CNN-Based Hybrid-Order Texture Segregation as Early Vision Processing and Its Implementation on CNN-UM

Chin-Teng Lin, *Fellow, IEEE*, Chao-Hui Huang, *Student Member, IEEE*, and Shi-An Chen

Abstract—In this paper, a biologically inspired, CNN-based, multi-channel, texture boundary detection technique is presented. The proposed approach is similar to human vision system. The algorithm is simple and straightforward such that it can be implemented on the cellular neural networks (CNNs). CNN contains several important advantages, such as efficient real-time processing capability and feasible very large-scale integration (VLSI) implementation. The proposed algorithm also had been widely tested on synthetic texture images. Those texture images are randomly selected from the Brodatz textures database [1]. According to our simulation results, the boundaries of uniform textures can be detected quite successfully. For the nonuniform or nonregular textures, the results also indicate meaningful properties, and the properties also are consistent to the human visual sensation. The proposed algorithm also has been implemented on the CNN Universal Machine (CNN-UM), and yields similar results as the simulation on the PC. Based on the efficient performance of CNN-UM, the algorithm becomes very fast.

Index Terms—Cellular neural networks (CNNs), early vision system, Gabor filter, retinex model, texture segregation.

I. INTRODUCTION

RESEARCHERS in recent decades have elucidated signal transduction in the retina and the function of the visual cortex. The highly flexible nature of the neural circuits in the visual cortex, especially during the critical period (the early vision period) has been an interesting subject for studying and developing the neural plasticity.

Early vision, also known as the pre-attentive vision, represents a set of information processing mechanisms of the first stage of the visual processing. These mechanisms are operated in parallel across the visual field, and are believed to be used for detecting the most basic visual features, such as contrast, edge, contour, grouping, corner, junction, brightness and lightness computation for surface perception. On the other hand, Jain and Farrokhnia [2] proposed their research results as follows: the human vision system included two fundamental features: one is that the difference of light intensity that projects into the retina, the other is that the behavior of the retina is similar to the band-pass filtering 0. Those facts implicate that the abrupt changes on the intensity of light will bring the visual cortex stronger stimulation, and those abrupt changes become the so-called boundaries. Sometimes those boundaries have been called the first-order features. In general, the first-order features equal the average of the local areas. According to some investigations, the boundary detection algorithms based on this features are related to the zero crossing of the *Laplacian of Gaussian* (*LoG*) and, sometimes, the Canny's boundary detector [2]. There is one common issue in computer vision that it does not make a distinction between the contours of objects which are the actual primitives needed in the most application. It reveals that there are still some other mechanisms during the visual perception. Thus, this paper will focus on the texture segregation.

The segregation of the visual scenes (better known as boundary detection) is the fundamental process of the early vision. That is also an important and fundamental issue in image processing. Boundary detection can be applied in many areas, such as object tracking, stereo vision, pattern recognition, etc. Basically, boundary detection is used to separate the specific partition of the images into related regions, and thus the boundaries will be extracted. Or in other words, boundary detection is used to compare the difference between the spatial related local features of the images. The procedure of those processing is not an easy problem in the image processing area. However, it seems to be an intuitive ability in the human vision system, obviously.

The human visual system is able to effortlessly integrate the local features to construct our rich perception of patterns, despite the fact that the visual information is discretely sampled by the retina and the cortex. According to some biological and computational evidences, some kinds of data compression procedure occur at a very early stage in the human vision system [2]. Moreover, many physiological evidences imply that some form of this compression mechanism is involved in the mechanism locating the boundaries in the image [3]. In the early 1960s, there was a research result which implied that the majority of neurons in the primary visual cortex respond to a line or a boundary of a certain orientation in a given position of the visual field [2]. In 1981, Hubel and Wiesel found two types of orientation-selective neuron, one of that is sensitive to the areas of lines and boundaries, called simple cells, and the other is not, called complex cells [4]. The receptive field (RF) of the simple cells can be modeled by the Gabor function, and it has been widely used for information extraction, which is the so-called second-order feature [3], [5]. The mechanism of second-order feature extraction is more commonly known as the filter-rectify-filter cascade.

The authors are with the Department of Electrical and Control Engineering, National Chiao-Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: ctlin @mail.nctu.edu.tw; chhuang.ece90g@nctu.edu.tw; cuteless.ece87g@nctu.edu. tw).

This consists of the early linear filtering subunits, a nonlinearity (e.g., rectification), and a late linear filter.

At first, give a definition to texture. Unfortunately, in the author's knowledge, there are not any precise and identical definitions to texture. Although there doesn't have any best definition for texture, those features are so obvious that it can not be neglected. Some researchers have proposed different texture descriptions. The "definition" of texture is formulated depending on the particular application and there is no generally agreed upon definition. Some perceptually motivated examples have been given as follows.

- "We may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule [6]."
- "A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constant, slowly varying, or approximately periodic [7]."
- "The notion of texture appears to depend upon three ingredients: i) some local 'order' is repeated over a region which is large in comparison to the order's size; ii) the order consists in the nonrandom arrangement of elementary parts; iii) the parts are roughly uniform entities having approximately the same dimensions everywhere within the textured region [5]."

Bovic *et al.* (1990) give a very detailed analysis of the Gabor function which uses localized spatial filters for the texture feature extraction [8]. Bovic mentioned three supervised approaches to select the filter locations using the empirical information based on the power spectrum characteristics of the individual textures: first, for stronger oriented textures, the most significant spectral peak along the dominant orientation direction is used as a filter location. Second, pick the lower fundamental frequency to identify periodic textures. Third, for nonoriented textures, using the center frequencies of the two largest maximum is recommended. It is clear that an automated method is more attractive.

Dunn and Higgins developed a method for selecting optimal filter parameters [9]. This is a supervised approach that focused mainly on using the minimal number of filters. Only the specific filter, that separates two textures optimally, is used to partition an image. The optimal filter responds strongly to one class, and a lack of textural information of the other class may be also expressed. This kind of class is not identified to have a particular characteristic but lacking a characteristic of another. A more global solution to the problem is spreading filters throughout the frequency-domain field to capture salient information.

By providing the uniform coverage of the spatial-frequency domain with the Gabor filters, the problem of the selected central frequency is avoided. Jain and Farrokhnia [2] implemented the real Gabor filters for the texture segmentation using frequency bandwidth of one octave, center frequency spacing of one octave, angular bandwidth of $45°$, and angular spacing of $45°$ 0. The frequencies used for filters are as follows: $1\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, \ldots, (N_c/4)\sqrt{2}$ (cycles per image width).

For the textures with distinct spectral peaks which correspond to some global regularity, Tan proposed a useful method [10] to design the Gabor filters automatically. The central step in the algorithm is a spectral detection. It detects a global spectral peak, and repeatedly detects conspicuous peaks by erasing-operation on the spatial frequency plan. The power spectrum of a small neighborhood (e.g., 7×7) around the detected peak is set to zero. The iteration of the peak detection terminates when the ratio of the magnitude of the current peak to that of the first (e.g., the highest) peak is less than a pre-specified value (e.g., 80%).

