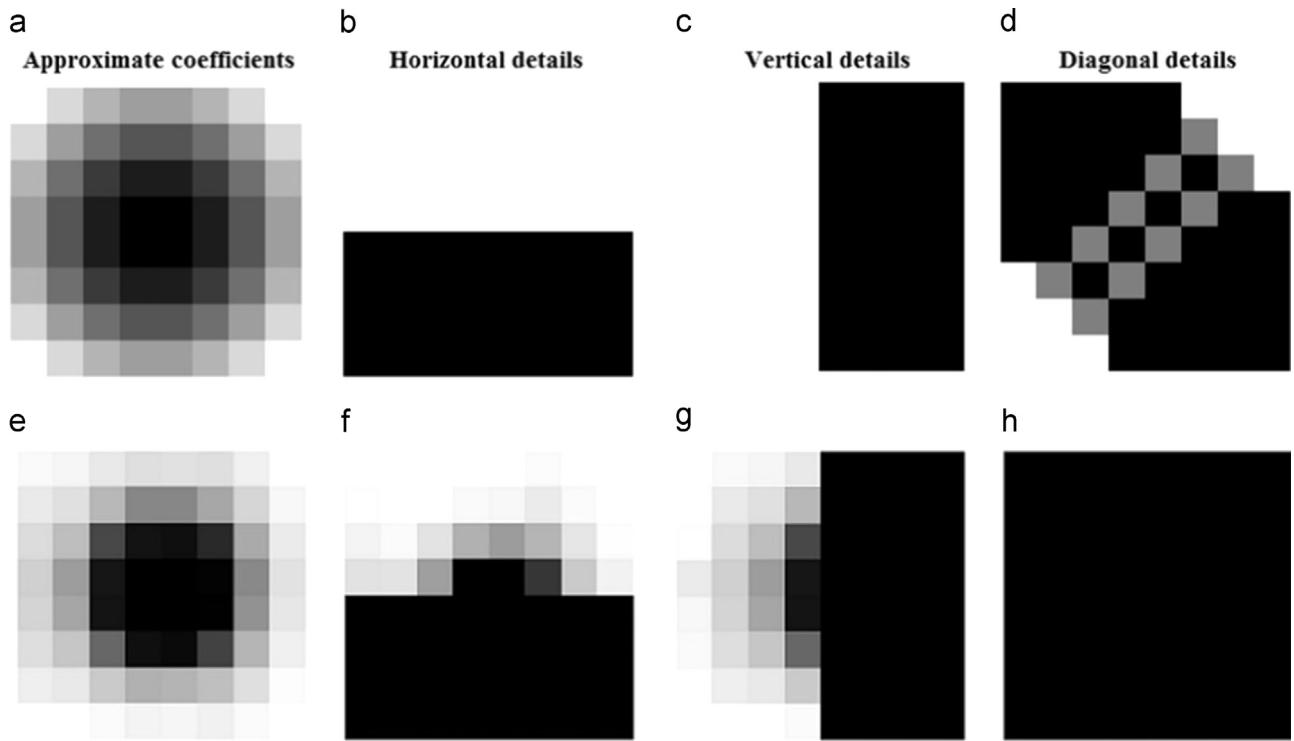


Fig. 5. Wavelet transform by Mallat method (a–d) and by SNN-FWT (e–h) of Gaussian image.

