Journal of Electronic Imaging 6(3), 329–336 (July 1997).

Wavelet-based feature-adaptive adaptive resonance theory neural network for texture identification

Jian Wang
Golshah Naghdy
Philip Ogunbona
University of Wollongong
Department of Electrical and Computer Engineering
Northfield Avenue
Wollongong, New South Wales 2522, Australia
E-mail: jiwa@elec.uow.edu.au

Abstract. *A new method of texture classification comprising two processing stages, namely a low-level evolutionary feature extraction based on Gabor wavelets and a high-level neural network based pattern recognition, is proposed. The design of these stages is motivated by the processes involved in the human visual system: low-level receptors responsible for early vision processing and the high-level cognition. Gabor wavelets are used as extractors of "low-level" features that feed the feature-adaptive adaptive resonance theory (ART) neural network acting as a high-level "cognitive system." The novelty of the model developed in this paper lies in the use of a self-organizing input layer to the fuzzy ART. Evaluation of the model is performed by using natural textures, and results obtained show that the developed model is capable of performing the texture recognition task effectively. Applications of the developed model include the study of artificial vision systems motivated by the human visual system model.* [S1017-9909(97)00403-0]

1 Introduction

Texture analysis plays an important role in vision research and is an important feature of object recognition and classification. There have been several attempts to develop systems based on the multi-channel model of the human visual system (HVS).[1–5] Such attempts have been facilitated by the notion of Gabor functions having properties similar to the receptive field profile of the simple cells found in the visual cortex of cats.[6,7] More recently, Gabor wavelets have been proven to be powerful tools in texture analysis.[1,4,5]

Most of the works in HVS-based vision research, however, are limited to the investigation of the properties of the low-level cells and are concentrated at the feature extraction level.[1–4] In this paper, a new neural network model, feature-adaptive adaptive resonance theory (ART), based on the fuzzy ART network, is proposed. The conventional fuzzy ART is modified into a self-organizing network not only at the output layer and the weights but also at the input layer. This is to ensure a self-expanding evolutionary model of an artificial vision system motivated by the human visual system.

The proposed scheme starts by learning simple frequency components. As the test environment becomes more complex, the system augments and refines itself by employing more feature extractors at the input layer. This arrangement ensures a robust functioning of the proposed system in a random environment and reduces the computational burden in real-time applications.

The proposed system has been tested using up to 50 natural textures. Results obtained show that the Gabor wavelets/feature-adaptive ART system can be used for texture classification.

The organization of the rest of the paper is as follows: Section 2 reviews the fuzzy ART system; Section 3 describes the proposed system; the results of simulation and a discussion are presented in Section 4; and a summary of the work is given in Section 5.

2 Fuzzy Adaptive Resonance Theory

The adaptive resonance theory (ART) network was first developed by Carpenter and Grossberg in the 1980s.[8] The initial network, ART1, is capable of processing binary inputs only, but subsequent networks, ART2 and ART3, added capabilities of handling continuous valued inputs.[8,9] Fuzzy ART, which has similar architecture as ART1, can process continuous-valued data like ART2 and ART3.[10] Furthermore, fuzzy ART has a fast and stable learning procedure[10] and is employed widely in object recognition tasks.[11,12] More recently, fuzzy ART has been further developed into ARTMAP and ART-EMAP networks for object recognition.[13,14] In the next section, ART1, which is the fundamental of the ART system, is reviewed.

2.1 *ART1*

The ART1 system, as described by Carpenter and Grossberg,[8] consists of two elements, attentional and ori-

Paper NNT-04 received Jan. 20, 1997; revised manuscript received Mar. 3, 1997; accepted for publication Mar. 5, 1997.
1017-9909/97/$10.00 © 1997 SPIE and IS&T.

