

Neural Network for Region Detection

Giorgio Cucurachi , Guido Tascini , Francesco Piazza

Istituto di Informatica - Dipartimento di Elettronica e Automatica
Università di Ancona
Via Breccie Bianche , 60131 Ancona (Italy)
E-Mail : tascini@inform.unian.it

Abstract : The paper proposes a neural network organized in three structures , each of which is constituted by a set of levels . The lower structure is made up of two layer groups : the first one filters the high frequency noise , while the second one is sensitive to scarcely lighted images . Finally the third structure detects contour and position of regions . The network uses neurons of C , S and V type in analogy to Fukushima Neo-Cognitron . A simulation program has been implemented , which shows good throughput in spite of network complexity .

Keywords:Neo-Cognitron,Pre-Processing,Filtering,Vectorial Quantization

1 Introduction

Neo-Cognitron is a multilayered neural network particularly suitable for image processing [8],[9] . It has been developed by Fukushima , taking inspiration from natural visual neural structure [6] . It is able to operate a classification of the objects in the image [8] , performing a features extraction and their relative linking in a manner more and more complex as features pass from a layer to another one . During the learning phase some cell-plans are generated : each of them extracts a specif feature , in order to recognize all the possible variations found out in the learning period . Cell-plans widening represents a critical point for Neo-Cognitron : in fact the increase of width does not permit investigation of high number of prototypes and reduces actually Neo-Cognitron application to hand-written characters recognition .

Image processing implies large dimensions matrices , that Neo-Cognitron is not able to manipulate : Fukushima structure operates efficiently with 19x19 pixels images but presents serious convergence troubles with larger matrices . It is necessary to operate pixels number reduction to investigate . It has to manipulate image , i.e. with a pre-processing phase , to decrease matrices dimensions , preserving image informative content .

Many filtering and edge-detection techniques [4] exist . Our idea is to realize Pre-Processing with a neural structure , inspired to biological models . Compared to traditional tecniques , neural model shows the advantage of a high parallelism , that allows remarkable computational performances and the possibility of managing larger images (see Fig.1) . Moreover the analogy with the biological structures exhibits high compability with the Fukushima's model , confirming the neuro-biological reading-key of the Neo-Cognitron . Toponomastic, graphic aspects and formalism of Neo-Cognitron have been preserved . It has been used a new kind of cell , called S-Cell Type B , to differentiate it from that of A-Type commonly used by Fukushima .

Obviously the cells that control the S-Cell Type B are called V-Cells Type B . The parallellism concept has been increased , making a neural structure with a lot of layers . Each layer is characterized by particular functions . The Pre-Processing structure is composed by two parallel ways that like natural visual system [2] , find out image different aspects (see Fig.2) .

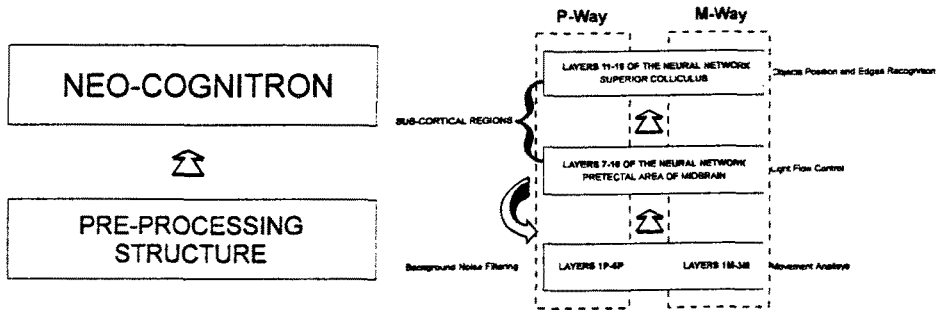


Fig.1:Image is first processed by Pre-Processing Structure and then by NeoCognitron

Fig.2:Scheme of the Pre-Processing structure

The P way [2] , in analogy with the biological Parvocellular System , attends to pattern recognition with a high luminosity condition : it permits a detailed vision under an aspect both morphologic and cromatic of objects in the visual field . The M way , in analogy with the Magnocellular System attends to shape recognition in a low signal-noise rate situation : it has an important rule in the objects moving analysis[2]. Both the informative ways can be divided into two blocks , that show both functional and structural different characteristics :

- 1) The network first layers , in analogy with Retina cells , make a filtering operation , by cleaning image from noise . Layers 1P-6P clean high frequency noise . Layers 1M-3M are able to distinguish , in low lighting situation , objects from the background .
- 2) Layers 11-18 find out the edges and the centres of the objects in the scene . Attention is then pointed on edge , that has got a higher informative content , while the internal part is reduced to a single point , in correspondance with the figure centre . This is useful in the analysis of objects within the visual field .