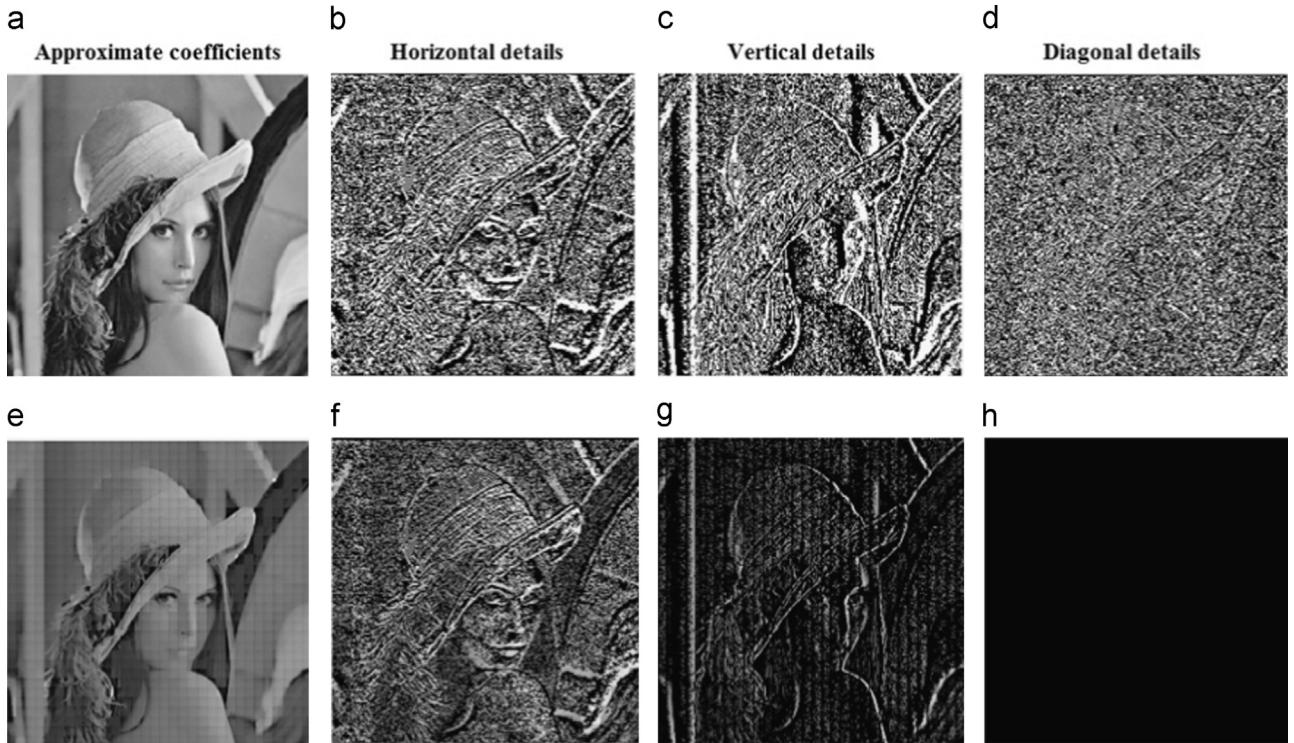


Fig. 6. Wavelet transform by Mallat method (a–d) and by SNN-FWT (e–h) of Lena.

These parameters can be adjusted to get a good quality output image.

The two images as shown in Fig. 4 are used to test the network model. Fig. 4(a) is Gaussian gray scale distribution in 16×16 area. Fig. 4(b) is Lena image which is widely used as a benchmark image in the image processing domain.

Fig. 5(a)–(d) shows the four coefficients of the wavelet transform of Gaussian image obtained by the Mallat method. Fig. 5(e)–(h)

displays similar results obtained by the spiking neural network. In Fig. 5, the dimensions of all the images are 8×8 and the resolution of these results is a quarter of the original image.

In order to demonstrate this behavior of the network further, the Lena image was used for transfer. Since the Lena image (512×512) exceeds the Matlab predetermined matrix dimension, therefore the image has been divided into 32×32 blocks and each block contains 16×16 pixels. The operation results are

shown in Fig. 6. Fig. 6(a)–(d) shows the result of the mathematical method, while Fig. 6(e)–(h) comes from the spiking neural network.

The above examples show that the proposed network can perform feature extraction similar to wavelet transform and focus on key features ignoring the weak noise. From the obtained results



Fig. 7. (a) Original image; approximate coefficients obtained by SNN-FWT in the first scale (b) and in the second scale (c).

