NEURAL VISION SYSTEM AND APPLICATIONS IN IMAGE PROCESSING AND ANALYSIS

Ling Guan, Stuart Perry, Raniero Romagnoli, Hausan Wong and Haosong Kong

Department of Electrical Engineering
The University of Sydney
Sydney NSW 2006 Australia

Abstract

We present a computer vision system based on an integrated neural network architecture. In the low level vision subsystem, a network of networks - a biologically inspired network is used to recursively perform filtering, segmentation and edge detection; in the intermediate level and the high level, hierarchically structured arrays of self-organizing tree maps - extension of the popular self-organizing map are utilized to carry out image/feature analysis. The system has been applied to solve a number of real world problems. Some interesting and encouraging results will be reported.

1 Introduction

Since the rebirth of neural computing in the middle eighties, there have been heated debates as if neural networks are merely suited only for solving trivial problems. We acknowledge that our limited biological insights in the natural neural networks restrict our ability to closely mimic the biologic system, vision system in particular. However, numerous experimental and theoretical research outputs confirm that if the biological facts are used wisely, neural networks are capable of addressing some large and complex problems effectively [1-3].

This paper presents a computer vision system based on an integrated neural network architecture. In the low level vision subsystem, a network of networks (NoN) [4] - a biologically inspired network is used to perform filtering, segmentation and edge detection recursively; in high level analysis, an array of self-organizing tree maps [5] - extension of the popular self-organizing map and a structured feedforward neural network are utilized to carry out image analysis and pattern recognition.

The motivation behind using the NoN and the SOTM in the design of the vision system is the compatibility of the networks with vision processing. With a modularized hierarchical architecture, the NoN is ideal for parallel and homogeneous processing tasks such as those in low level vision. On the other hand, the SOTM uses a tree structure with minimum nodes to describe the objects in the scene, providing a sensible representation of images for analysis and recognition. A common strength of the NoN and the SOTM is that they are capable of adaptively identifying optimal network architecture for the problem on hand. The identified optimal architecture leads to extremely fast training and superior generalization which are both critical in real world applications.

The system has been used for a number of applications: computer-assisted diagnosis of breast cancers in digital mammograms; biomedical vision to analyze neurological disorders; identification of underwater objects through 3D sonar image processing and visualization, etc. The performance of the system demonstrates that this neural computing approach can provide robust and efficient solutions to the problems on hand. Some interesting and encouraging results will be reported.

2 System Architecture

The proposed vision system consists of two processing stages. In the first stage – low level vision, an recursive processing model based on the NoN integrates filtering, segmentation and edge detection together. The second stage – image representation and analysis, uses an array of the SOTMs to extract features for pattern classification where a structured feedforward neural networks is used. These processing stages will be described in the following subsections.

2.1 Low Level Vision

Although many algorithms are available in the area of low level vision for the extraction of edges [6], segmentation [7] and filtering [8], robustness and efficiency characteristics are still typically lacking in these methods. One possible reason for this is that they do not capture the adaptive and learning characteristics which occur in biological vision. For example, in computational vision. filtering is popularly accepted as the first processing step. However, the seemingly progressive aspect of human perception suggests that there exists interplay between filtering and segmentation/edge detection.

To bring such paradigms into low level vision processing, the biologically inspired neural computing model - NoN is adopted. It proposes that by using clustering method, adaptive filtering and segmentation/edge detection are naturally linked to one another. Based on this concept, the model recursively processes image data to facilitate visualization and high level vision processing.

The NoN incorporates an important biological fact - sparse, hierarchical clustering of neurons in the cortex - into the design of network architecture. The model is then cast into a three-level NoN. The recorded image data is categorized into the first level clusters (homogeneous areas) with each data unit forming one or more zeroth level clusters. The second level connections represent the transition dynamics between the first level clusters. The processing model iteratively identifies network architecture, adaptively filters the image, and perform segmentation/edge detection in an recursive fashion. The model is schematically illustrated in Figure 1. It groups the image pixels (neurons) into clusters, and determines the processing hierarchy. The processing is initialized by estimating the local statistical properties of the input image I_o to produce, in the order of, a network cluster structure C_1, a filtered image I_1, a segmentation map S_1 and an edge graph E_1. Then the recursive procedure starts. At the ith iteration,