The action of these two components is controlled by a photosensory system that , according to lighting level, activates the P way or the M way . Layers 7-10 find out objects light intensity in relation to the background and in a similar manner to the Praetector Area of Midbrain , control the two ways activity .

Layers 1P-6P , 1M-3M , 11-18 (and 7-10 that have a supervisional rule) make a Pre-Processing structure : as we have already told , it does a filtering action , then identify the edges and the centres of objects in the scene . So a simplified image is

produced where point number , is considerably lower than in the original image . In other words , we obtained a Vectoral Quantization action , that simplifies greatly the recognition process .

Images so treated are passed to multilayer neural structure , similar to the Fukushima's Neo-Cognitron [9] , that identifies edges corners and straight lines . The recognition of these features allows objects clustering , after a training period with a supervised learning .

Neural Networks are suitable to this project : they were born in fact with the willing of riproducing natural neural structures and have been developed paralleling to neurobiological discovers [1] .

Neural Networks application to patterns recognition problem shows two advantages :

- 1) High parallelism level ;
- 2) Experience improving skill .

The first aspect permits a high computational speed and gives a reading-key to a hardware realization . The second factor represents a big advantage in objects clustering process .

2. Self-adapting neural structure

The network is formed by two parallel structures : the first one filters the high frequency noise , while the second one is sensitive to scarcely lighted images . A further neural structure realizes a feedback control mechanism , where brightness value is used to regulate sample window opening level . So the neural structure is able to adjust itself according to light intensity .

2.1 Background noise filtering (Parvocellular System)

Neural model shown in Fig.3 presents a functional analogy to Retina Parvocellular System cells [2] .

First layer C-Cells corresponds to Retina Photoreceptors . They get the grayscale of each sampled image pixel . For each first layer cell there is a Type A S-Cell group that controls the eight pixels of the examined cell boundary .

Each first layer cell is compared to the near ones by a Type A V-Cells set , that allows costant signal relay from second layer C-Cell to the third layer Type A S-Cells.

So , for every first layer pixel , is activated a number of third layer S-Cells corresponding to the number of cells of the examined cell boundary which have a gray level similar to that of the examined cell . Since signal relayed from third level cells has an unitary value , the fourth layer C-Cells , making their sum , go to identify the number of cells of the examined cell boundary that are in corrspondence with the central cell.

Fourth layer cells have an activation threshold . Hence they allow to generate a signal output only if the threshold value is exceeded. The threshold value is fixed to two . This value has been determined empirically , in accordance with the results obtained by a neural structure simulation program . It represents a good compromise between

filtering capabilities and selectivity in thin patterns research . In other words C-Cell bursts only if the examined cell has at least two contiguous cells with a similar gray level . The signal of the C-Cell excites at last a Type B V-Cell that enables signal trasmission coming from the first layer .

In this way only the signal that shows a not-casual relation with that of bordering cells is passed to the higher layers . On the contrary , background noise has a casual distribution and a large amount of it is eliminated by the neural structure .

This background noise filtering tecnique permits to reduce the computational weight of the neural structure , allowing it to investigate only the more significative pixels . This procedure is useful only in a high luminosity situation . Parvocellular System in fact realizes the diurnal vision , where it is able to make an image detailed analysis , showing a great sensibility to shapes and to the grayscale .

2.2 Magnocellular System

In a low signal-noise ratio condition it is more important to determine objects shapes rather than their details or their exact grayscale . It is necessary to define a new type of artificial neural structure , different from the previous one that operated in high luminosity situation . Like in the biological Magnocellular System , the M-way cells (see Fig.4) are located in the peripheral area of the visual field [2]. They are activated by the 7-10 layers of the neural structure that will be described in the following paragraph .Like the pupil in the natural eye, this system controls the light flow , by widening the sample window when the luminosity is low and by contracting it when the luminosity grows .