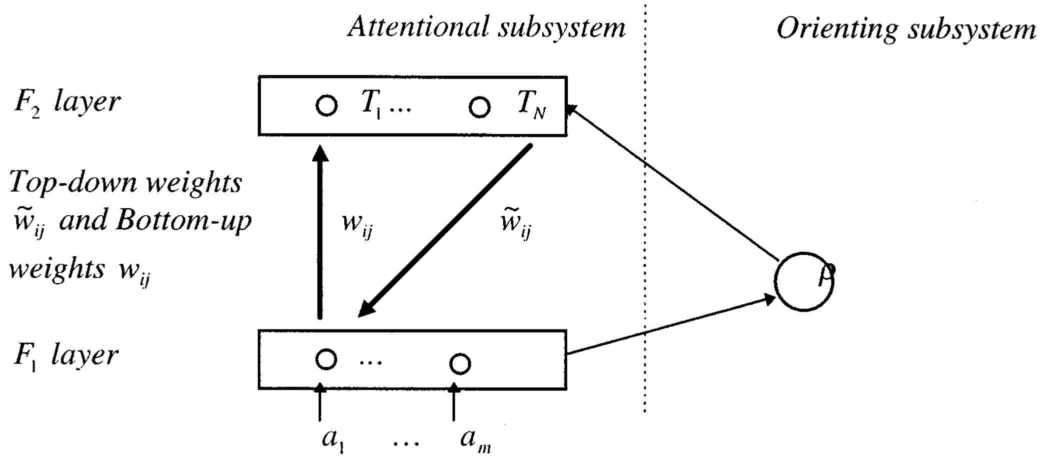


Fig. 1 ART1 system.

enting subsystems (see Figure 1). The attentional subsystem has two layers, F_1 and F_2, where F_1 is the input layer and F_2 is the output layer. The nodes in the different layers are fully connected as every node in the F_1 or F_2 layer is connected to all the nodes in the other layer. The pathway (weight) from the i'th node in the F_1 layer to the j'th node in the F_2 layer is w_{ij} and is called bottom-up weight. The weight from the i'th node in the F_2 layer to the j'th node in the F_1 layer is $\widetilde{w}_{ij}$, which is one element in the top-down vector. The vigilance parameter, ρ in the orienting subsystem decides whether the system goes into the resonance state or reset state. Its function will be introduced later.

For a given binary input data, say $I=(a_1,\ldots,a_m)$ (that is, bottom-up vector), passing through the bottom-up pathway (w_{ij}), an input for the F_2 layer of the form given in Eq. (1) is generated,

$$T_j(I)=\sum_{i=1}^{M} a_i w_{ij} \tag{1}$$

where $T_j(I)$ is the input of the j'th node in the F_2 layer. In the F_2 layer, a temporary ''winner'' node is chosen based on the ''*winner takes all*'' rule.[8] It means that only the pixel having a maximum input is chosen as the winner. While it is selected, only the winner is active and all the other nodes are inactive.

Let us assume that node J is chosen as the winner node. The current winner then goes back to the F_1 layer through the top-down pathway. In the F_1 layer, the top-down vector of node J, which is $\widetilde{w}_J$, is matched with the bottom-up vector (input vector). The degree of match is sent to the orienting subsystem. A match is compared to a preset threshold ρ. If the degree is greater than or equal to the threshold, then the current winner J is recommended as the output of the F_2 layer, and the system goes into a resonance state. Otherwise, a reset wave is generated, and the system goes into the reset state.

In the resonance state, the system starts to learn from the current input. In ART1, because all the data are binary, the following fast learning is operated:

$$\widetilde{w}_J^{new}=I\cap\widetilde{w}_J^{old} \tag{2}$$

where the symbol $\cap$ is the logical AND operator.[8] The bottom-up weights are also changed so that the pathway between two nodes in either bottom-up or top-down is the same.

In the reset state, the current winner is prohibited by the system and will never be the winner for the current input. The bottom-up vector is sent to F_2 layer again, and the ''winner takes all'' and matching procedures are repeated. The procedure is repeated until one of the following scenarios occurs:

- A winner is found with a matching degree greater than or equal to the vigilance parameter, and the system goes into the resonance state.
- None of the nodes in the F_2 layer can match the rule. Then a new node is inserted into the F_2 layer as the output of the system. This implies that the current input is a new kind of object, and the system goes into resonance again.
- All the memory of the system is used up, and the system stops working.

From the description above, it can be observed that the ART network has many characteristics that are unique among other learning schemes:

- ART constitutes an unsupervised learning scheme. The system self-organizes its bottom-up, top-down, and output nodes during the classification. Other systems need previously fixed answers for all the possible inputs.
- ART can work in a nonstationary circumstance. It can also assign the unpredictable input to a new class and protect the previously learned information from contamination. An unexpected input to some neural net-

work systems can wipe out their prior information, which makes them impracticable in real circumstances.

The ART network is, therefore, a powerful tool in real-time object recognition capable of dealing with binary inputs. The limitation of binary inputs has been solved in ART2 and ART3, however, the computational complexity of ART2 and ART3 prevents them from being used widely. Fuzzy ART, with its capability to cope with analog data, was consequently developed as a reliable and easily implemented network.[10] Fuzzy ART keeps the structure of ART1, but employs the fuzzy rather than crisp logical rules. The fuzzy rules in the computation make it possible to process the continuous data.

2.2 *Fuzzy ART*

Fuzzy ART has the same network structure as ART1 but it incorporates a fuzzy set computational theory in the processing steps. Some definitions, concepts, and notations regarding fuzzy ART are now given.