In this paper, first, the related knowledge from physiology and psychophysics will be introduced. Next, proposed algorithm, so-called CNN-based hybrid-order texture segregation will be presented. Third, some experiments will be shown. Finally, the conclusions follow.

II. METHOD

A. Cellular Neural Networks

The cellular neural network (CNN), also known as the cellular nonlinear network, first introduced by Chua and Yang [11] as an able to implement alternative to fully connected neural networks, has evolved into a paradigm for those types of array [12]. The CNN paradigm provides the framework for the computation of an algorithmically programmable array computer on a chip: named the CNN Universal Machine (CNN-UM). Its powerfully computing characteristic enables the realization of complex image processing tasks. However, it is not necessary to construct complex analog-logical circuits, such as the CNN-UM for a special CNN application. Thus, this study aims at the property of the proposed algorithm to design a suitable CNN-based circuit.

1) CNN-Based Hybrid-Order Texture Segregation: In this paper, a new boundary detection algorithm is proposed. This algorithm combines the first and the second-order features for modeling the pre-attentive stage of human visual system. Fig. 1 shows the flow chart of the proposed approach: first, the first-order features have been extracted by the Gaussian low-pass filters and the second-order features have been extracted by the Gabor filters, respectively. Assume that each pixel of the output is defined by an $N + 1$ dimensional vector. After the first-order features extraction, the vector contains N Gabor filters and *1* Gaussian filter. Next, measure the difference of each pixel with its neighbor. Since the pixels, which belong to the same region have similar features, the level of difference among those pixels should be smaller than the difference to pixels existing in other regions of the image. Third, these pixels with values larger than a specific threshold are kept, and set the others to zero. Coarse boundaries which are look like Bell-shaped distribution would be got. Consequently, these boundaries can be thin by local peak detection. Finally, the boundaries which are similar to human visual system can be obtained. Followed sections are going to introduce each block of the proposed algorithm.

First-Order Feature Extraction: As the introduction in the previous section, the ganglion is accomplished by the so-called "center-surround" organization of the RF, in which it's excitatory and inhibitory subfields are organized into circularly symmetric regions. This fact implies the RF of the ganglion is similar to the *Difference of two Gaussians (DoG)*.
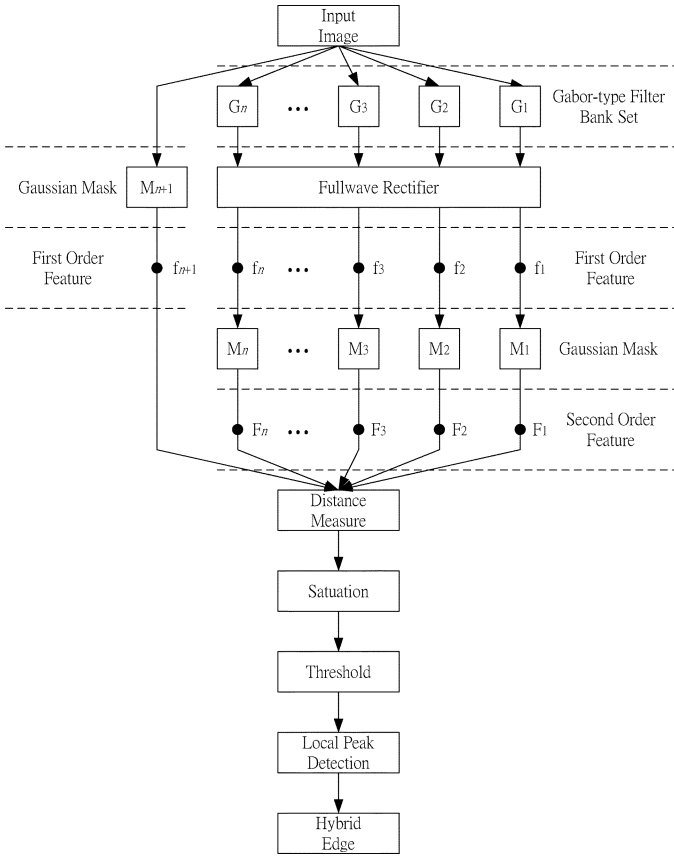
Fig. 1. Proposed algorithm. First-order features $f_1, f_2, f_3, \ldots, f_n$ are generated by a Full-wave Rectifier, and a set of Gabor-type filtering. The first-order feature f_{n+1} is generated by a Gaussian Mask. The second-order features $F_1, F_2, F_3, \ldots, F_n$ are generated by a set of Gaussian Mask. The following processors, including Distance Measure, Saturation, Threshold, Local Peak Detection, and Hybrid Edge, use the features $F_1, F_2, F_3, \ldots, F_n$, and f_{n+1} as the input and determine the boundaries as the description in the proposed algorithm.

Subsequently, this paper describes how the DoG function detects boundaries: first, two Gaussian filters with different values of σ are applied in parallel to the images. Afterwards the difference of the two smoothed instances is computed. It can be shown that the DoG operator approximates the LoG one which has been widely used in boundary detection.

The RF of the ganglion cells can be considered as the linear spatial weighting function. That is, the retinal ganglion cell can be modeled as a linearized function, where the RF implies where the weights are. Using the function to characterize the shape of the RF based on the DoG model, the output of the retinal ganglion cells can be calculated as follows:

$$O = \sum_{x,y} R(x,y)I(x,y) \tag{1}$$

where $I(\cdot)$ is the input image and $R(\cdot)$ represents the operation of the retinal ganglion cell. $R(\cdot)$ in (1) and (2) represents a biological structure named the RF of retinal ganglion cells. Unfortunately the exactly model of this biological structure is still unknown. Some physiological experiments indicated that the RF of the ganglion cells exhibits a center/surround charac-
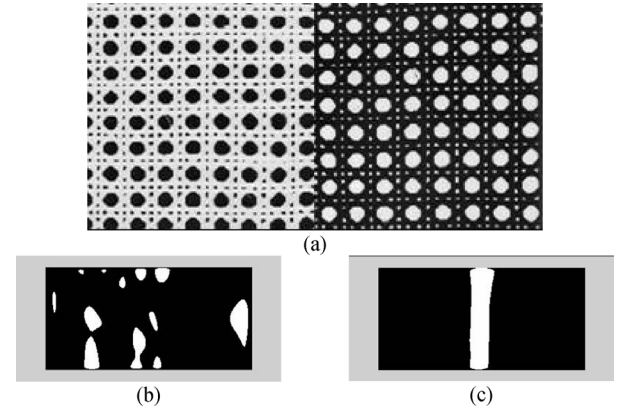

Fig. 2. Example which demonstrated the effect of first-order feature in boundary detection. (a) Input images which can be obtained from Brodatz texture database (D101-D102) [1]. (b) Boundary has been detected by second-order feature. (c) Boundary has been detected by first-order feature.

teristic. Furthermore, Thiem state that the RF of ganglions cells can be modeled as

$$R(x,y) \approx \Delta\left(g_\sigma(x,y)\right)$$

where $\Delta(\cdot)$ represents a Laplace operator and $g_\sigma(\cdot)$ is a Gaussian function. According to Hubel, under the optimum lighting condition, the central part of the RF is about 10 μm (4 csp) [5]. Thus, Thiem also recommends a standard deviation of $\sigma_G = (5\,\mu\text{m})/(\sqrt{2}) = \sqrt{2}$ (csp) [6].