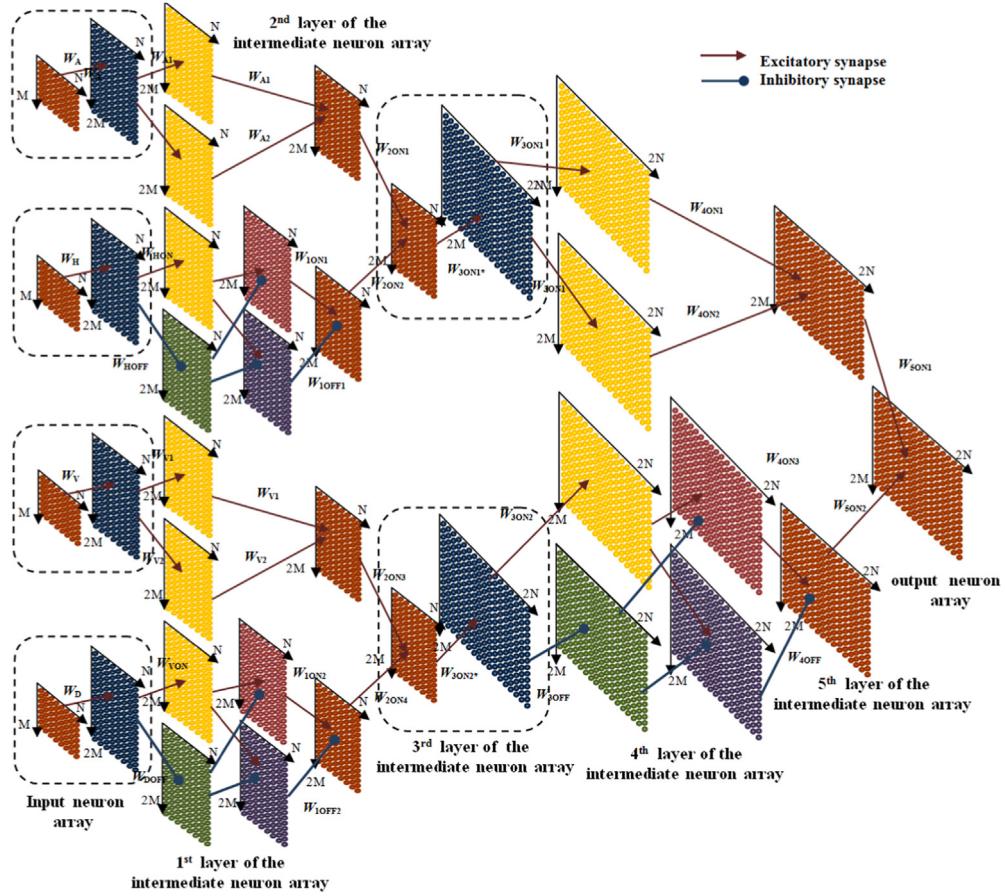


Fig. 8. Spiking neural network for inverse wavelet transform.

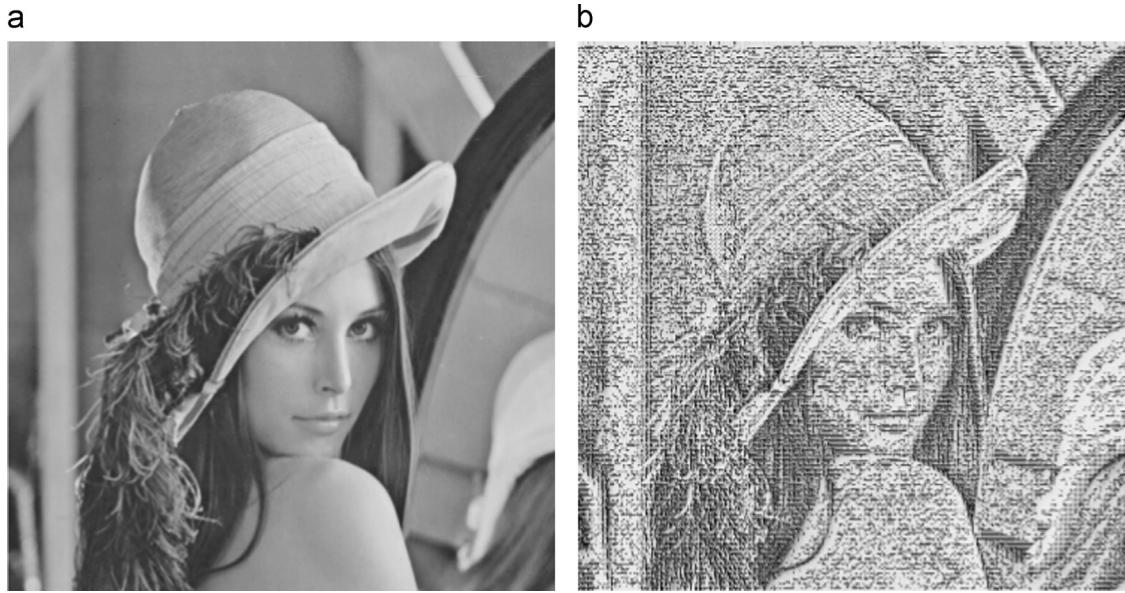


Fig. 9. Comparison of the original Lena image (a) and the reconstructed image (b).