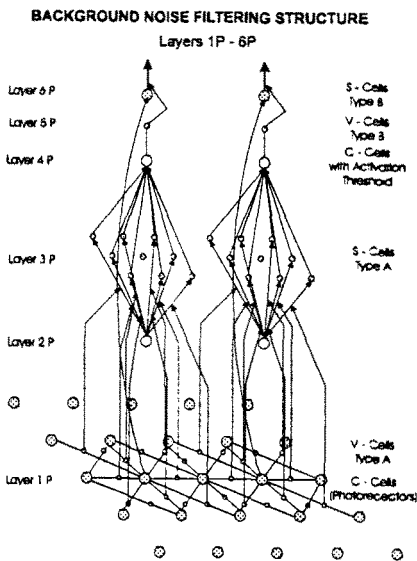


Fig.3:Background filtering structure (Layers 1P-6P)

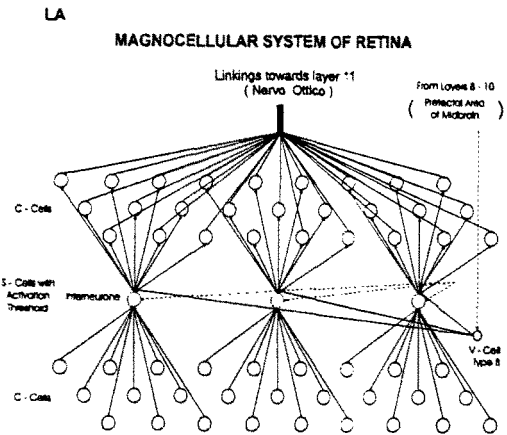


Fig.4:Layers 1M-3M of the Artificial Neural Network

The C-Cells , in the 8-10 layers of the Pretectal Area simulating structure , determine the averaged brightness , considering mainly the visual field peripheral area , because of the photosensories distribution . S-cells of the 1M layer reproduce Magnocellular System behaviour , calculating the average value of the photoreceptors signals that they control . This mean value will be relayed from Interneurons to the 3M layer cells only if it exceeds the computed threshold value .

Thanks to this architecture will be enphatized images that , even if hardly lightly and scrambled , have signals stronger than background ones . Because of the convergence of synaptic endings of each area cells in a single S-cell , the night vision allows only objects shapes and general properties recognition [5] .

Since every region , in which the visual field is divided by the M-System , is invariant to little signal oscillations related to background noise , is quite evident the M-System importance in movement analisys . A big luminosity variation of such a region is often related to the movement of the object that previously occupied that area . In case of high enviroment luminosity , control system (layers 7-10) automatically disables M-Way and assigns visual functions to P-Way that , like biological Fovea , operates a high resolution images analisys .

2.3 Sample-window control system

Photosensors system , like that described in Fig.5 , determines the whole image average brightness. Photosensors are distributed so that such a value is influenced principally from the peripheral area . Praetectal Area C-Cells get informations coming from photosensors , collocated in circular regions farther and farther from the visual field centre . C-Cells have an Activation Threshold higher and higher as it is passed from border to centre . These cells activation produces sample window contractness . In this way more luminosity is high and more the sample window opening is reduced. Signals leaving C-Cells arrive to Type B V-Cells , that inhibit Magnocellular System cells . M-Way 2M layer Interneurons blocking represents the control action of this neural structure . Such an archititue realizes a feedback control mechanism , where brightness value is used to regulate the sample window opening level . The sample window can scan gradually the whole image surface , or can be used a selective attention tecnique (based on the neural structure described in paragraph 2.4) that reduces the complexity and that , for the sake of brevity , we will not describe furtherly .

Analogally to what happens in Retina, P-System cells are concentrated in the central area of the window and M-System ones are distributed in the peripheral area [2] . So P-way is particularly suitable to details recognition while M-Way to nightly vision and movement analysis . The type M cell group , collocated in the visual field central area , continues to be active even when enviroment luminosity achieves high values . It is evident that in these conditions they don't give any contribute to vision , in comparison with that we already know through P-Way . Anyway these cells are important in movement analysis. A big luminosity variation of a certain region in fact is often related to the movement of the object that previously occupied that area [5] .