For the whole of retinal ganglion cells with identical RFs, the output of each cell in the array can be obtained as follows:

$$O(x_0, y_0) = \sum_{x,y} R(x - x_0, y - y_0)I(x,y) \tag{2}$$

where $O(x_0, y_0)$ is one of the output of the retinal ganglion cells, whose RF is centered at position x_0 and y_0.

The operation of DoG function can be divided into two stages: the Gaussian convolution and the gradient operation. The Gaussian convolution is like extracting the mean of the local region which is the so-called first-order feature, and the gradient operation as a measure for the variation of the first-order feature.

For the sake of combining of the first-order and the second-order features, first, use the Gaussian convolution for extracting the first-order features. And then, the gradient processing will be applied after combining the second-order features. Fig. 2 illustrates the coarse boundary between two patterns based on first-order features.

There is more than one type of features that are mixed, and only the first- and second-order features are insufficient. Fig. 2 demonstrates the situation for the patterns (D101-D102 from Brodatz textures [1]) with hybrid-order feature. Note that the first-order feature is dominant.

For convenience, a general low-pass operation has been used to emulate the Gaussian function in the CNN. The template can be defined as follows [13]:

$$A = \begin{bmatrix} 0.1 & 0.15 & 0.1 \\ 0.15 & 0 & 0.15 \\ 0.1 & 0.15 & 0.1 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, z = 0. \tag{3}$$

Fig. 3.   Example of 2-D Gabor type filtering. (a) Standard 2-D Gabor type filter in time domain. (b) In frequency domain.

Second-Order Feature Extraction: As the description in the previous section, the RFs of V1 cells are orientation selective, and it can be modeled by the Gabor function [3], [5]. The Gabor function is an adaptive bandpass filtering method which constructs a complete but nonorthogonal basis set. On the other hand, the Gabor function consists of a Gaussian function which is modulated by a sinusoidal function and can be described as follows [14]:

$$g_{K,S}(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \quad (4)$$

and the Fourier transform of (4) is as follows:

$$G_{K,S}(u,v) = \exp\left[-\frac{1}{2}\left(\frac{(u-W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right] \quad (5)$$

where $\sigma_u = 1/(2\pi\sigma_x)$ and $\sigma_v = 1/(2\pi\sigma_y)$.

An example of the Gabor function has been presented in Fig. 3, where Fig. 3(a) is the impulse response of a standard Gabor filtering in time domain, and Fig. 3(b) is its frequency-domain representation. Expanding the signal using this basis provides a localized frequency description.

In practice, the Gabor function can be divided into real (even) part and imaginary (odd) part as follows:

$$g_c(x,y) = h(x,y)\cos(2\pi Fx') \quad (6)$$
$$g_s(x,y) = h(x,y)\sin(2\pi Fx') \quad (7)$$

where $x' = x\cos\phi + y\sin\phi$, and $h(x,y)$ is a Gaussian function.

Clearly, (6) and (7) are very similar to the each other. In fact, (7) is just a phase shifted version of (6) Both of them can be used to extract the local features of the spatial domain. Based on some of the psychophysical theories, Malik and Perona provide justifications which indicate that only even-symmetric filters can be used. In this paper, real-valued, even-symmetric Gabor filters are used.

For implementation of the 2-D Gabor-type filtering, according to Shi [15], the template in the CNN for the 2-D Gabor-type filtering can be represented as follows [15]:

$$A = \begin{bmatrix} 0 & e^{-jw_{y0}} & 0 \\ e^{jw_{x0}} & -(4+\lambda^2) & e^{-jw_{x0}} \\ 0 & e^{jw_{y0}} & 0 \end{bmatrix}$$
$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$z = 0. \quad (8)$$

Shi's CNN-based Gabor-type filtering provides good results. However, the complex template becomes a problem in the implementation on the CNN-UM. Thus, a linear-region CNN-based Gabor-type filtering is proposed. The proposed algorithm can be separated into a two steps CNN processing . These can be presented as follows.

First Step:

$$A_1 = \begin{bmatrix} a & b & c \\ d & -(2+\lambda^2) & d \\ c & b & a \end{bmatrix}, B_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, I_1 = 0 \quad (9)$$

Second Step:

$$A_2 = \begin{bmatrix} c & d & a \\ b & (2+\lambda^2) & b \\ a & d & c \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, I_2 = 0 \quad (10)$$

where a, b, c, and d is used to determine the orientation of the Gabor-type filtering. Here λ can be used to determine the bandwidth of the filtering. For example, Fig. 4 presents the frequency response of the proposed CNN-based Gabor-type filtering with the following parameters: $a = 0$, $b = 1$, $c = 0$, $d = 0$, and $\lambda = 0.9$. The frequency response is given by the following.
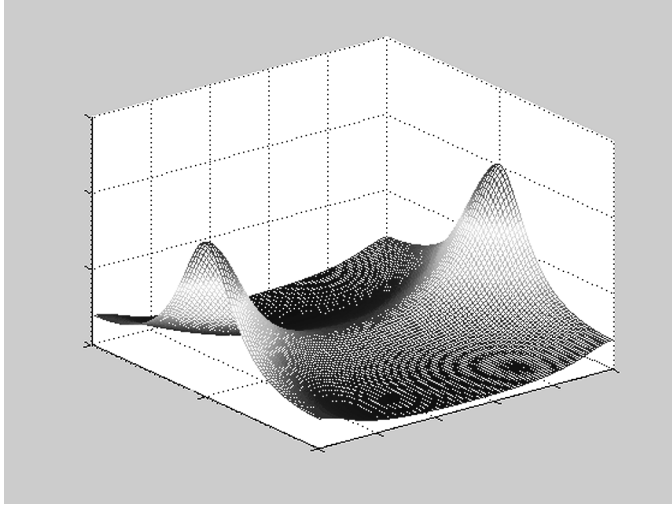
Fig. 4. Example of proposed Gabor-type filtering.