Input vector. The input vector is an M-dimensional vector of real values, $\mathbf{I}=(I_1,\ldots,I_M)$, where $I_j\in[0,1]$, and $j=1,\ldots,M$.

Weight vector. The top-down and bottom-up vectors and their definitions are maintained in the Fuzzy ART. The initial values of all the weights are set to unity (1.0).

Committed and uncommitted nodes. A node in the F_2 layer is said to be committed if it has been chosen as the output of the system, otherwise, it is termed uncommitted.

Parameters. There are three parameters in the fuzzy ART system, namely:

- choice parameter $\alpha>0$,
- learning parameter $\beta\in[0,1]$,
- vigilance parameter $\rho\in[0,1]$.

Category choice. For each input vector, the system will produce an output based on one of the nodes in the F_2 layer; the ''winner.'' The choice of the ''winner'' is calculated as:

$$T_j=\frac{|\mathbf{I}\wedge\mathbf{w_j}|}{\alpha+|\mathbf{w_j}|} \tag{3}$$

$$T_J=\max\{T_j:j=1,\ldots,N\} \tag{4}$$

where T_j is the input of the j'th node in the F_2 layer. The bottom-up weight is $\mathbf{w_j}$. The symbol ''$\wedge$'' is the fuzzy and minimum operator,[10] which is defined as:

$$(\mathbf{x}\wedge\mathbf{y})_\mathbf{i}=\min(x_i,y_i) \tag{5}$$

and the norm ''$|\,|$'' operator is defined as:

$$|\mathbf{x}|\equiv\sum_{i=1}^{M}x_i. \tag{6}$$

If the J'th node is chosen as the ''winner'' in the first instance, it is changed from an uncommitted node to a committed node when the next input data is presented.

Resonance or reset. When a node is chosen as the ''winner,'' a match parameter m is calculated as:

$$m=\frac{|\mathbf{I}\wedge\mathbf{w_j}|}{|\mathbf{I}|} \tag{7}$$

indicating the degree of match between the input vector and the top-down vector. The match parameter is compared to the vigilance parameter ρ. If $m\geqslant\rho$, the system goes into the resonance state, otherwise, it goes into the reset state.

In the resonance state, the network starts to learn the current input adaptively as:

$$\mathbf{w_J^{new}}=\beta(\mathbf{I}\wedge\mathbf{w_J^{old}})+(1-\beta)\mathbf{w_J^{old}}. \tag{8}$$

A fast learning occurs when the learning parameter is equal to unity (1.0). The weight therefore learns from the input as much as possible. Usually, for an uncommitted node, the learning parameter is set to unity (1.0). Once, it becomes a committed node, the learning parameter is taken as $\beta<1.0$. This technique is called fast commitment option. It ensures that the system can learn a new kind of input quickly as all the weights are greater than or equal to the components in the input vector. The current input can be fully ''caught'' in this case based on Eq. (5). When there is an input whose winner in the F_2 layer is a committed node, the system records it slowly without sweeping the existing memory. After learning, the bottom-up and the top-down weights between two nodes are kept unchanged.

In the reset state, a reset wave is generated by the orienting subsystem (see Figure 1). It prohibits the current ''winner'' from being active again. The winner's output is always equal to minus one (-1.0) for the current input. The system starts to learn the input again the same way as ART1. All the prohibited nodes can revive only after the occurrence of resonance.

Input normalization option. The proliferation problem described by Moore,[15] in which the weights of the ART system are eroded when a large number of inputs with different norms are introduced, is resolved in fuzzy ART in two ways. The first method is to normalize all the inputs before they are sent to the system, that is:

$$\mathbf{I}=\frac{\mathbf{a}}{|\mathbf{a}|}. \tag{9}$$

The other option is called complement coding. This technique realizes the normalization without losing the information of the amplitude. A vector has its on-response and off-response in the system. The on-response is defined as itself, and the off-response is defined as:

$$a_i^c=1-a_i. \tag{10}$$

In Eq. (10), a_i^c is the component of the off-response vector. In this situation, the input vector $\mathbf{I}$ is constructed as:

$$\mathbf{I}=(\mathbf{a},\mathbf{a}^{\mathbf{c}})=(a_1,\ldots,a_M,a_1^c,\ldots,a_M^c) \tag{11}$$

where M is the number of on-response (off-response) components in the input vector. It can be easily concluded that:

$$|\mathbf{I}|=\sum_{i=1}^{M}(a_i+a_i^c)=M. \tag{12}$$

If the input vector is made of binary data, fuzzy ART will be equivalent to ART1. In this paper, fuzzy ART is further developed into a feature-adaptive ART, and the new scheme is exploited to accomplish the classification task.