**LAYERS 7 - 10 OF THE ARTIFICIAL NEURAL NETWORK
PRETECTAL AREA OF MIDBRAIN**

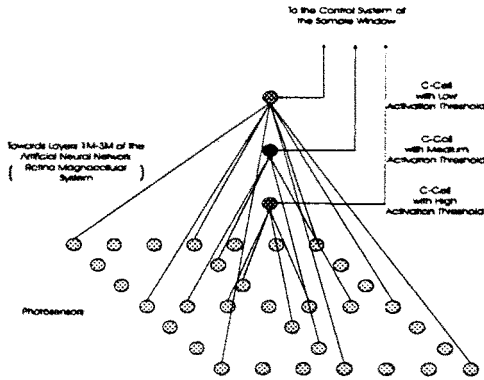


Fig.5: The Sample Window opening is regulated by the Neural Structure that simulates the Pretectal Area of Midbrain

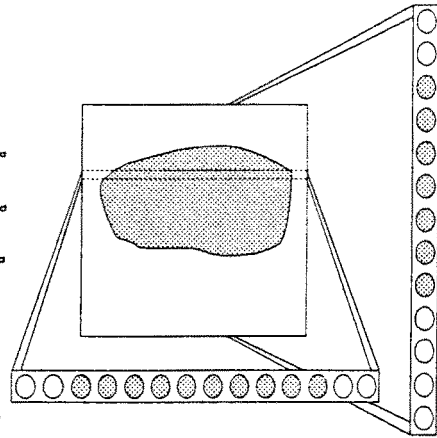


Fig.6: Layers 11-18 (Superior Colliculus) operate on parallel and independent sections

3 Artificial Neural Structure for Region and Centre detection

This neural structure, like Superior Colliculus, has an important rule in the recognition of edges and centres of objects in the scene [2]. The neural architecture (Fig.7) is composed of an array of neural structures that work, in an independent manner, on single sections, both horizontal and vertical, into which we can divide the whole visual field. In Fig.6 it has been evidenced the technique, used by the artificial neural structure in order to extract the image features. It is easy to see the cells that participate to the recognition process. In any of the two sections, randomly chosen, we can notice the division among the homogeneous regions that constitute the image. Each section is made up of a 3D neural structure, composed of eight cells layers, that we can divide functionally in four blocks. First layer C-Cells receive afferences from 6P and 3P layers (like the optic nerve): the signal that they receive has already been filtered from background noise, thanks to P-Way and M-Way actions. Each C-Cells couple relays signal to a Type B V-Cell, that compares the grayscale of the two nearby cells. Inputs to V-Cells have opposite synaptic weight. So V-Cells burst if they feel different grayscale values of the input-layer cells. The third layer S-Cells receive signals that have informations only about the spatial position of the visual field area that they are examining. The third layer S-Cells relay signals to fifth layer S-Cells if they are enabled by the second layer V-Cells. Hence only the S-Cells that are in correspondence with objects edges will be abilitated.

In order to find out the centre of every area in which the section is divided, the artificial neural network converges signals, that come from boundaries, towards the central point of each homogeneous range. The fifth level S-Cells are composed by a

couple of cells that are controlled by two Type A V-Cells of the fourth layer . These cells control the grayscale of the input-layer cells . So we have the right S-Cell activation or the left one , according to the examined point is the left or the right ending of the examined region . In other words the network control structures enable only the signals that are going to the central point of each omogeneous region into which the section is divided . The synaptic linkings between the fifth layer S-Cells and the sixth level Type A V-Cells have weight 0.5 .

The V-Cells compare their sum with the signal coming from an external source that identifies the spatial position of the examined cell . The V-Cell activation shows that we are in corispondence with the central point of the examined region : The transaction of the grayscale from the input-layer to the eight level S-Cell is then allowed .