First Step:

$$H_1\left(e^{j\omega_x}, e^{j\omega_y}\right) = \frac{\lambda^2}{2 + \lambda^2 - 2\cos\left(\omega_y\right)}, \text{ for all } \omega_x \quad (11)$$

Second Step:

$$H_2\left(e^{j\omega_x}, e^{j\omega_y}\right) = \frac{\lambda^2}{-2 - \lambda^2 - 2\cos\left(\omega_x\right)}, \text{ for all } \omega_y. \quad (12)$$

Thus, the frequency response of the proposed CNN-based Gabor-type filtering can be represented as shown in (13) at the bottom of the page. An example of proposed Gabor-type filtering is shown in Fig. 4

In fact, the proposed linear-region, CNN-based Gabor-type filtering is not able to adjust the frequency-related parameters. However, it can be implemented by the CNN-UM. Consequently, the entire proposed approach can be realized.

Gabor Filtering Bank Set: Besides the orientation selectivity, the Gabor filters are also frequency selective. With these two properties, Daugman extended the original Gabor filter to a two-dimensional (2-D) representation [16]. There are many researches which focus on the Gabor filter bank. Jain and Farrokhnia (1981) suggested a bank of Gabor filters, i.e., Gaussian shaped bandpass filters, with dyadic coverage of the radial spatial frequency range and multiple orientations 0. Fig. 5 shows an example of the Gabor filtering bank set.

Gabor has first recognized and introduced a time-frequency version of Heisenberg's inequality as follows:

$$\sigma_t \sigma_f \geq \frac{1}{4\pi} \quad (14)$$

where σ_t and σ_f are the time and frequency standard deviations, respectively. The Gabor filter is just the modulation of the Gaussian function. For the Gabor function it has been proven that the action only causes a shift in the frequency domain, and it wouldn't affect the resolution of the Gaussian function in both the spatial and the frequency domain. It means that the Gabor function inherit the properties of the Gaussian possessing optimal resolution in both domains, and those properties imply the Gabor filter is suitable for the texture segregation.

Since the goal of this paper is designing an algorithm which can be implemented on the CNN-UM, the structure could not be complex-valued. In this paper, four Gabor filters are used for extracting the second-order features in the experiments. All of these Gabor filters have the same Gaussian shape in the frequency domain and scatter uniformly in four orientations.

Full-Wave Rectification: Just like the other filter-rectify-filter model, the rectifying operation is taken after the operation of the Gabor filters. It has been generally acknowledged that V1 cells have a property which looks like the half wave rectification. The intervening rectification ensures that the fine-grain positive and negative portions of the carrier will not disable another when the smoothing operation is performed.

Fig. 6(b) demonstrates the output of the Gabor filtering without rectification, and Fig. 6(c) is the result of the same operation but with rectification. The white pixels in the image represent the pixels that the matching features have been detected by the Gabor function. This result is similar to the behavior of the V1 cells. Because of the restriction of display, there are some pixels with negative response which do not appear in Fig. 6(b). In Fig. 6(c), the boundary of the regions is more apparently, and that is because of the rectification turning the negative response to positive.

For the implementation of the full-wave rectification by the CNN, the template can be defined as follows:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 0 \end{bmatrix}, z = 0 \quad (15)$$

where

$$b = \begin{cases} 1, & \text{if } v_{u_{ij}} \geq 0 \\ -1, & \text{otherwise.} \end{cases} \quad (16)$$

Gaussian Post Filtering: After the cells were stimulated by a specific signal, for example, a bar with specific orientations, the

$$H\left(e^{j\omega_x}, e^{j\omega_y}\right) = H_1\left(e^{j\omega_x}, e^{j\omega_y}\right) H_2\left(e^{j\omega_x}, e^{j\omega_y}\right)$$

$$= \frac{\lambda^4}{-\lambda^4 - 4\lambda^2 - 2\lambda^2\cos\left(\omega_x\right) + 6\cos\left(\omega_y\right) - 4\cos\left(\omega_x\right) + 4\cos\left(\omega_x\right)\cos\left(\omega_y\right) - 4} \quad (13)$$
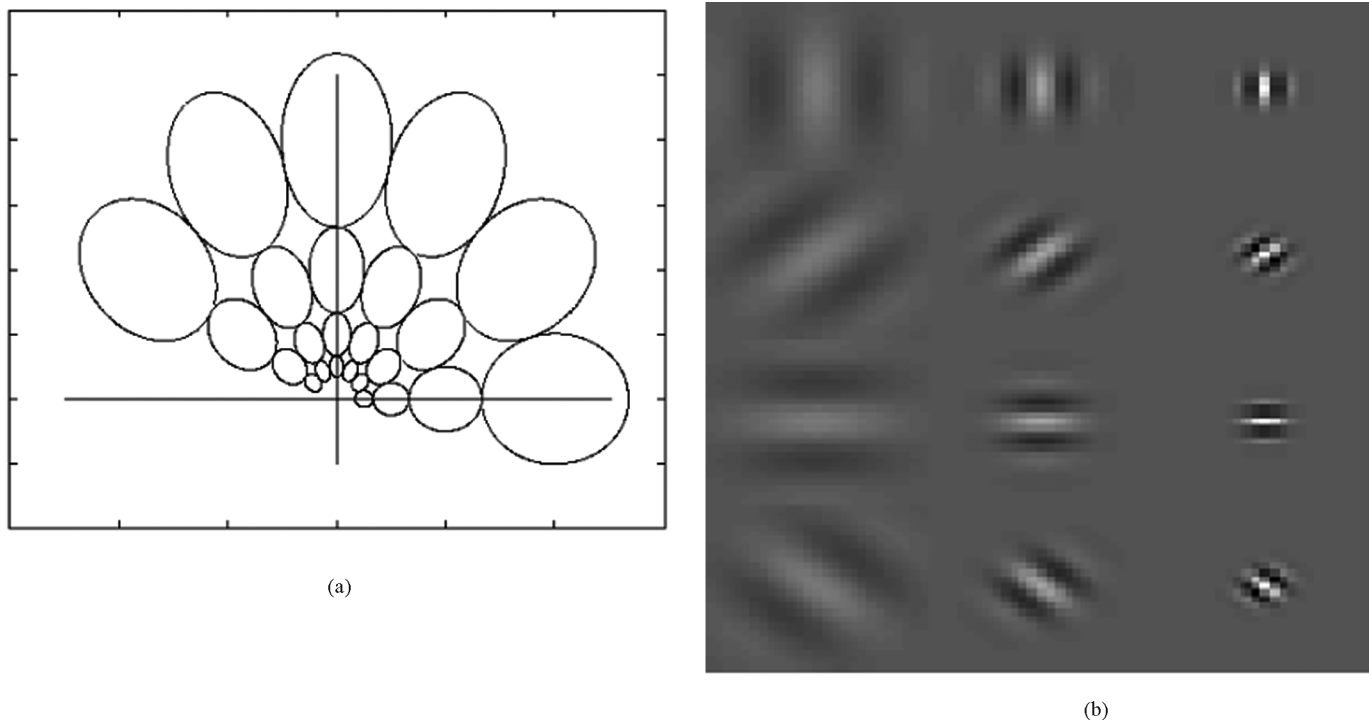
(a)