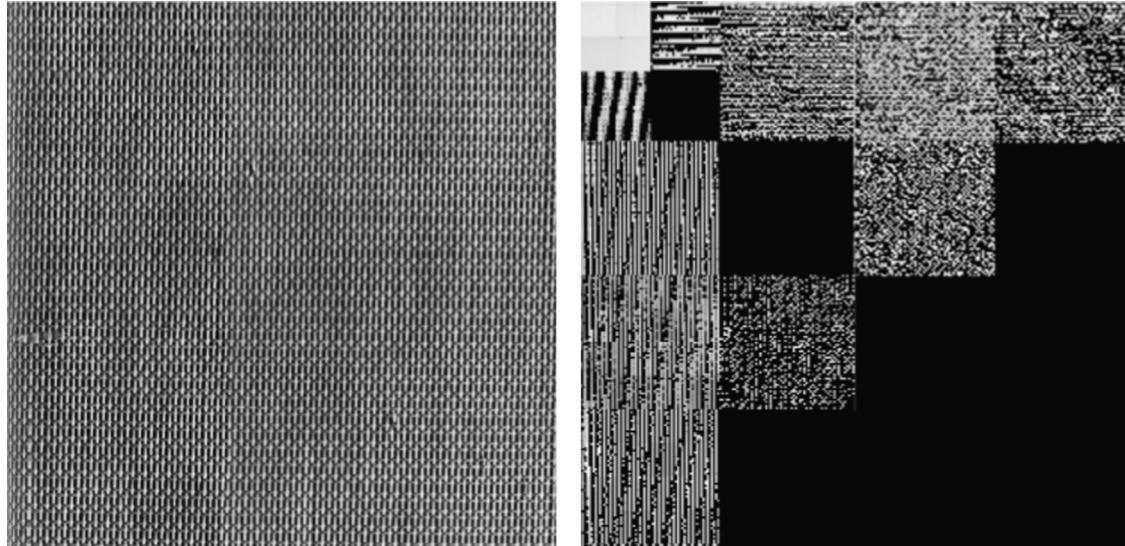


Fig. 10. D53 and its three-level SNN wavelet transform.

by the SNN-FWT method, it can be seen that the low-frequency and middle-frequency information contained in the approximate coefficients, horizontal details, vertical details are well preserved. Because pulse transmission has a threshold limit in the spiking neural networks, the information contained in the weak high-frequency component diagonal details is ignored. It is known that the high-frequency components of a picture include little information and the main information is included in the low-frequency and middle-frequency components. So, relative to the wavelet transform performed through mathematical methods, although the visual image signals pass through a complex spiking neural network and some details are ignored, the main information is still retained to achieve the purpose of the key feature extraction and image reconstruction.

Fig. 7 shows the comparison of the original image and the sub-images which are decomposed by SNN-FWT in the first scale and in the second scale. As the signal through the two sets of spiking neural networks, the key information of the image is extracted.

4. Spiking neural network model for inverse wavelet transform (SNN-IWT)

Using the four wavelet transform coefficients obtained by SNN-FWT, combined with the inverse wavelet transform algorithm and the ON/OFF pathways mechanism, another spiking neural network for image reconstruction is proposed, which is shown in Fig. 8.

In Fig. 8, there are four input neuron arrays corresponding to the four wavelet coefficients, each neuron in these arrays generates spikes induced by a synapse current in terms of pixel brightness. The dimension of these arrays is $M \times N$. In order to achieve the up-sampling, we have inserted N columns of new neurons into each input array at intervals. In addition, the setting mode of the synapse strength distribution of the SNN-IWT is shown in Fig. 8, which has been set in a manner consistent with the synapse strength distribution of the SNN-FWT. Furthermore, the synapse currents, neuron potential and spike train generated by the spiking neurons in the network in different arrays can be governed by Eqs. (6)–(8), and the excitatory synapse strength for the ON