LAYERS 11-18 OF THE NEURAL NETWORK SUPERIOR COLLICULUS

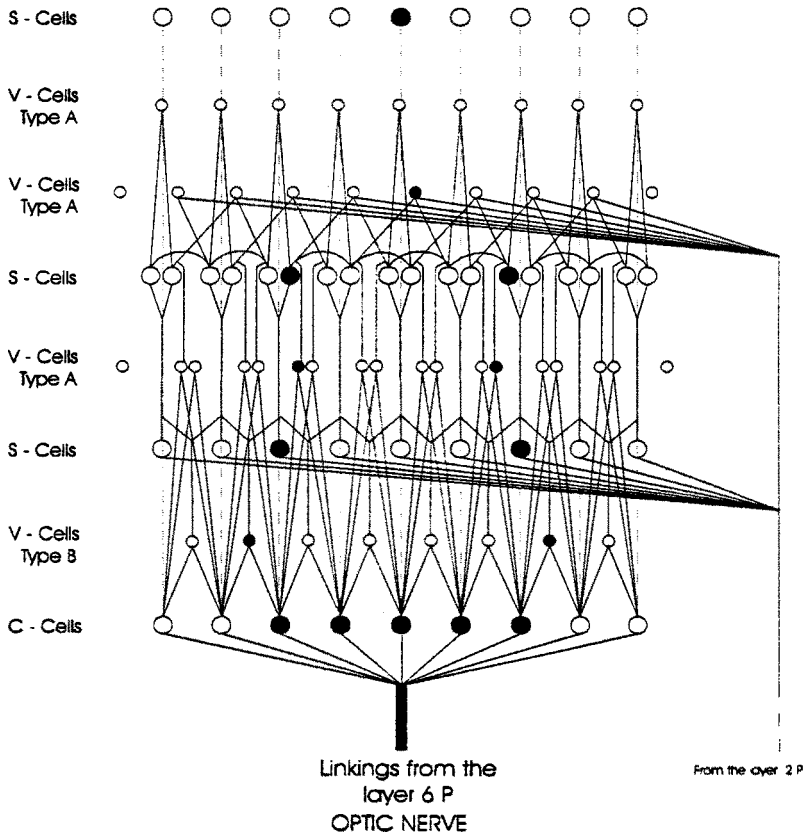


Fig.7:Architecture of the layers 11-18

The same afferences of the sixth level Type A V-Cells are directed even to Type B V-Cells that enable the transaction of the signal of the fifth level S-Cell towards the near S-Cell . In this way , the signal of the edge S-Cells is made converging towards Central S-Cells , until sixth level Type A V-Cells point at the final accomplishment of the region central point . Seventh layer V-Cells burst if both the fifth level S-Cells are active . This happens if the examined object has unitary width . These V-Cells detect very thin patterns . Then they communicate it to upper visual structures that investigate separately this particular type of visual structure . At last the artificial neural structure detects the presence of objects within the visual field , by recognizing the edge and the geometric centres .

3.1 We analyze through an example the action of layers 11-18 on visual signals

We consider an image with dimensions 300x300 and we observe the action of the layers 11-18 on it . The artificial structure , that simulates retina , sends an image already filtered from noise . We can recognize in Fig.8A a square , a rectangle , a circle and a pentagon . Objects are randomly positioned in the space and can have a whatever grayscale value .

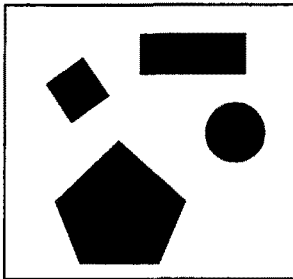


Fig.8A : Image at the input layer of the layers 11-18

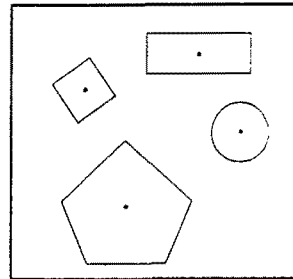


Fig. 8B : The artificial neural structure finds out edges and centres

Neural Network , by simulating the sub-cortical nuclei behaviour , identifies edges and geometric centres of objects .

3.2 Example of movement analysis

A four images sequence is shown in Fig.9 . Each frame identifies object position in a specific instant. We can see that :

- 6P layer (parvicellular system) filters high frequency background noise ,
- 13 and 18 layers (superior colliculus) detect object border and centre .