3 Wavelet/Neural Network in Texture Classification

A new wavelet/neural network system is proposed to accomplish the texture recognition task. In this scheme, there are two layers: a Gabor wavelets system that extracts the space/spatial-frequency features of the input texture images, and the feature-adaptive ART neural network that processes the extracted feature vectors to perform the classification/recognition task.

3.1 *Feature Extractor Module*

The wavelet/neural network system is motivated by the preattentive and attentive processes of the human visual system. The first layer simulates the feature extraction task performed by the simple cells found in the visual cortex, while the neural network layer is more akin to the higher level cognition processes performed by the brain. Gabor wavelets have been shown to resemble the receptive field profile of the simple cells,[6,7] and are capable of texture feature extraction.[1,2,4,5] Gabor wavelets have been used as the feature extractors in the proposed system.

The Gabor wavelets used for image feature extraction are exactly like those used in Ref. 1 and are defined as:

$$h(x,y)=\exp\left[-\alpha^{2j}\,\frac{x^2+y^2}{2}\right]\cdot\exp[j\pi\alpha^j(x\cos\theta+y\sin\theta)] \tag{13}$$

where $\alpha=1/\sqrt{2}$, $j=0, 1, 2, \ldots,$ and $\theta=k\pi/N$, $N=0, 1, 2, \ldots,$ $k=0, 1, 2, \ldots, N-1$.

The four orientations used in the work reported in this paper are 0, $\pi/4$, $\pi/2$, and $3\pi/4$. The choice of the frequency components is, however, adaptive. Initially, the j in Eq. (13) is chosen as zero (0). Therefore, at first, four wavelets are used resulting in a feature vector with four components representing four orientations and one frequency. As the system develops, the family of the wavelets is enlarged by adding more frequencies.

The Gabor wavelets are used to construct a bank of spatial domain filters. Each filter is made of a pair of filters that are the real and imaginary part of the complex sinusoid. For each image, several (100 is chosen in this paper) pixels are picked for consideration, and the filter pairs are convolved with the texture in these positions. The output of a filter pair is calculated as:

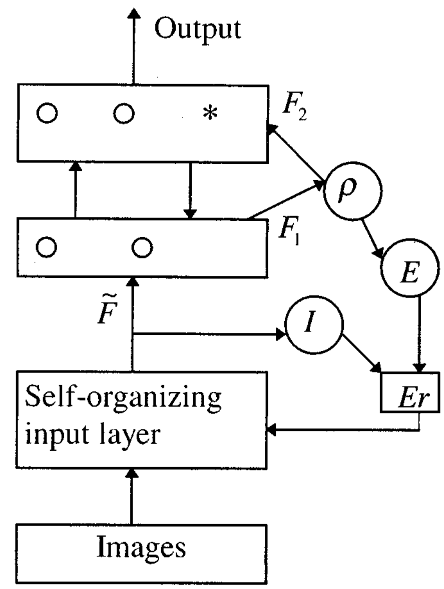


Fig. 2 The wavelet-neural network system.

$$\text{Output}=\sqrt{R_{output}^2+I_{output}^2}. \tag{14}$$

In Eq. (14), the R_{output} and I_{output}, respectively, represent the response of the real and imaginary parts of the Gabor filter pair. The mean of the outputs of one filter pair at different positions is stored as one feature of the texture. In other words, every filter pair is employed to capture one feature of the texture. For each texture, a multidimensional feature vector is constructed based upon the filters used, and for every new input image, the same set of filters is used to construct the feature vector.

3.2 *Feature-Adaptive ART in Texture Classification*

Several proposed classification methods have been proven to be effective in the post-processing of the feature vectors.[1,2,5] However, they mostly reflect the human knowledge view, rather than the neurological structure of the human visual system and the brain. Neural networks are currently regarded as the closest artificial structure to the human brain. Since the main aim in this work is to develop an artificial vision system motivated by some of the known mechanisms in the human visual system, a neural network is chosen as the high-level processing module. The selection of the fuzzy ART among the numerous other neural network architectures is due to its self-organizing property and the possibility of real-time operation.

Some modifications, however, have been made to the conventional fuzzy ART to add a new self-organization capability at the input layer; this addition turns out to make it more robust to noisy inputs. In a network with a fixed input layer, the possibility of acquiring more features to discriminate a noisy input is obviated. The new structure of the neural network is depicted in Figure 2. Two new counters, ''I'' and ''E,'' which keep track of the number of inputs and errors of the system, respectively, are added to the fuzzy ART system. A special node in the F_2 layer, denoted as ''*'' and referred to as error node, helps to indicate when the system is unable to classify the current input. The

''Er'' in Figure 2 will be introduced later in Eq. (16). Other structures in the F_1 and F_2 layers, such as top-down and bottom-up weights, and the reset wave in the orienting system (''ρ'' in Figure 2) are the same as in the conventional ART. The calculation rule in the fuzzy set membership theory is retained. The fundamental knowledge of the inputs is provided by initially introducing several feature vectors of each texture to the system and is stored in the bottom-up and top-down weights.

