1-1-1996

New wavelet based ART network for texture classification

Jiazhao Wang
University of Wollongong, jiazhao@uow.edu.au

Golshah Naghdy
University of Wollongong, golshah@uow.edu.au

Philip O. Ogunbona
University of Wollongong, philipo@uow.edu.au

Abstract

A new method for texture classification is proposed. It is composed of two processing stages, namely, a low level evolutionary feature extraction based on Gabor wavelets and a high level neural network based pattern recognition. This resembles the process involved in the human visual system. Gabor wavelets are exploited as the feature extractor. A neural network, Fuzzy Adaptive Resonance Theory (Fuzzy ART), acts as the high level decision making and recognition system. Some modifications to the Fuzzy ART make it capable of simulating the post-natal and evolutionary development of the human visual system. The proposed system has been evaluated using natural textures. The results obtained show that it is able to effectively perform the object recognition task and will find useful application in the study of the human visual system model for artificial vision.

Keywords

texture, classification, network, wavelet, art

Disciplines

Physical Sciences and Mathematics

Publication Details

Wang, J., Naghdy, G. & Ogunbona, P. O. (1996). New wavelet based ART network for texture classification. Proceedings of the Australian and New Zealand Conference on Intelligent Information Systems (pp. 250-253). IEEE.

Proc. 1996 Australian New Zealand Conf. on Intelligent Information Systems, 18-20 November 1996, Adelaide, Australia. Editors, Narasimhan and Jain

A New Wavelet Based ART Network For Texture Classification

J. Wang, G. Naghdy, and P. Ogunbona
Department of Electrical and Computer Engineering
University of Wollongong
Australia

Tel: +61-42-214689, *FAX*: +61-42-213236
email: jiwa@elec.uow.edu.au

Abstract: A new method for texture classification is proposed. It is composed of two processing stages, namely, a low level evolutionary feature extraction based on Gabor wavelets and a high level neural network based pattern recognition. This resembles the process involved in the human visual system. Gabor wavelets are exploited as the feature extractor. A neural network, Fuzzy Adaptive Resonance Theory (Fuzzy ART), acts as the high level decision making and recognition system. Some modifications to the Fuzzy ART make it capable of simulating the post-natal and evolutionary development of the human visual system. The proposed system has been evaluated using natural textures. The results obtained show that it is able to effectively perform the object recognition task and will find useful application in the study of the human visual system model for artificial vision.
Keywords: Human visual system, Fuzzy ART, Gabor Wavelets, texture classification.

1. INTRODUCTION

Texture analysis plays an important role in vision research and there are attempts to develop systems based on the multi-channel model of the human visual system [1] [2]. However, most of the studies in this area are limited to the investigation of the properties of the low level cells in the human visual system and are concentrated on the feature extraction level. In this paper, a neural network, Fuzzy Adaptive Resonance Theory (Fuzzy ART), is employed to simulate the capacity of the high level cells in the attentive process of the human visual system. The proposed scheme starts by learning simple frequency components. The system is able to refine itself in a manner similar to the post-natal development in the visual system, as the test environment becomes more complex. Several natural textures have been tested on the proposed system. Results obtained show that Gabor wavelets-Fuzzy ART system can be used as a multi-tier model of the human visual system.

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

2. WAVELET-NEURAL NETWORK IN TEXTURE CLASSIFICATION

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

2.1 Feature extractor module

The wavelet-neural network system is proposed to simulate the pre-attentive and attentive processes of the human visual system. The first layer simulates the feature extraction task performed by the simple cells in the visual cortex while the fuzzy 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 in cats [3]; this observation motivates its choice as the feature extractor in the proposed system.

The Gabor wavelets used for image feature extraction are exactly like those used in [1], and are defined as:

$$h(x, y) = \exp[-\alpha^{2j}\frac{x^2+y^2}{2}]\cdot\exp[j\pi\alpha^{j}(x\cos\theta + y\sin\theta)] \qquad (1)$$

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

Four orientations are used in the work reported in this paper, $0, \frac{\pi}{4}, \frac{\pi}{2}$, and $\frac{3\pi}{4}$. The choice of the frequency components is, however, adaptive. Initially, the j in equation (1) is chosen as zero. As the system develops, the family of the wavelets will be enlarged by adding more frequencies. This represents the evolutionary and post-natal developments similar to the human visual system. Each wavelet filter is made of a pair of filters that are the real and imaginary part of the complex sinusoid. These pairs are convolved with the texture images separately. The output of a filter is the magnitude of the output of complex sinusoid. It is calculated as:

$$Output = \sqrt{R^2_{output} + I^2_{output}} \tag{2}$$

The mean of the outputs of one filter in different positions is stored as one feature of the texture. In other words, every filter 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 are used to construct the feature vector.