a) the structure of the network is updated, $C_{i-1} \rightarrow C_i$, using the segmentation map S_{i-1} and the edge graph E_{i-1} obtained from the previous iteration, via the clustering method described in the next section;

b) the structural information is then sent to the adaptive filter to improve the quality of the filtered image, $I_{i-1} \rightarrow I_i$;

c) based on the enhanced image I_i (and the structural information C_i), the segmenter and the edge detector update the segmentation map, $S_{i-1} \rightarrow S_i$, and the edge graph, $E_{i-1} \rightarrow E_i$, respectively.

The recursive procedure terminates when the optimization criterion adopted is satisfied. The output of the low level vision system: the filtered image I_{out}, the segmentation map S_{out}, and the edge graph E_{out}, are the inputs to visualization, and/or image analysis and feature extraction.

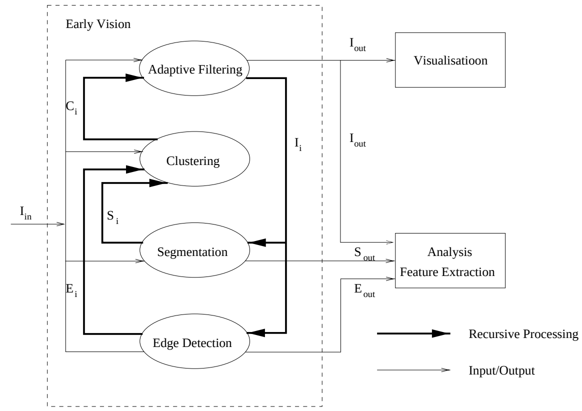

Figure 1: The low level vision model

It is worth noting that the first level cluster representation relates to texture/gray-scale segmentation of the image, and the second level inter-cluster connections represent the transition dynamics between clusters – the edges between homogeneous areas. Thus, the processing model produces adaptive filtering, segmentation and edge detection in a recursive fashion.

2.2 Representation and Analysis

Kohonen's Self-Organizing Map (SOM) is as popular method for prototype generation and clustering in order to organize unlabeled feature vectors into natural clusters in such a way that the entities within a cluster are more similar to each other than those in different clusters. However, the SOM has some undesirable properties. When an input vector distribution has a prominent shape, the results of the best-match computations tend to be concentrated on a fraction of nodes in the map. Therefore, the reference vectors lying in zero-density areas may be affected by input vectors from the surrounding nonzero distribution areas. This may cause statistical instability [9].

Based on the idea of SOM, we proposed a new mechanism called self-organizing tree map in which the relationships between the output nodes can be defined adaptively during learning [5]. The clustering algorithm starts from an isolated node and coalesces the nearest patterns or groups according to a hierarchy control function from the root node to the leaf nodes to form the tree. The SOTM mapping projects an input pattern $x = (x_1 ... x_N) \in R^N$ onto a tree node. With every node i, a weight vector $w_j = [w_{1j}, ... w_{Nj}]^T \in R^N$ is associated. The proposed approach has the advantage of accurately locating cluster centers, and preserving topological relationships. In training of SOTM, within a given period of the hierarchy control function, the weight vectors of the tree nodes converge to the mean of their corresponding input vectors as the learning rate $\alpha(t)$ decreases. The SOTM provides a better and faster approximation of prominently structured density functions.

Learning in the SOTM takes place in two phases: the locating phase and the convergence phase. The adaptation parameter $\alpha(t)$ controls the learning rate which decreases with time as weight vectors approach the cluster centers. During the locating phase, global topological adjustment of the weight vectors w_j takes place. $\alpha(t)$ is maintained relatively large during this phase. A small $\alpha(t)$ for the convergence phase is needed for fine adjustment of the map.

A hierarchy control function $H(t)$ controls the levels of the tree. It begins with a large value and decreases with time. It adaptively partitions the input vector space into smaller subspaces.

With the decreasing of hierarchy control function $H(t)$, a subnode forms a new branch. The evolution process progresses recursively until it reaches the leaf node. The entire tree structure preserves topological relations from the root node to the leaf nodes.