With each new input, the input counter is incremented. As aforementioned, the input to the network is the feature vector (with four components), which is $F=\{f_1,f_2,f_3,f_4\}$ and computed as the output of the Gabor wavelets system. Each feature vector is normalized according to Eq. (9) as:

$$\tilde{f}_i=\frac{f_i}{\Sigma_i f_i} \tag{15}$$

where $i \in [1,4]$. Complement coding, as shown in Eq. (11), is also applied to the input, thus the input vector has four on-response and four off-response, which is $\tilde{F}=\{\tilde{f}_1,\tilde{f}_2,\tilde{f}_3,\tilde{f}_4,\tilde{f}_1^c,\tilde{f}_2^c,\tilde{f}_3^c,\tilde{f}_4^c\}$, where $\tilde{f}_i^c=1.0-\tilde{f}_i$. This will eliminate the detrimental effects of contrast variance of the input images and the proliferation of the weights as described in Ref. 15.

The input vector $\tilde{F}$ then passes through the bottom-up pathway, and a temporary winner is generated in the F_2 layer according to Eqs. (3) and (4). Each node in the F_2 layer is an index of the ''shape'' vector. The real ''shape'' of the vector is in its top-down weights. The top-down weights of the temporary winner are compared to the input by the match parameter [Eq. (7)]. If $m \geqslant \rho$, the system enters a resonance state, otherwise, it goes to reset state. The value of the vigilance parameter ρ is preset to be very high (more than 0.95 normally). This ensures that the inputs cannot be classified into wrong groups.[8] However, the choice of a high vigilance parameter may cause a problem by generating excessive classes of textures for the same textures captured in slightly different environment. This problem will be resolved in the proposed system by increasing the feature components in the feature vector at the self-organizing stage.

When a resonance occurs, the system starts to learn from the input data. The learning parameter in Eq. (8) is set as 0.5. This protects the previously learned information in the system from being wiped out by the new input.

When the top-down vector fails to match the input vector, the system will go to the reset state and stops the search for the matched winners in the F_2 layer. This is due to the fact that the temporary winner in the F_2 layer is the one that is most similar to the input vector by the ''winner takes all'' theory. If the winner fails, it is intuitive that no other vectors (nodes) will win, otherwise, the new winner should be the winner in the first place. Hence, termination of the search process saves time and calculation without affecting the normal functioning of the system. The reset wave from the orienting subsystem excites the error node in the F_2 layer. The system produces a signal that signifies its inability to recognize the current input, and at the same time the error counter is incremented.

At the end of these steps, the system checks the content of the counters. If the number in the input counter is smaller than a threshold value, for example 200 in this work, the system prepares for the next input, otherwise, it looks at the error ratio Er, which is given as:

$$Er=\frac{\text{number of errors}}{\text{number of inputs}}. \tag{16}$$

The value of Er is compared to a preset threshold (set by the user) that satisfies the system requirements. If Er is smaller than some threshold (92% in this work), the system resets both counters to zero, otherwise, the system goes into a self-organizing state. The threshold here is set as 92% because that is the highest accuracy of classification in the test of four features (refer to the results in Table 2); it will ensure the system self-organizes its input at least once. If the threshold level is set to the maximum (i.e., 100%), the system will self-organize too frequently and may not reach the desired accuracy. If, on the other hand, the threshold level is too low, the system will temporarily be stable until it cannot satisfy the requirement of the user, whereupon it self-organizes.

In the self-organizing state, the system increases the number of wavelets; one more frequency component is added to the Gabor wavelets family each time. Hence, in Eq. (13), j equals 0 and 1. The new frequency is always one octave above the existing ones in the system. Nodes in the F_1 layers of the neural network and the bottom-up and top-down weights are increased accordingly. All other parameters in the system are retained. Samples from each texture are tested during the training of the new system.

Increasing the number of features improves the discriminability of the feature set, thereby leading to a more accurate classification. The misclassified images are grouped together; this reflects the procedure in human post-natal development in which the visual system of a new born baby is far less developed than that of an adult. As time goes on, the post-natal development fine-tunes the visual system, thus enabling it to detect and recognize many more objects. From the computational point of view, the proposed system is efficient as it increases the components stepwise.