2.2 Fuzzy ART in texture classification

In this work, the main aim is to develop an artificial vision system that models the mechanism in the human visual system. Much of the processing taking place in the human visual system is non-linear and a neural network is able to capture such non-linearity. The selection of the Fuzzy ART among the numerous neural networks available, is due to its properties of self-organising and real-time learning.

The structure of the Fuzzy ART well described in [4] [5]. It is also depicted in Figure 1. Two counters which keep track of the number of inputs and errors of the system, "*I*" and "*E*" respectively in Figure 1, are added to the Fuzzy ART system. A special node in the F_2 layer, denoted as "*" in Figure 1, and referred to as *error node*, signifies when the system is unable to classify the current input. The operator "*/*" in Figure 1 will be introduced later in equation (4). Other structures in the F_1 and F_2 layer, such as top-down and bottom-up weights, and the reset wave in the orienting system, that is "ρ" in Figure 1, are the same as in the conventional Fuzzy ART [5]. The calculation rule in the theory of fuzzy set membership is observed [5]. However some modifications have been made to make the system more reliable. The fundamental knowledge of the inputs is provided by initially introducing several feature vectors of each texture to the system. This knowledge is stored in the bottom-up and top-down weights.

With every new input, the input counter is incremented by one. The input vector is passed through the bottom-up pathway, and a temporary winner is generated in the F_2 layer. The top-down weights of the temporary winner are compared to the input by the match function [5] as:

$$m = \frac{|\mathbf{I} \wedge \mathbf{w}_J|}{|\mathbf{I}|} \tag{3}$$

where m is the degree of match between the input vector $\mathbf{I}$ and the top-down vector $\mathbf{w}_J$. The match degree is compared to the vigilance parameter, ρ. If $m \geq \rho$, the system enters a *resonance stage*, where it starts to learn the knowledge from the input, otherwise, it goes to *reset* stage. The value of ρ is preset to be very high, normally more than 0.95. This ensures that the inputs cannot be mis-classified [4]. However, the choice of high a vigilance parameter may cause problem by generating excessive classes of textures for the same textures captured in a slightly different environment. This problem will be resolved in the proposed system by increasing the feature components in the feature vector at the self-organising stage.

In the reset stage, the system discontinues the search for the matched winners in the F_2 layer. 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 recognise the current input. At the same time, the error counter is incremented by one.

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,

$$Er = \text{number of errors} / \text{number of inputs} \tag{4}$$

The ratio, *Er*, is compared to a preset threshold that satisfies the requirement of the system user. If *Er* is smaller than some threshold (92% in this work), the system resets both the numbers in the error counter and input counter to zero and prepares for the new input, otherwise, the system goes into a *self-organising stage*.

In the self-organising stage, the system increases the number of wavelets. One more frequency component is added into the Gabor wavelets family each time. Hence, in equation (1), j equals 0 and 1. The new frequency is always one octave above the existing ones in the system. This is analogous to the evolutionary and post-natal development of the human visual system [3]. Nodes in the STM layers of the neural network and the bottom-up and top-down weights are increased accordingly. All the parameters in the system are retained. Samples from each texture are tested during the training of the new system. The counters are reset to zero and the system starts again.

Increasing the number of features improves the discriminability of the feature set thereby leading to a more accurate classification. The mis-classified images are grouped together. This reflects a similar procedure in the human post-natal development. The visual system of a newly 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 recognise many more objects. From a 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 neural network system dissimilar to the ART proposed in [5]. In the ART system proposed by Grossberg and Carpenter [4] [5], the system can self-organise 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 enough to differentiate 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 is extended as well as the output layer as the exemplar increases. 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-organises not only the output layer but also the input layer. This is a crucial difference between the conventional ARTs and the one proposed here.

The results of the experiments also corroborate the logic of the proposed system.