(b)

Fig. 5. Example of Gabor filter dictionary. (a) Gabor-type filter bank set, and (b) Feature space of Gabor filter dictionary.
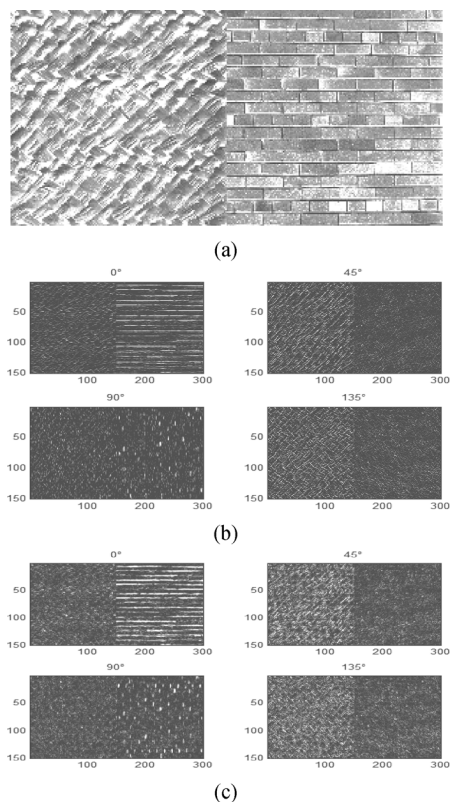

(a)

(b)

(c)

Fig. 6. Example demonstrates the effect of rectifying. (a) Input. (b) Output without rectifying. (c) Output with rectifying.

(a)

(b)

(c)

Fig. 7. (a) Input. (b) Output before rectification. (c) Output after rectification.

output of the V1 cells responding to same direction will aggregate together. The region of cells which contain the same propertie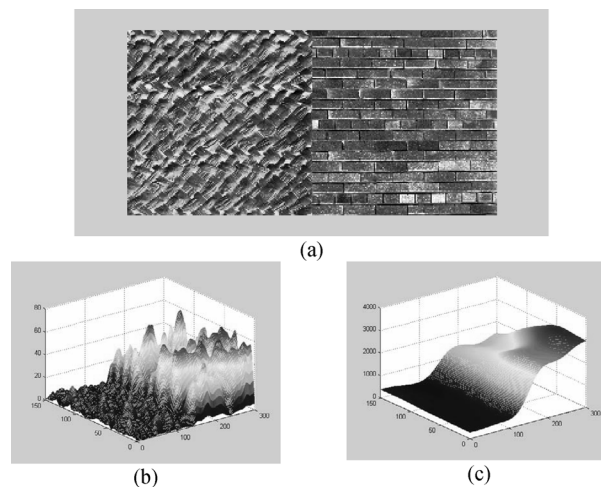s will respond stronger than the other regions. It is consistent with the "localization" properties of the textures. This effect can be simulated by a Gaussian post filters. It looks like the averaging with different weighting which is inverse proportional to the distance from the center of the post filter.

Fig. 7(b) shows the result after the processing of rectification, and Fig. 7(c) is the result that Fig. 7(b) has been processed by the Gaussian post filtering. Note that in Fig. 7(c) there is a ramp-like feature profile. The next step is for detecting the position where the variation of difference is maximal.

For the implementation of the Gaussian post function in the CNN, the template can be defined as the same as in (3) [13].

Difference Measure: The features which have been extracted by Gaussian post filtering can be described by an N-dimen-

sional vector. Each feature vector can be regard as a point in N-dimensional space. According to Chen [17], the difference is represented by the distance in N-dimensional space [17].

There is an important property of the textures. That is the pixels which are aggregated together usually contain similar features. Based on this property, some algorithms use gradient for extracting the features [10]. In this paper, only the difference between the features of each pixel to pixels right behind and below to it has been calculated.

Saturation: In the proposed algorithm, when there are more than two kinds of textures in the test patterns, there will be more than one boundary. Because these boundaries usually do not have similar intensity, choosing threshold becomes an important problem.

For the sake of finding the threshold, the mean of the difference of the total pixels is used as the threshold. Usually, some boundaries with relative lower magnitude are eliminated. This is because first, a relative huge region be considered for measuring the local feature. Next, the scale of difference between different patterns varies enormously. Obvious boundaries and cause relatively larger difference and raise the mean of difference. The boundaries which are not so obvious causing relative lower difference will be eliminated.

For attacking this problem, a natural-log transform can be used for simulating the saturation effect. It can suppress stronger responses which may affect the threshold too much. Meanwhile, it still keeps the location of maximum difference where the boundaries lying.

The strength of the responses reflects the level of differences between two local regions, but it may be not so linearly consistent to our perceptual feeling. According to some biological theories, in the human vision system, the dynamic range of response is limited, and the range of response is not linearly proportional to stimulate [18]. Natural log transform is an ordinary and important operation and it stretch the range of lower responses where it need to judge whether there are boundaries or not.

Local Maximum Detection: The coarse boundaries detected after taking threshold generates a range where the boundaries are probably located on. Thus, local maximum detection is used to detect the best assumption of the location of the boundaries. It assumes that the difference among different patterns should be maximal at their boundary.

Local maximum detection can be implemented by the CNN quiet easily. Fig. 8 is an example which demonstrates the peak detection in the algorithm. Fig. 8(a) is an input image, and Fig. 8(b) is the detected coarse boundary. Fig. 8(c) is the 3-D version of Fig. 8(b), and in this figure the vertical axis is intensity. Fig. 8(d) is the result of Fig. 8(c) by taking peak detection. Fig. 8(e) is the superposition of Fig. 8(a) and (c). From Fig. 8(e) the detected boundaries have high accuracy which is consistent to our assumption can be observed.

For the implement of the local maximum detection in CNN, the template can be defined as follows:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} b & b & b \\ b & 0 & b \\ b & b & b \end{bmatrix}, z = 3.5 \quad (17)$$
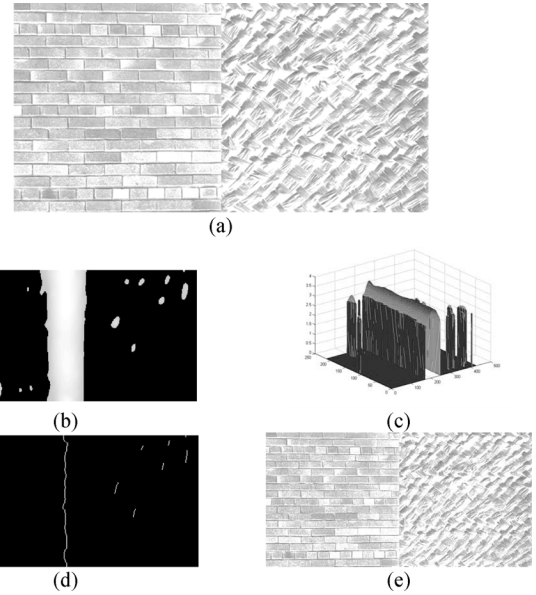

Fig. 8. (a) Input image. (b) Coarse boundary. (c) 3-D version of (b). (d) Results after peak detection.

where

$$b = \begin{cases} 0.5, & \text{if } v_{u_{ij}} - v_{u_{kl}} \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

Equation (17) and (18) performs what the algorithm needs. However, that can not be implemented in CNN-UM because of nonlinear features. Thus, some other solutions for CNN-UM have been required. In this paper, BLACK_AND_WHITE_SKELETONIZATION has been suggested for this operation if the algorithm needs to implement this operation in CNN-UM [19].