From the ongoing description, it is clear that the proposed feature-adaptive system is distinct from the fuzzy ART proposed in Ref. 10. In the ART system proposed by Grossberg and Carpenter,[8–10] the system can self-organize the nodes only in the output layer, and top-down and bottom-up weights. However, in some cases, the features in the input layer may not be adequate in differentiating various textures; for example, two different textures may have the same frequency and orientation in some degree. In the proposed system, however, the input of the system as well as the output layer are extended as the test cases are expanded. This is akin to the human visual system where the feature detectors undergo a stepwise refinement as the experiences increase. Post-natal development is essential to the construction of the visual system. In the proposed scheme, the system is enlarged as more failures occur. It self-organizes not only the output layer but also the input layer. This is a crucial difference between the conventional ARTs and the one proposed here.

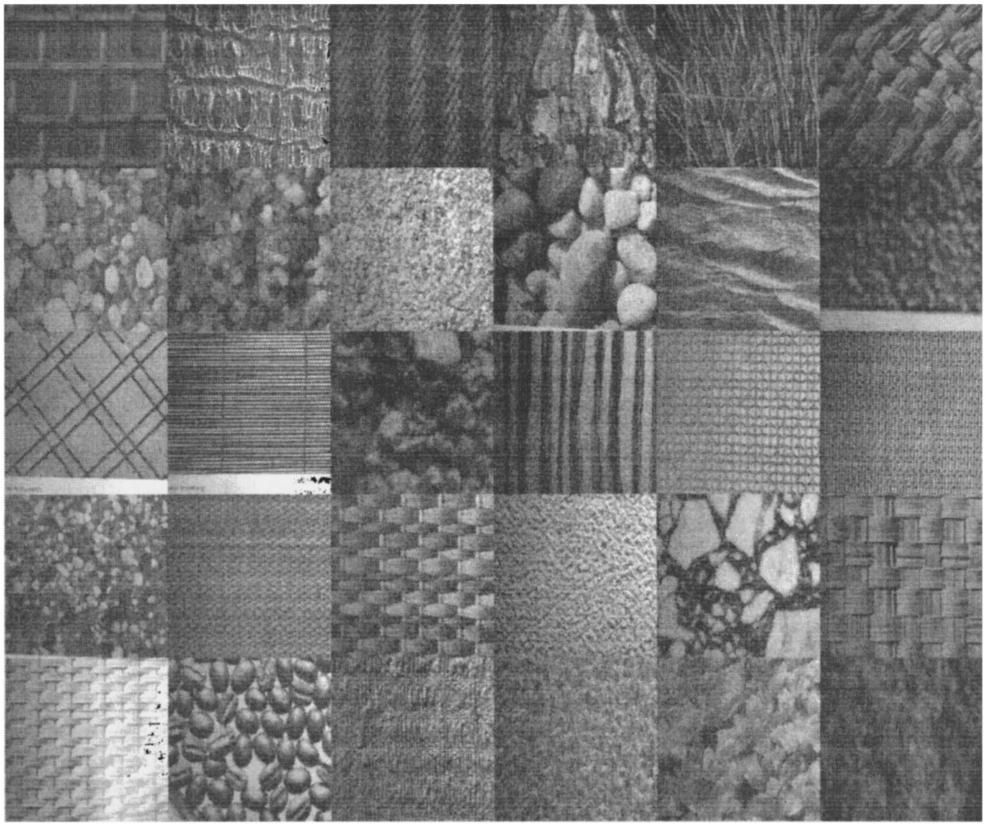


Fig. 3 Some texture images.

The results of experiments, presented in Section 4, show that the new system is capable of coping with situations in which it is desirable to have a variable number of input features. A system with a fixed number of features obviously cannot perform effectively as the number of test cases expands.

4 Results and Discussion

Fifty natural textures from the Brodatz album[16] are used in testing the proposed system. Some of the textures are shown in Fig. 3. Twenty samples of each texture are captured by a CCD camera under varying environments that include changing orientations and background lighting, and different distances between the camera and the textures. The samples are all 256×256 pixels images quantized to 256 gray levels and stored separately. Images from the same texture are grouped together; there are therefore fifty groups of images.

Results in Table 1 show the results of five tests with 1000 images in which the accuracy of classification before and after self-organizing are compared. The accuracy of classification is defined as the number of the images that are classified correctly out of the number of the total input images. Eight features are exploited in the ''before self-organizing'' tests, while twelve features are used in the ''after self-organizing.'' The improvement obtained is evident when the last two columns of the table are compared. Two hundred randomly selected texture images are used in each test because the system checks its Er value after every 200 inputs. In each test, the threshold for Er is set as 1.0, so as to ensure that the system self-organizes after 200 samples. The improvement obtained with the new system can be attributed to the extension of the input as well as the output layer during the self-organization process.