The dynamic SOTM topology is demonstrated in the following example. In Figure 2, the learning of the tree map is driven by sample vectors uniformly distributed in English letter "K" as shown in Figure 2.a). The tree mapping starts from the root node and gradually generates its subnodes as H(t) decreases. Each time as $H(t)$ decreases, $\alpha(t)$ starts from the initial state again. For certain $H(t)$, $\alpha(t)$ decreases with time. By properly controlling the decreasing speed of $\alpha(t)$, the SOTM will find the cluster center just like K-means does. During vector quantization, an N-dimensional vector in Euclidean space is approximated by its closest representative among the finite set of the tree nodes (reference vectors). In the node organizing process, from visualizing the tree map evolution the optimal number of the output nodes can be obtained as shown in Figure 2.b). The SOM is also used to work on this example. as shown in Figure 2.c). Although the SOM's topology exhibits the distribution of the structured input vectors, it also introduces false representations outside the distribution of the input space.

In image analysis and feature extraction, an array of three SOTMs are used to represent the image to be analyzed, one for grey scale representation, one for texture representation, and one for edge graph representation. Based on the characteristics of the images being processed, the features extracted from different SOTMs are properly weighted, according to their relevance with respect to the problem on hand. For example, if the texture significantly characterize the image,

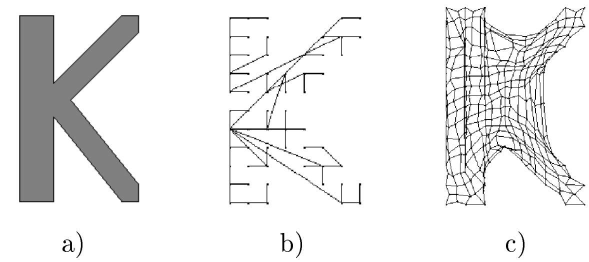

Figure 2: SOTM for representation: a) English letter "K"; b) SOTM representation of "K"; c) SOM representation of "K".

the corresponding features will be scale up. Otherwise, they will be scale down. The weighted features are then fused together by a structured feed-forward neural network (SFFNN). Unlike the other feedforward neural networks, the SFFNN has an transparent structure between the input layer and the hidden layer, and conventional fully connection between the hidden layer and the output layer. The argument for this architecture is as follows. A unary super feature can be formed from the features extracted from the same SOTM. On the other hand, a binary super feature can be formed by the fusion of features from two SOTMs. The six super features are finally fused together by the fully connected second level to produce the final result. This level of processing is schematically illustrated in Figure 3.

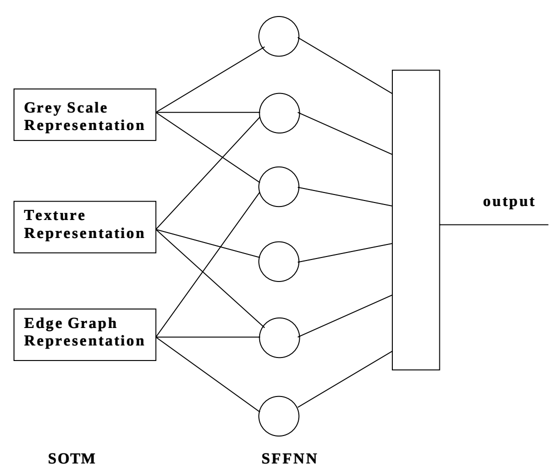

Figure 3: Image representation and Analysis.

3 Experiment

The system has been used for a number of real world applications. In this section, we will use representation of underwater objects through 3D sonar image processing and visualization to demonstrate its performance. An unknown underwater object was detected by a high resolution sonar. Due to the complex imaging conditions, the image was degraded by space-variant degradation as shown in Figure 4.a). Major processing results by our system are shown in Figure 4.b), 4.c) and 4.d), respectively. Figure 4.b) and 4.c) shows the recursive filtering and the segmentation, and Figure 4.d) is the final description based on texture and grey scale representations. Edge representation is of little use due to the nature of the image.