III. EXPERIMENTS

In this paper, the images which consist of a number of different test patterns have been applied to the algorithm. Most of those test patterns are synthesized by textures from Brodatz album [1], and it also has become a standard for evaluating texture algorithms. Each texture pattern that has been used here are $640 * 640$ pixel 8-bit gray-scale images, respectively.

In the experiments, proposed algorithm has been simulated both on MATLAB and CNN-UM. Followed section will discuss the results, respectively.

A. Simulations

In our experiments, first, implement proposed algorithm in a computer. There are some parameters need to be selected:

The number of Gabor filters and their parameters (U, V, σ) which decide the shape and orientation of the Gabor filters in the frequency domain. The Gabor filtering is computation intensive, and increasing the number of the Gabor filters will increase computation loading dramatically. On the other hand, unnecessary and useless feature extracted by the wrong-designed Gabor filters may cause wrong boundaries.
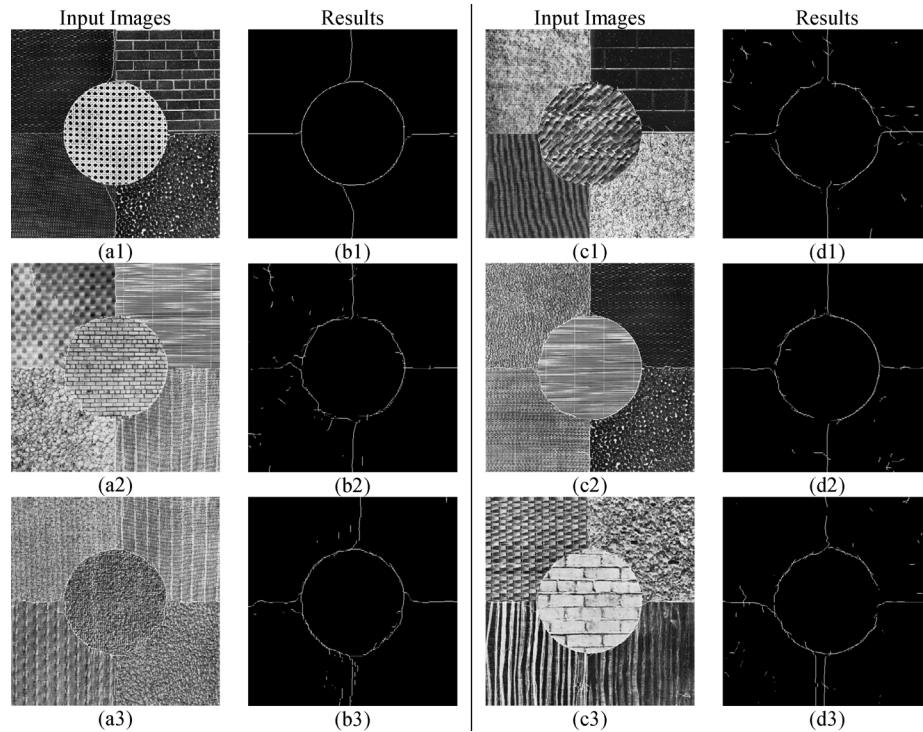
Fig. 9. Simulation results of proposed algorithm.

The σ_g of the post Gaussian filter, which decides the smoothing level. Increasing σ can eliminate more noise, but the accuracy of the boundary may decrease. Because both the Gabor filters and the Gaussian filters have spatial information, the values of σ_g must cooperate with σ to obtain a better result.

Designing above parameters is an important but sophisticated issue. Designing center frequencies of the Gabor filters is discussed in filter-design approaches. There is including the unsupervised methods; such as the algorithm proposed by Jain, and supervised methods; such as algorithm proposed by Dunn [20]. The algorithm in this paper is a kind of unsupervised method. That means all of the information of input patterns is unknown. Nevertheless, the emphasis of this paper is not on optimizing the design of the Gabor filters, but rather proposing a simple algorithm modeling early vision and being able to implement on CNN. Parameters as follows are empirically chosen and they are all the same in the following experiments without indicating specifically:

> pattern size(Brodatz texture): $640 * 640$ pixels;
> orientation ϕ: $0°$, $45°$, $90°$, $135°$;
> center frequency F: 1/32, 1/16, 3/32, 1/8 cycles/pixel;
> σ of Gabor filter: 16 pixels;
> down sampling rate M: 3;
> σ_g of post Gaussian filter: 25 pixels;
> mask sizes of Gabor and Gaussian: 3σ, $3\sigma_g$.

The simulation results have been shown in Fig. 9. In the results, most cases of texture boundaries are able to be extracted. Only few cases are not. On the other hand, in the cases which boundaries are not able to be extracted, the results also indicate meaningful properties which consistent to human visual sensation.
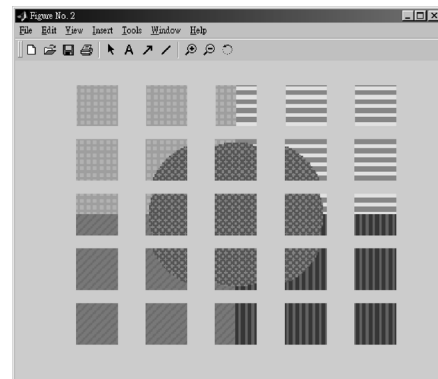

Fig. 10. Example of dividing of the input images.