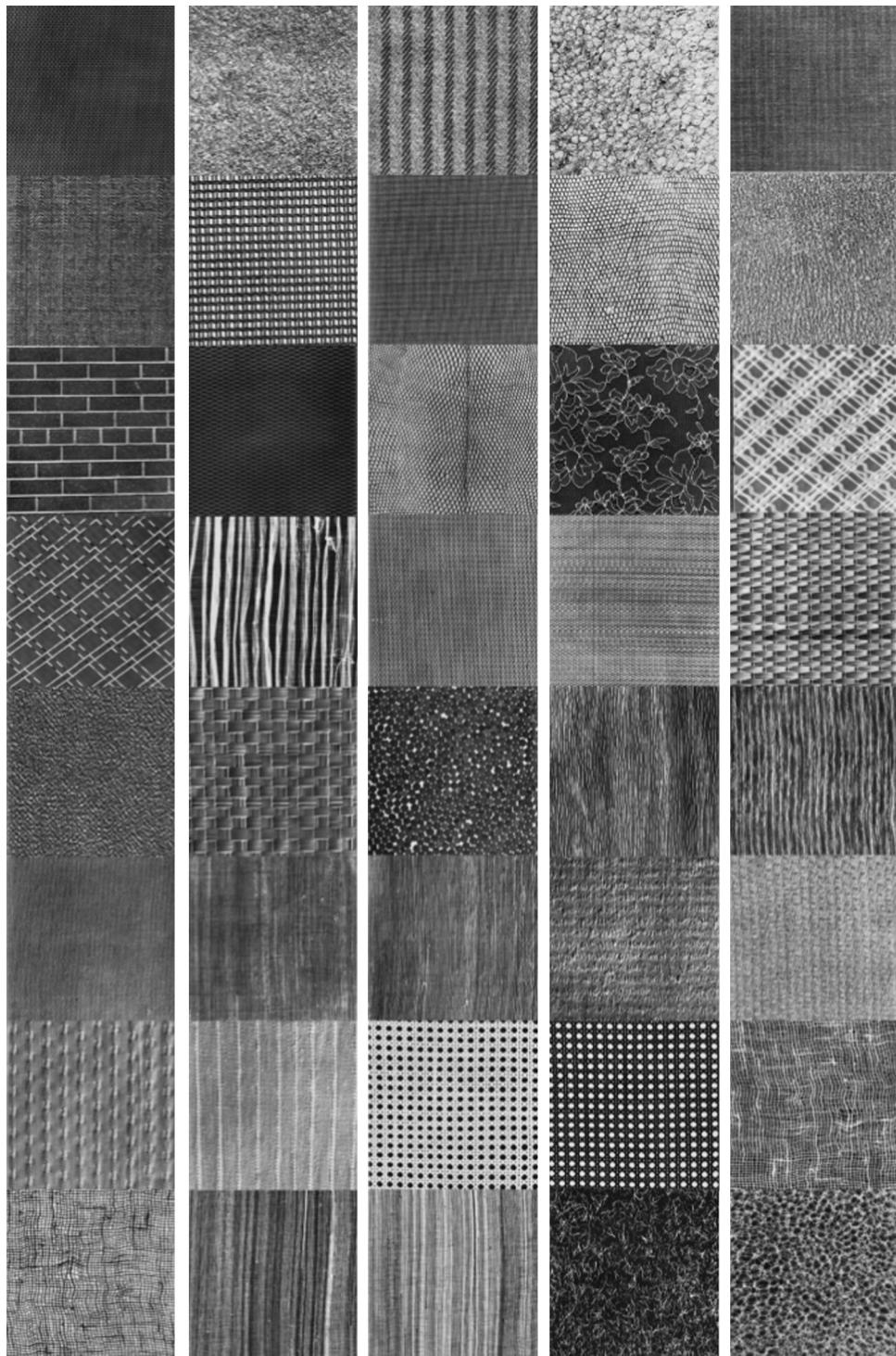


Fig. 11. Forty textures used in the experiments.

neurons and inhibitory synapse strength for the OFF neurons are calculated by expressions (14) and (15).

Finally, the firing rate for output neuron layers is calculated by formula (26). A firing rate map can be obtained by plotting the firing rate. After normalization, the final results with 255 Gy scale levels can be obtained.

The reconstructed Lena image in Fig. 9 indicates that, since part of the signal is ignored during transmission between the different arrays of the spiking neuron networks, making the reconstructed image is vague relative to the original image, but the reconstructed image still retains most of the important information. For example,

it contains the outline of the main character, facial features. The outline and detailed features of the hat and background profile are also well restored. Based on this property of the network, the important information can be extracted for texture classification to achieve a good performance.

5. Texture classification based on SNN-FWT

Texture provides essential information for many image classification tasks. Extensive research has been done on texture classification during the past three decades [28–30].

Table 1

Comparison of the experimental results using different methods.

ID	SNN-FWT (%)	LRM-WT (%)	F16b (%)	Wavelet and GLCM (%)	TSWT (%)	PSWT (%)
D6	100	100	95.122	100	88.225	68.596
D9	80	97.531	80.488	95.122	50.941	39.136
D11	100	97.531	68.298	85.366	55.756	39.074
D14	100	93.827	100	100	90.154	73.318
D16	100	98.765	95.122	100	95.278	74.799
D17	100	95.062	80.488	97.651	60.509	44.398
D20	100	98.765	95.122	100	98.241	88.812
D21	97.5	100	100	100	100	96.343
D22	100	93.827	92.683	97.561	84.969	68.889
D24	100	98.765	70.732	95.122	58.241	42.207
D26	90	98.765	97.561	100	92.114	68.858
D34	100	98.765	97.561	100	81.728	70.833
D36	100	100	95.122	100	95.679	51.25
D41	100	90.123	82.927	92.683	45.324	33.827
D46	100	98.765	100	100	96.96	90.278
D47	100	98.765	100	100	97.886	74.367
D51	100	100	100	100	92.207	69.228
D53	100	96.296	100	100	92.577	70.247
D55	100	97.531	78.049	100	83.704	57.238
D56	100	97.531	97.561	100	91.574	73.133
D57	100	86.42	51.22	87.802	75.725	60.694
D64	97.5	98.765	100	100	94.383	61.713
D66	97.5	93.827	100	97.561	87.315	73.58
D68	95	93.827	100	92.683	87.361	62.685
D76	100	96.296	92.683	97.561	67.022	46.713
D77	100	98.765	97.561	100	77.824	47.824
D78	100	98.765	85.366	92.683	67.963	46.142
D79	100	96.296	80.488	90.244	61.188	43.688
D80	100	100	85.366	87.805	62.114	37.253
D82	100	98.765	65.854	100	73.904	50
D83	100	100	70.732	100	71.019	39.182
D85	100	97.531	87.805	100	62.901	38.92
D101	100	100	100	100	100	38.904
D102	100	93.827	100	100	90.478	87.809
D103	100	98.765	100	100	99.907	90.571
D104	100	98.765	100	100	99.846	92.114
D105	100	98.296	82.927	95.122	76.049	54.815
D106	100	95.062	95.122	92.683	66.343	52.346
D109	100	97.531	92.683	90.244	66.235	52.346
D111	100	100	87.805	80.488	56.991	39.676
Average	98.938	97.151	90.061	96.707	79.166	61.588