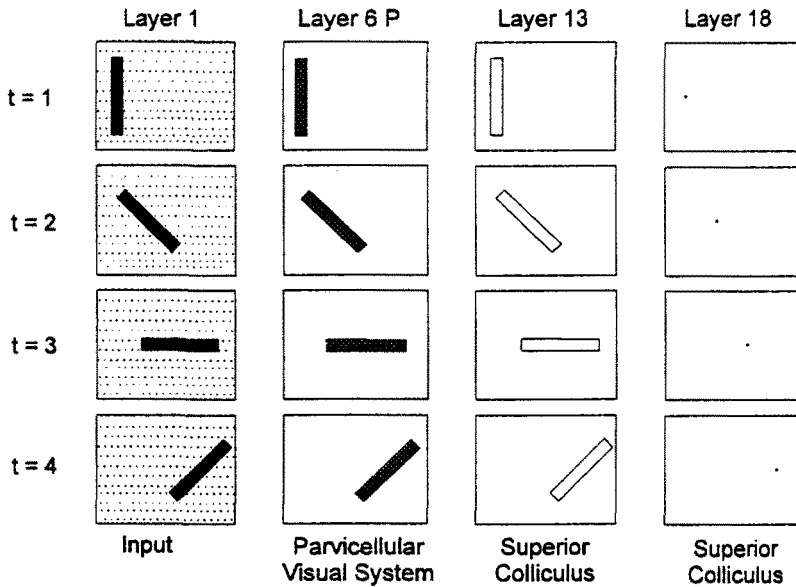


Fig.9 : The layers of the neural network identify object contour and centre during its movement

4 Conclusions

The Biologically-inspired Pre-Processing Neural Structure presents a high level of parallelism, with a total distribution of elaboration, control and data.

The network architecture has been configured like a Parallel Distributed Processing System [3]. We can see the structure, forgetting the neurobiologic theories that have inspired the design, like a data elaborating system, that is able to offer high computational performances. The whole network architecture shows a vertical development in Hypercolumn. So signals don't diffuse themselves randomly but follow parallel routes that pass through all the layers.

The neural structure is able to manipulate high dimensions matrices. Because of the highly parallel formulation, this kind of neural architecture can manage data in a local environment. The Neural Structure Simulation Program, that we have realized, can process images with dimensions of 400x400. It is quite an important result, since it allows to get over the limitations of about all neural networks. In fact traditional neural networks hardly manipulate matrices with dimensions over 128x128. The advantage of this structure application to Neo-Cognitron is quite evident: it permits to applicate Fukushima's model not only to characters recognition but also to wide dimensions images processing. This network architecture needs limited hardware resources. Because of its 3D development, this neural architecture divides the data set into small groups that manages independently, through parallel ways.

This Pre-Processing structure is completely independent from Translation, Scale and Rotation of objects within the visual field. Moreover it is quite insensible to noise

and to image dimensions . The neural structure simulation program is simple to be managed . It is written in Pascal 7 and can work on a common PC . It ensures high performance , in terms of computational speed and image complexity . For example , we can say that a Pentium based PC operates the Pre-Processing procedure on a 300x300 image in about 3 seconds .

Bibliography

- [1] S.Zornetzer,J.Davis,C.Lau : An Introduction to Neural and Electronic Networks. Academic Press Inc. (1991)
- [2] E.Kandel,J.Schwartz,T.Jessel : Principles of Neural Science . Prentice Hall (1991)
- [3] J.McClelland,D.Rumelhart and the PDP Research Group : Parallel Distributed Processing . Explorations in the microstructure of cognition Vol.2 Psychological and Biological Models . MIT Press (1986)
- [4] Y. Pao : Adaptative Pattern Recognition and Neural Networks Reading (Mass.) . Addison Wesley Publishing Co. (1989)
- [5] D.Hubel : Eye , Brain and Vision . Scientific American Library (1988)
- [6] K.Fukushima : Neo-Cognitron : a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position . Biological Cybernetics 36 : 193-202 (1980)
- [7] K.Fukushima : Neo-Cognitron : a neural network model for a mechanism of visual pattern recognition . IEEE Transaction on Systems , Man and Cybernetics Vol.13 N.5 (1983)
- [8] K.Fukushima : Neo-Cognitron : a neural network model for selective attention in visual pattern recognition . Biological Cybernetics 55 : 5-15 (1986)
- [9] K.Fukushima : Neo-Cognitron : a hierarchical neural network capable of visual pattern recognition . Neural Networks,Vol.1 119-130 (1988)
- [10] W.K.Pratt : Digital Image Processing , Wiley , New York (1991)