The performance of the proposed system is compared to that of the fixed ART (three feature vector sizes are used: 4, 8, 12). Table 2 summarizes the comparison. The first column in Table 2 indicates the number of groups of images used, and the second column is the accuracy of the classification (the size of the feature vector is given in parenthesis) obtained via the feature-adaptive ART. The third, fourth, and fifth columns summarize the accuracy of the classification obtained by using the fixed ART with feature vector size of 4, 8, and 12, respectively. The reader's attention is drawn to the fact that as the number of groups of

Table 1 Accuracy of the classification before and after self-organization (numbers in parenthesis are the number of features used in the test).

	Before self-organizing	After self-organizing
Test 1	0.90 (8)	0.94 (12)
Test 2	0.92 (8)	0.95 (12)
Test 3	0.84 (8)	0.89 (12)
Test 4	0.87 (8)	0.92 (12)
Test 5	0.91 (8)	0.93 (12)

Table 2 Comparison of the classification accuracy of the feature-adaptive and fixed-feature systems (numbers in parenthesis are the number of features used in the test).

Number of groups	Adaptive feature	Fixed feature	Fixed feature	Fixed feature
10	0.92 (4)	0.92 (4)	0.95 (8)	0.96 (12)
20	0.91 (4)	0.91 (4)	0.93 (8)	0.94 (12)
30	0.92 (8)	0.83 (4)	0.92 (8)	0.93 (12)
40	0.90 (8)	0.71 (4)	0.90 (8)	0.92 (12)
50	0.92 (12)	0.65 (4)	0.83 (8)	0.92 (12)

textures used in the test increases, the accuracy of the fixed ART decreases while that of the feature-adaptive ART is maintained. The feature-adaptive system self-organizes itself. In the test with fifty groups of textures, up to twelve features are used, corresponding to $j=0,1,2$ in Eq. (13) and the four orientations. The four orientations are exploited in all the fixed systems with different frequencies. In this experiment, all the samples from the randomly selected groups are presented to all the systems five times. Each time, ten more groups are added.

In Table 3, the performance of the new system is compared with that of the scheme proposed in Ref. 1. In Ref. 1, a minimum distance classifier is used; the mean feature vector of each group is obtained and used to construct a codebook of feature vectors. During the test, the feature vector of the current image is compared to the ones in the codebook, and the one with minimum distance is found. The current image is then ''coded'' or classified by the codevector. In this experiment, both systems are tested five times. Each time, ten more groups of the images are tested. The proposed system self-organizes its structure, and up to twelve features are exploited when fifty groups are presented to it. The minimum distance based system has twelve features that are four orientations and three frequencies, where $j=0,1,2$ in Eq. (13).

The results of the experiments are presented in Table 3; the wavelet/neural network system outperforms the system proposed in Ref. 1. The main difference between these two systems is the feature classifier.

5 Summary

In this paper, a two-layer system, motivated by the pre-attentive and attentive as well as the post-natal processes of the human visual system, is proposed. Gabor wavelets are applied to obtain the space/spatial-frequency characteristics of the texture images. It ''simulates'' the low-level feature extraction function in the human visual system. Feature-adaptive ART, a self-organizing neural network, is exploited as a classifier ''simulating'' the learning and cognition processes. The proposed system is tested extensively through the presentation of up to 50 natural textured images from the Brodatz album.[16] A stable classification rate of about 93% is achieved with extremely compact feature vector and efficient real-time processing. The system outperforms all the previous generations of the Gabor based texture classification methods. The results show that the wavelet/neural network system is a promising technique in the artificial vision research area.

Table 3 Comparison of the classification accuracy of the wavelet/neural network and wavelet/minimum distance systems (numbers in parenthesis are the number of features used in the test).

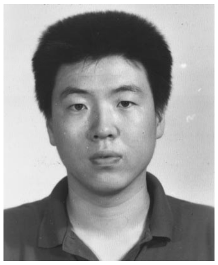
Number of groups	Wavelet/neural network	Wavelet/minimum distance
10	0.93 (4)	0.94 (12)
20	0.92 (4)	0.93 (12)
30	0.92 (8)	0.90 (12)
40	0.92 (8)	0.87 (12)
50	0.93 (12)	0.66 (12)

Acknowledgment

This work was supported by a grant from the Industrial Automation Research Centre at the University of Wollongong.