The traditional FWT recursively decompose sub-signals in the low-frequency channels [31]. However, since the most significant information of a texture not only appears in the low-frequency channels, it may also appear in the middle-frequency channels. The further decomposition of the conventional wavelet transform just in the lower-frequency region may not help much for the purpose of classification. In order to prove the validity of the proposed SNN-FWT in the above, we use this network to perform multi-scale wavelet transform in different frequency channels. The results of the transformation are used in texture classification.

In the experiments, the common texture sources Brodatz album [32] are used as test images. The D53 is as an example and its three scale SNN wavelet transform is shown in Fig. 10. Firstly, the feature of D53 is extracted by SNN wavelet transform and four wavelet coefficients are obtained. Since most of the information of diagonal details was lost, the energy of the remaining three coefficients is extracted and regarded as the characteristic values. The energy of the sub-images can be represented as

$$e = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x(i,j)| \quad (27)$$

where $x(i,j)$ is the pixel value of the sub-image, $1 \leq i \leq M$ and $1 \leq j \leq N$. Then, continue to decompose the three remaining

coefficients in the second scale and receive energy values of nine new sub-images, which are also as characteristic values. Finally, the last three energy characteristic values are obtained after decomposing the approximate coefficients in the third scale.

In this section, the performance is verified using the KNN classification algorithm based on the fifteen feature values obtained from SNN-FWT. 40 quasi-periodic textures, as shown in Fig. 11, are from the Brodatz album. Every original image is of size 640×640 pixels with 256 Gy levels. 100 sample images of size 256×256 are extracted from each original image and used in the experiments. These 4000 texture images are separated into two sets. The training set includes 2400 images and the test set includes 1600 images.

The proposed method is compared with other traditional multi-resolution methods [33–34], such as the combination of the wavelet transform the gray level co-occurrence matrices (GLCM) [35], the linear regression model (LRM) wavelet transform [30], the pyramid-structured wavelet transform (PSWT) [21], the tree-structured wavelet transform (TSWT) [14] and the F16b filter bank, which is the quadrature mirror bank designed by Johnston and was used for texture analysis by Randen [28,36,37].

Table 1 shows the retrieval accuracy of these different multi-resolution methods for 40 textures.

From the experimental results shown in Table 1, the method of using the SNN-FWT to gain characteristic values of texture images and achieving image classification by KNN algorithm shows an outstanding performance. The classification rate of most images reached 100%, except for the individual image, the classification accuracy is slightly lower. The experimental results show that the proposed method has a higher average classification accuracy than other traditional multi-resolution methods. In addition, KNN algorithm has also proved that it is very suitable for features classification in this paper. The results also indicate that the proposed SNN-FWT can reserve the main information of images and perform texture image classification very well.

6. Discussion

In this paper, we propose an integrate-and-fire spiking neuron network combining visual ON/OFF neuron pathways and synapse current mechanism to extract features from a visual image. The algorithm of the network is based on the spiking neuron model. It is shown that the network performs a transform similar to Mallat fast wavelet.

In the process of building the model, different neuron arrays are used to perform the convolutions of wavelet transform, while the firing neurons are selected according to down-sampling algorithm, and new neurons are added to achieve up-sampling algorithm. Meanwhile, it is also very important to select appropriate synapse strength distribution coefficients in the operation. The simulation results show that spiking neuron-based system is able to perform fast wavelet transform and image reconstruction. In addition, it is shown that the key information can be obtained when the visual image signals pass through a complex spiking neural network; perhaps this explains how the human brain extracts an attention map by volition-controlled signals from a high level in the brain.