3. RESULTS AND DISCUSSION

Fifty natural textures from the Brodatz album [6] are tested in this research. They are all 256x256 pixels images quantised to 256 grey scales. Twenty samples of each texture are captured under varying environmental conditions and grouped together. Therefore, there are fifty groups of images being used.

The performance of the proposed system is compared to that of a fixed ART (three feature vector sizes are used: 4, 8, and 12). Table 1 summarises the comparison test . The first column indicates the number of groups of images used, and the second column is the accuracy performance of the feature-adaptive ART. Notice the increase in the feature vector size. The third, fourth and fifth columns summarises the accuracy performance of the fixed ART with vector sizes of 4, 8 and 12 respectively. Notice that the accuracy of the fixed ART decreases as the number of groups used in the test increases while the feature-adaptive ART sustains its performance. The feature-adaptive system self-organises itself. In the test of fifty groups of textures, up to twelve features are exploited, which are j=0,1,2 in (1) and 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 sent to all the systems five times. Each time, ten more groups are added.

From the results, it can be seen that as more groups are included, the fixed feature system losses its classification ability. However, in the feature-adaptive system the accuracy is almost stable as more textures are included.

In Table 2, the new system is compared to the scheme proposed in [1]. In the scheme in [1], the features are classified by the minimum distance method. The mean feature vector of each group is obtained and used to construct the codebook of feature vector. During the test, the feature vector of 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-organises its structure, and up to twelve features are exploited when fifty groups are sent to it. The minimum distance classifier based system has twelve features that are four orientations and three frequencies, where j=0,1,2 in (1).

It can be seen from Table 2 that the wavelet-neural network system outperforms the system described in [1]. The main difference between these two systems is the feature classifier. It seems that a classifier more akin to the human brain performs better in object recognition tasks.

4. SUMMARY

In this paper, a two-layer system, which simulates the pre-attentive and attentive as well as post-natal processes of the human visual system, is proposed. Gabor wavelets is applied to obtain the spatial/frequency characters of the texture images. thus simulating the function of low-level feature extraction in the visual system. Fuzzy ART, a self-organising neural network, is exploited as a classifier to simulate the learning and cognition processes. Modifications to the conventional ART makes the proposed network work in a stable output environment, while the property of self-organising the feature components in the system saves computation in real-time applications. Several tests have been performed on the new system. The results show that the wavelet-neural network system is a promising technique in the artificial vision research area.

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)

Table 1, the accuracy of the classifications in adaptive feature and fixed feature.

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

Table 2, the accuracy of classifications by wavelet-neural network system and wavelet-minimum distance.

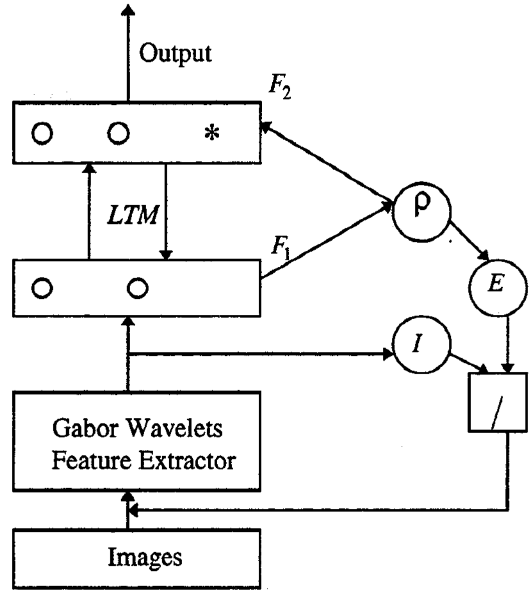

Figure 1, the Wavelet-Neural network system.

REFERENCE

[1] G. Naghdy, J. Wang, and P. Ogunbona, "Texture analysis using Gabor wavelets", IS&T/SPIE Symposium on Electronic Imaging, Vol. 2657-10, San Jose, USA, Jan. 27 -- Feb. 2, 1996.

[2] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 1, pp. 55 -- 73 January 1990.

[3] 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., Vol. 2, No. 7, pp. 1160-1169, July 1985.

[4] G. A. Carpenter and S. Grossberg, "The ART of adaptive pattern recognition by a self organising neural network", IEEE Computer, pp. 77 -- 88, March 1988.

[5] 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, Vol. 4, pp. 759 -- 771, 1991.

[6] P. Brodatz, "Textures: A Photographic Album for artist and designers", New York, Denver, 1966.