4 Conclusions and Future Work

We present a vision system based on a neural computing architecture. The low level vision part recursively integrates adaptive filtering, segmentation and edge detection. The high level part uses an array of self-organizing tree maps to represent the image for feature extraction, and a structured feed-forward neural network to fuse the features to produce pattern classification. The system has been used in a number of real world applications. A sonar image processing example is given to demonstrate the performance of the system. Through the research and development activities of this work, we conclude that, by incorporating biological facts in architecture design, neural networks can be used to resolve large scale and complex problems.

References

1 S.Y. Kung, *Digital Neural Networks*, Prentice Hall, 1993.

2 T.M. Caelli, D.McG. Squire and T.P.J. Wild, "Model-based neural networks," *Neural Networks*, vol. 6, pp. 613-625, 1993.

3 L. Guan, J. Anderson and J. Sutton, "A network of networks model for image regularization," *IEEE Trans. on Neural Networks*, vol. 8, no. 1, pp. 169-174, Jan, 1997.

4 J.A. Anderson and J.P. Sutton. "A network of networks: Computation and neurobiology," *World Congress of Neural Networks*, vol. 1, pp. 561-568, 1995.

5 H. Kong and L. Guan, "A self-organizing tree map for eliminating impulse noise with random intensity distributions," to appear in *J. of Electronic Imaging*, vol. 6, no. 1, 1998.

6 S.W. Lu and A. Szeto, "Hierarchical artificial neural networks for edge enhancement," *Pattern Recognition*, vol. 26, no. 8, pp. 1149-1163, 1993.

7 T. Wang, X. Zhuang, and X. Xing, "Robust segmentation of noisy images using a neural network model," *Image and Vision Computing*, vol. 10, no. 4, pp. 233-240, May, 1992

8 L. Guan, J. Anderson and J. Sutton, "A network of network model for image regularization," *IEEE Trans. on Neural Networks*, vol. 8, no. 1, pp. 169-174, Jan, 1997.

9 T. Kohonen, *Self-Organizing Maps*, Springer, Berlin, 1995.

10 W.K. Pratt, *Digital Image Processing*, John Wiley & Sons, New York, 2nd Edition, 1991.

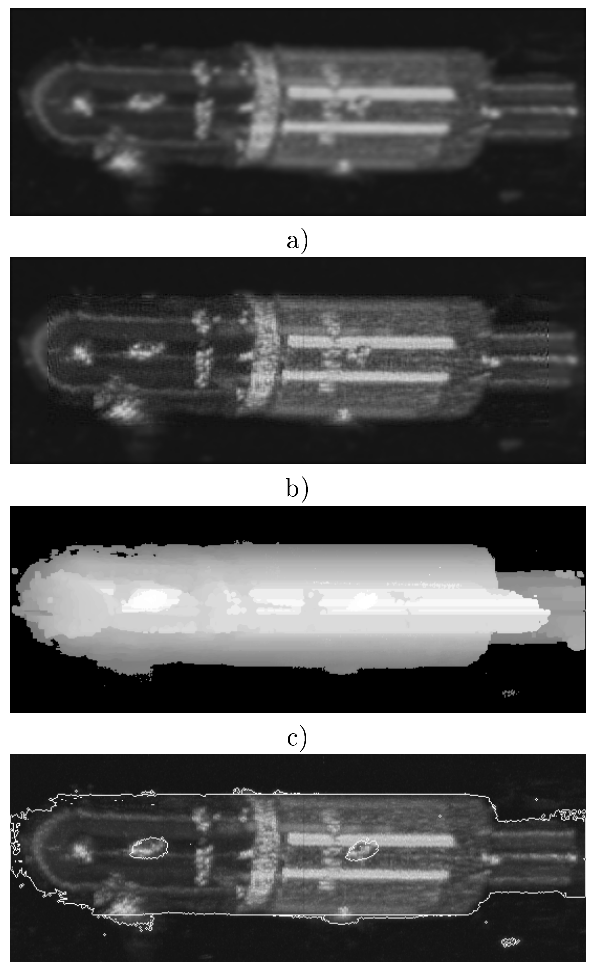
a) b) c) d)

Figure 4: Processing of a underwater image: a) The recorded "mine" image; b) filtered a); c) segmentation of a); d) description of the image.