The proposed network is inspired by the human visual system and can be used for spiking neuron-based intelligent system to extract key information from the visual images. In this paper, we designed a new approach to texture analysis and classification with the information extract from SNN-FWT combined with the KNN algorithm, and this model is presented and its good performance on the classification accuracy is demonstrated in the experiments. Although the traditional multi-resolution methods, like PSWT, TSWT and f16b filter bank, are suitable for some textures, the proposed

method is not only natural and effective for spiking neuron-based system, but also obtains outstanding performance.

However, there are still some issues that need further discussion. Further research can focus on application of the behaviors for object recognition. In this paper, we set the synapse strength distributions based on principles taken from the fast wavelet transform theory. It can be tried for obtaining the strength using the spike timing-dependent plasticity of synapses in future.

Acknowledgments

The authors gratefully acknowledge the fund from the Natural Science Foundation of China (Grant nos. 61179011 and 61070062) and the Science and Technology Major Projects for Industry-Academic Cooperation of Universities in Fujian Province (Grant no. 2013H6008).

References

- [1] Q.X. Wu, T.M. McGinnity, L. Maguire, R.T. Cai, et al., A visual attention model based on hierarchical spiking neural networks, *Neurocomputing* 116 (2013) 3–12.
- [2] E. Niebur, C. Koch, Computational architectures for attention, in: R. Brain (Ed.), *The Attentive*, Parasuraman, MIT Press, Cambridge, MA, 1998.
- [3] A.L. Hodgkin, A.F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* 117 (4) (1952) 500–544.
- [4] E. Müller, *Simulation of High-Conductance States in Cortical Neural Networks* (Master's thesis), University of Heidelberg, Heidelberg, 2003.
- [5] E.R. Kandel, J.H. Schwartz, *Principles of Neural Science*, Edward Arnold (Publishers) Ltd, London, 1981.
- [6] R.H. Masland, The fundamental plan of the retina, *Nat. Neurosci.* 4 (9) (2001) 877–886.
- [7] W.R. Taylor, D.I. Vaney, New directions in retinal research, *Trends Neurosci.* 26 (7) (2003) 379–385.
- [8] B. Jonathan Demb, Cellular mechanisms for direction selectivity in the retina, *Neuron* 55 (2) (2007) 179–186.
- [9] R. Nelson, H. Kolb, On and off pathways in the vertebrate retina and visual system, in: L.M. Chalupa, J.S. Werner (Eds.), *The Visual Neurosciences*, MIT Press, Cambridge, London, 2003, pp. 260–278.
- [10] Q.X. Wu, T.M. McGinnity, L. Maguire, A. Ghani, et al., Spiking neural network performs discrete cosine transform for visual images, *Emerging Intelligent Computing Technology and Applications: With Aspects of Artificial Intelligence*, 5755, (2009) 21–29.
- [11] C.K. Chui, *An Introduction to Wavelets*, Academic Press, New York, 1992.
- [12] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [13] C.L. Liu, A Tutorial of the Wavelet Transform, (<http://disp.ee.ntu.edu/>), (2010).
- [14] T.H. Chang, J. Kuo, Texture analysis and classification with tree-structured wavelet transform, *IEEE Trans. Image Process.* 2 (4) (1993) 429–441.
- [15] C. Bouman, B. Liu, Multiple resolution segmentation of textured images, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (2) (1991) 99–113.
- [16] A.C. Bovik, M. Clark, W.S. Geisler, Multichannel texture analysis using localized spatial filters, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1) (1990) 55–73.
- [17] T.R. Reed, H. Wechsler, Segmentation of textured images and Gestalt organization using spatial/spatial-frequency representations, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1) (1990) 1–12.
- [18] S. Hu, C. Xu, W.Q. Guan, Y. Tang, et al., Texture feature extraction based on wavelet transform and gray-level co-occurrence matrices applied to osteosarcoma diagnosis, *Bio-Med. Mater. Eng.* 24 (1) (2014) 129–143.
- [19] M.S. Zaid, R.J. Jadhav, P.J. Deore, An efficient wavelet based approach for texture classification with feature analysis, in: Proceedings of the 2013 3rd IEEE International Advance Computing Conference (IACC), 2013, pp. 1149–1153.
- [20] L. Cohen, Time-frequency distributions—a review, *Proc. IEEE* 77 (7) (1989) 941–981.
- [21] S. Mallat, *A wavelet tour of signal processing. The Sparse Way*, 3rd edition, Salt Lake City, Academic Press, 2008.
- [22] S.G. Mallat, A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, 11, (1989) 674–693.
- [23] Q.X. Wu, M. McGinnity, L. Maguire, B. Glackin, et al., Learning mechanisms in networks of spiking neurons, studies in computational intelligence, *Studies in Computational Intelligence* 35 (2006) 171–197.
- [24] Q.X. Wu, T.M. McGinnity, L. Maguire, A. Belatreche, et al., 2D co-ordinate transformation based on a spike timing-dependent plasticity learning mechanism, *Neural Netw.* 21 (9) (2008) 1318–1327.
- [25] Q.X. Wu, R.T. Cai, T.M. McGinnity, L. Maguire, et al., Remembering key features of visual images based on spike timing dependent plasticity of spiking neurons, in: Proceedings of the 2009 2nd International Congress on Image and Signal Processing, vol. 1–9, 2009, pp. 2168–2172.
- [26] Q.X. Wu, T.M. McGinnity, L. Maguire, A. Belatreche, et al., Edge detection based on spiking neural network model, in: D.-S. Huang, L. Heutte, M. Loog (Eds.), *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence*, Springer, Berlin Heidelberg, 2007, pp. 26–34.
- [27] Q.X. Wu, T.M. McGinnity, L.P. Maguire, B. Glackin, et al., Learning under weight constraints in networks of temporal encoding spiking neurons, *Neurocomputing* 69 (16–18) (2006) 1912–1922.
- [28] T. Randen, J.H. Husoy, Filtering for texture classification: a comparative study, *IEEE Trans. Pattern Anal. Mach. Intell.* 21 (4) (1999) 291–310.
- [29] J.G. Zhang, T.N. Tan, Brief review of invariant texture analysis methods, *Pattern Recognit.* 35 (3) (2002) 735–747.
- [30] W. Zhi-Zhong, Y. Junhai, Texture analysis and classification with linear regression model based on wavelet transform, *IEEE Trans. Image Process.* 17 (8) (2008) 1421–1430.
- [31] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *IEEE Trans. Syst., Man Cybern. SMC-3* (6) (1973) 610–621.
- [32] P. Brodatz, *Texture: A Photographic Album for Artists & Designers*, New York: Dover, New York, 1966.
- [33] S. Gai, G.W. Yang, S. Zhang, Multiscale texture classification using reduced quaternion wavelet transform, *AEU-Int. J. Electron. Commun.* 67 (3) (2013) 233–241.
- [34] G. Akbarizadeh, A new statistical-based kurtosis wavelet energy feature for texture recognition of SAR images, *IEEE Trans. Geosci. Remote Sens.* 50 (11) (2012) 4358–4368.
- [35] G. Van de Wouwer, P. Scheunders, D. Van Dyck, Statistical texture characterization from discrete wavelet representations, *IEEE Trans. Image Process.* 8 (4) (1999) 592–598.
- [36] J.D. Johnston, A filter family designed for use in quadrature mirror filter banks, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '80, 1980.
- [37] T. Randen, J.H. Husoy, Multichannel filtering for image texture segmentation, *Opt. Eng.* 33 (1994) 2617–2625.