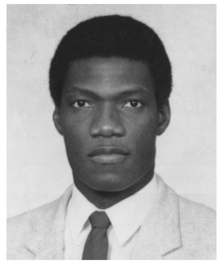
References

1. G. Naghdy, J. Wang, and P. Ogunbona, ''Texture analysis using Gabor wavelets,'' *IS&T/SPIE Symp. Electronic Imaging, Proc. SPIE* **2657**, 74–85 (1996).
2. A. C. Bovik, M. Clark, and W. S. Geisler, ''Multichannel texture analysis using localized spatial filters,'' *IEEE Trans. Patt. Anal. Mach. Intell.* **12**(1), 55–73 (Jan. 1990).
3. M. Porat and Y. Y. Zeevi, ''Localized texture processing in vision: analysis and synthesis in the Gaborian space,'' *IEEE Trans. Biomed. Engineering* **36**(1), 115–129 (Jan. 1989).
4. A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, ''Localized measurement of emergent image frequencies by Gabor wavelets,'' *IEEE Trans. Info. Theory* **38**(2), 691–712 (March 1992).
5. H. Greenspan, R. Goodman, R. Chellappa, and C. H. Anderson, ''Learning texture discrimination rules in a multiresolution system,'' *IEEE Trans. Patt. Anal. Mach. Intell.* **16**(9), 894–901 (Sep. 1994).
6. J. G. Daugman, ''Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters,'' *J. Opt. Soc. Am. A* **2**(7), 1160–1169 (July 1985).
7. S. J. Anderson and D. C. Burr, ''Receptive field properties of human motion detector units inferred from spatial frequency masking,'' *Vision Research* **29**(10), 1343–1358 (1989).
8. G. A. Carpenter and S. Grossberg, ''The ART of adaptive pattern recognition by a self organizing neural network,'' *IEEE Computer* **21**(3), 77–88 (March 1988).
9. G. A. Carpenter and S. Grossberg, ''ART 3: hierarchical search using chemical transmitters in self-organizing pattern recognition architectures,'' *Neural Networks* **3**, 129–152 (1990).
10. G. A. Carpenter, S. Grossberg, and D. B. Rosen, ''Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system,'' *Neural Network* **4**, 759–771 (1991).
11. A. Dubrawski and P. Reignier, ''Learning to categorize perceptual space of a mobile robot using fuzzy-ART neural network,'' IROS' 94, *Proc. IEEE/RSJ/GI Inter. Conf. Intelligent Robots and Systems* **2**, 1272–1277 (1994).
12. H. M. Lee and C. S. Lai, ''Supervised fuzzy ART: training of a neural network for pattern classification via combing supervised and unsupervised learning,'' *1993 IEEE Inter. Conf. Neural Networks* **1**, 323–328 (1994).
13. G. A. Carpenter, S. Grossberg, and J. H. Reynolds, ''A fuzzy ARTMAP nonparameteric probability estimator for nonstationary pattern recognition problems,'' *IEEE Trans. Neural Networks* **6**, 1330–1336 (Nov. 1995).
14. G. A. Carpenter and W. D. Ross, ''ART-EMAP: A neural network architecture for object recognition by evidence accumulation,'' *IEEE Trans. Neural Networks* **6**, 805–818 (July 1995).
15. B. Moore, ''ART1 and pattern clustering,'' in, D. Touretzky, G. Hilto, and T. Sejnowski, Eds., Proc. 1988 Connectionist Models Summer School, Morgan Kaufmann Publishers, San Mateo, CA (1988).
16. P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Denvor, New York (1966).



Jian Wang received the BE degree in biomedical engineering in 1993 at the Capital Institute of Medicine, Beijing, China. He is now pursuing his PhD study in the Department of Electrical & Computer Engineering, University of Wollongong, Australia. His research interests are in image compression, computer vision, and medical image processing.

Golshah Naghdy is a senior lecturer at the Department of Electrical and Computer Engineering, University of Wollongong, Australia. She received a BSc degree in electrical and electronic engineering from Aryamehr University of Technology in Tehran in 1977. She received an MPhil in control engineering from Bradford University and a PhD in electronic engineering from Portsmouth University, England, in 1982 and 1986, respectively. She was a senior lecturer at Portsmouth University before emigrating to Australia in 1989. Her major research interests are in biological and machine vision and medical image processing.



Philip Ogunbona received the BSc (Hons) degree in electronic and electrical engineering from University of Ife, Nigeria, in 1981 and the DIC, PhD in electrical engineering from Imperial College of Science, Technology and Medicine, University of London, United Kingdom, in 1987. He is currently a senior lecturer in the Department of Electrical and Computer Engineering, University of Wollongong, Australia. His current research interests include wavelets, image and video compression, and multimedia databases. He is a member of the IEEE.