B. Implementation on CNN-UM

In the experiments, proposed algorithm has been implemented on ACE4K Chip. Because of the limitation of current technology, the size of cell array has been limited. That becomes to the major problem in the implementation of proposed algorithm on CNN-UM. For example, the chip that has been used for implementing proposed algorithm contains cell array 64 by 64. It is too small to analyze the texture. For the sake of higher resolution for obtaining better performance, some necessary operations have been performed before proposed approach: first, each input image has been divided into several sub-images. Next, process each sub-image, respectively, with same parameters. Fig. 10 shows how to divide the input images. Note that there are overlapped areas between sub-images because the boundary effects of the cell array have to be avoided.
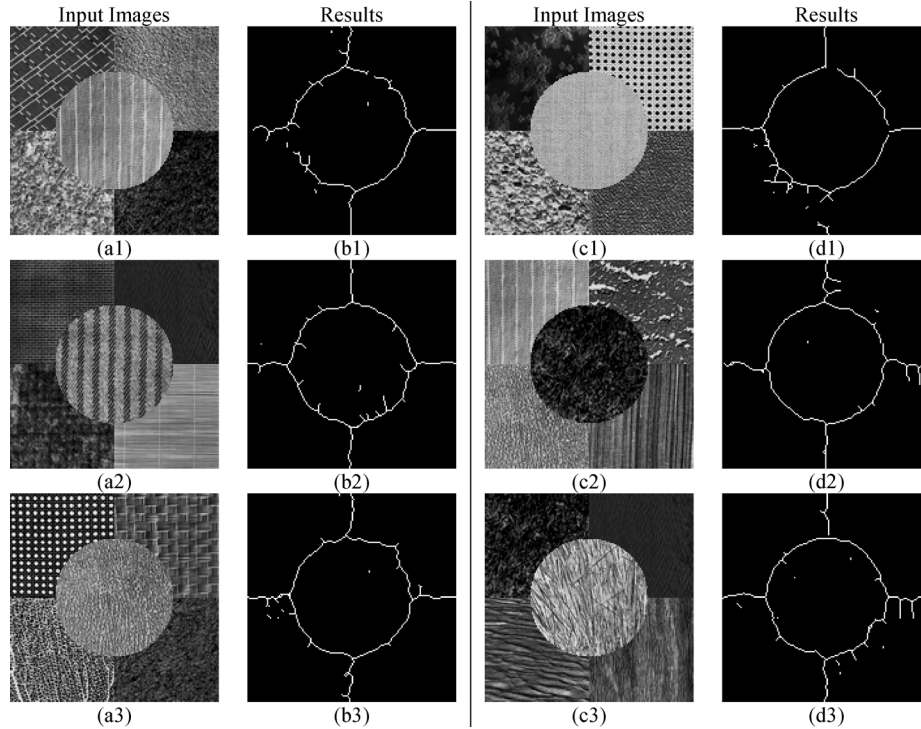
Fig. 11. Implementation results of proposed algorithm.

Another issue in the implementation of proposed algorithm on CNN-UM is that some nonlinear templates are not able to be implemented on the chip. Thus, several proposed operation is not able to be implemented directly. For those operations, another operation which is able to be implemented on the chip is chosen and the result is similar to the results of the simulations. Even more, some operations in the proposed algorithm need to be disabled and take the trade-off. For example, local maximum detection operation contains nonlinear template and thus, it can not be implemented on ACK4K Chip. Hence, BLACK_AND_WHITE_SKELETONIZATION is chosen to replace it. Fig. 11 shows the results of implementation of proposed algorithm on ACE4K Chip.

C. Discussions

In this section, some properties of the proposed approach is discussed. The following describes how to estimate the error.

Only the case that synthesizes two texture patterns in Brodatz texture is considered. In the algorithm, every boundary is independently. It is hard to judge the accuracy if multi-boundaries simultaneously is considered. Especially when some boundaries are detected and the others are not.

The distance between the answer and the result detected by the algorithm is measured in the condition of boundary which is detectable. The error by dividing measured distance into the number of total pixels is defined.

For simplicity, 70 textures of Brodatz textures, which are generally consistent to our definition of textures, are picked. Each test image is synthesized by choosing two textures from the 70 textures randomly, 500 of combinations have been tested. Some examples have been shown in Fig. 12.
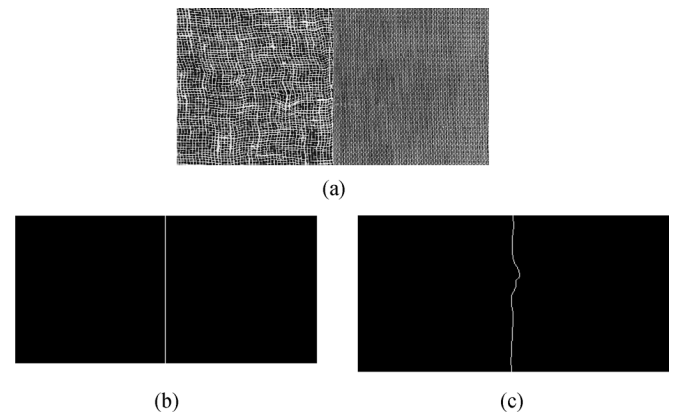

Fig. 12. Example of error estimation. (a) Input. (b) Answer (middle line). (c) Output.

Fig. 13 is a histogram of error estimation in our experiment, and the results with error less than 5% is account for 85% for test images. The average of the error is less than 5%. The smallest error is 0.76%. Note that the images with bigger estimated errors are reasonable. This kind of examples has been shown in Fig. 14. In these examples, the boundaries between different textures (middle line) exist, but they are weaker than local boundaries caused by nonuniform regions. For the sake of simplicity, only the largest peaks are kept during error estimation, so the boundaries in the middle are not kept in the results. Although in these examples, the outputs are consistent to human visual perception. Their errors are quite big. It is hard to define a generally "correct answer" for all test images in human vision system, and the method that used to measure the error is probably not suited for those kind of test images. For this reason, the measurement
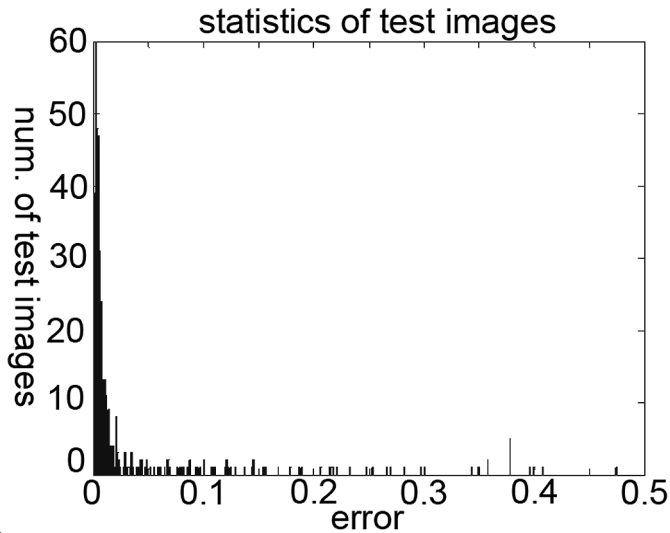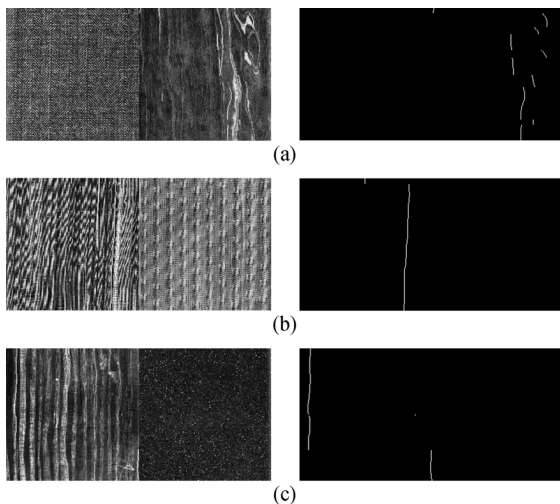
Fig. 13.　Histogram of error estimation.