ZhenMin Zhang received the B.S. degree from LanZhou University and the M.S. degree from XiaMen University, in China, in 2006 and 2010, respectively. She is currently pursuing the Ph.D. degree in the College of Photonic and Electronic Engineering, Fujian Normal University. Her current research interests include digital signal and image processing, wavelet theory and applications, and spiking neural networks.



QingXiang Wu received the Ph.D. degree in intelligent systems from the University of Ulster in the United Kingdom in 2005. He is a professor in the College of Photonic and Electronic Engineering, Fujian Normal University, in China. He has been a RCUK academic fellow in the School of Computing and Intelligent Systems in the University of Ulster in the UK. He has been a research fellow in School of Electronics, Electrical Engineering and Computer Science, Queen's University in the UK. He is a member of the IEEE and a member of International Rough Sets Society. He took part in designing of the IFOMIND autonomous robot control system which won the Machine Intelligence Prize of 2005 competition organised by the British Computer Society. His research interests include intelligent systems, robotics, spiking neural networks, visual image processing, machine learning, knowledge discovery, data mining, and intelligent internet of things. He has more than 90 publications on these topics in international journals, Chinese journals, and proceedings of international conferences.



ZhiQiang Zhuo received the B.S. degree from Anyang Institute of Technology, China, in 2011. He is currently a post graduate student in the College of Photonic and Electronic Engineering, Fujian Normal University, in China. His research interests are CUDA and spiking neural networks.





XiaoWei Wang received the B.S. degree from Tianjin University of Technology and Ed, China, in 2011. She is currently a post graduate student in the College of Photonic and Electronic Engineering, Fujian Normal University, in China. Her research interests are intelligent visual image processing algorithms.



Liuping Huang received the bachelor's degree in College of Photonic and Electronic Engineering from Fujian Normal University, China, in 2012. She is currently a post graduate student in the College of Photonic and Electronic Engineering, Fujian Normal University, in China. Her research interests are image processing and machine vision.