Fig. 14.　Three examples of test image with big estimation errors.

is not necessary for the input images synthesized by the rest 42 textures in Brodatz textures.

IV. CONCLUSION

In this paper, a simple framework for hybrid-order boundary detection is proposed. It mimics the mechanism of the early stage of the human vision, and experimental results are generally consistent to the human visual sensation. After post processing, the detected boundaries also have adequate accuracy for the other image processing applications such as stereo, and pattern recognition. By implementing the proposed algorithm on the CNNs, the computational time will greatly decrease. The real-time processing capability is critical in some applications, such as the object tracking.

Although the proposed algorithm is widely tested to detect boundaries of synthetic textures successfully, there are still some problems demanding to be overcome:

As same as the other algorithms for textures analysis which is also based on the Gabor filters, there are too many parameters need to be determined. Determining the parameters will much more complex when the synthesized texture patterns increased.

For the sake of keeping the structure simple and combining the hybrid-order features easily without the clustering methods, same resolution for all of the Gabor filters is used in the approach.

In this approach, only the first- and the second-order features are considered. According to some research results, there are still some higher order features that can be utilized. For example, color is one of them. The proposed approach can be extended to color textures by integrating color information.

REFERENCES

[1] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*.　New York: Dover, 1966.
[2] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern. Recogn.*, vol. 24, pp. 1167–1186, 1991.
[3] I. Chen, "Visual and visuomotor processes," in *Vision and Cognition*. Taipei, Taiwan, R.O.C,: Yuan-Liou, 1999, pp. 128–158.
[4] D. H. Hubel and T. N. Wiesel, "Sequence regularity and geometry of orientation columns in themonkey striate cortex," *J. Comput. Neurol.*, vol. 158, pp. 267–293, 1974.
[5] J. K. Hawkins, "Picture processing and psychopictorics," in *Textural Properties for Pattern Recognition*, B. Lipkin and A. Rosenfeld, Eds. New York: Academic, 1969.
[6] H. Tamura, S. Mori, and Y. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 460–473, 1978.
[7] J. Sklansky, "Image segmentation and feature extraction," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 237–247, 1978.
[8] A. C. Bovik, "Analysis of multichannel narrowband fiters for image texture segmentation," *IEEE Trans. Signal Process.*, vol. 39, no. 12, pp. 2025–2043, Dec. 1990.
[9] D. Dunn and W. E. Higgins, "Optimal Gabor filter for texture segmentation," *IEEE Trans. Image Process.*, vol. 4, no. 7, pp. 947–964, Jul. 1995.
[10] T. N. Tan, "Texture edge detection by modelling visual cortical channels," *Pattern. Recogn.*, vol. 28, pp. 1283–1298, 1995.
[11] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. 35, no. 12, pp. 1257–1272, Dec. 1988.
[12] L. O. Chua and T. Roska, "The CNN paradigm," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, no. 2, pp. 147–156, Feb. 1993.
[13] *ALADDIN CNN Software Library (Templates and Algorithms)*.　Budapest, Hungary: Analogic Computers Ltd., 2000.
[14] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429–457, 1946.
[15] B. E. Shi, "Gabor-type filtering in space and time with cellular neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 45, no. 2, pp. 121–132, Feb. 1998.
[16] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimised by two-dimensional visual cortical filters," *J. Opt. Soc. Amer.*, vol. 2, pp. 1160–1169, 1985.
[17] I. Chen, *Texture Perception: A Linear System Approach*.　Berkeley: University of California, 1994.
[18] A. Treisman, "Features and objects in visual processing," *Sci. Amer.*, vol. 254, pp. 114–125, 1986.
[19] *ALADDIN CNN Software Library for ACE4K Chip (Templates and Algorithms)*.　Budapest, Hungary: Analogic Computers Ltd., 2000.
[20] T. P. Weldon and W. E. Higgins, "Designing multiple Gabor filters for multi-texture image segmentation," *Opt. Eng.*, vol. 35, pp. 2852–2863, 1996.

Chin-Teng Lin (S'88–M'91–SM'99–F'05) received B.S. degree in control engineering from National Chiao-Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 1996 and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, Lafayette, IN, in 1989 and 1992, respectively.

Since August 1992, he has been with the College of Electrical Engineering and Computer Science, NCTU, where he is currently the Provost of Academic Affairs and the Chair Professor of Electrical and Control Engineering. He served as the Founding Dean of Computer Science College of NCTU from 2005 to 2007. He is the textbook author of *Neural Fuzzy Systems* (Prentice-Hall, 1996) and *Neural Fuzzy Control Systems with Structure and Parameter Learning* (World Scientific, 1994). He has published over 110 journal papers, including about 80 IEEE transaction papers. His research interests are in intelligent technology, soft computing, brain-computer-interface, intelligent transportation systems, robotics and intelligent sensing, and nano-bio-information technologies (NBIC) and cognitive science.

Dr. Lin was promoted to IEEE Fellow in 2005 for contributions to biologically inspired information systems. He is a member of the Board of Governors (BoG) of the IEEE Systems, Man, Cybernetics Society (SMCS) from 2003 to 2005, and the current BoG member of IEEE Circuits and Systems Society (CASS). He is the IEEE Distinguished Lecturer from 2003 to 2005. He currently serves as the Deputy Editor-in-Cheif of the IEEE TRANSACTIONS OF CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS. He was the Program Chair of the 2006 IEEE International Conference on Systems, Man, Cybernetics held in Taipei, Taiwan, R.O.C. Dr. Lin is a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi. Dr. Lin was the president of Board of Government (BoG) of Asia Pacific Neural Networks Assembly (APNNA) from 2004 to 2005. He has won the Outstanding Research Award granted by National Science Council (NSC), Taiwan, since 1997 to present, the Outstanding Professor Award granted by the Chinese Institute of Engineering (CIE) in 2000, and the 2002 Taiwan Outstanding Information-Technology Expert Award. He was also elected to be one of 38th Ten Outstanding Rising Stars in Taiwan (2000).

Chao-Hui Huang (S'05) received the M.S. degree in computer science and information engineering from Chung-Hua University (CHU), Hsinchu, Taiwan, R.O.C., and the Ph.D. degree in electrical and control engineering from National Chiao-Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 1999 and 2007, respectively.

His research interests are in the areas of artificial intelligence, soft computing, and bio-inspired information system.

Shi-An Chen received the M.S. degree in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 2004, where he is currently working toward the Ph.D. degree.

His primary research interests are analog integrated circuitdesign, signal processing, and application of neural networks. Currently, his research focuses on spatial and frequency properties of EEG and dual tasks analysis.