USING MULTI-KOHONEN SELF-ORGANIZING MAPS FOR MODELING VISUAL PERCEPTION

Michel Collobert

France Telecom R&D IRIS/VIA, F-22307 Lannion Cedex, France ---michel.collobert@orange-ft.com

Abstract:

This paper presents a new organizing principle for perceptual systems based on multiple Kohonen self organizing maps. These maps are arranged in order to model the global brain activity as seen on tomography pictures. In contrast to most of neural network models, a perceived object or knowledge is not represented by the final activity of a single neuron but by a <u>configuration of activity</u> of the whole neurons' set.

Key words: Distributed perception, configuration of activity, Kohonen maps, somatosensory areas, neuron, chaos, multisensoriality, grandmother neuron.

1. INTRODUCTION

In our human conscience, each representation of the external world (objects, knowledge, etc..) is automatically translated into one word, i.e. into *one* symbol. Because of this, researchers in artificial intelligence (AI, including perception, cognition and generation of action), tend to construct their algorithms in order to obtain a single code as 'output'. In the field of « neural networks », such as the well-known Kohonen's map [7] or Multi-Layers Perceptrons [9], or in the field of Machine Learning [3] the process is similar. Only one of the elements used for coding the representation (or at most a few) are active at the end of the process.

However, in living organisms with efficient and adaptive perceptual systems, the only important outputs are actually 'actions'. Thanks to new *K. Wojciechowski et al. (eds.), Computer Vision and Graphics*, 1031–1036. © 2006 *Springer*.

2 Michel Collobert

imaging techniques (such as PET or fMRI scans) we realize that in such a system a substantial portion of the neurons is activated at any given time (Fig.1). There is never only one neuron or one little neuron group (such as in the grandmother neuron concept) firing. Instead, a lot of neurons from different brain areas are always firing.

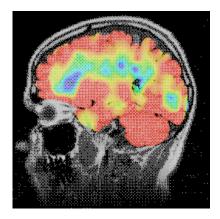


Figure 1. Brain activity of Dr. Christof Koch with his permission.

It is a great challenge to the scientific community (and has been for a long time), in particular in the A.I. and robotics communities, to understand and study the functionality of the perceptual systems. Even modeling the nervous system of a simple housefly [5] would be a great success. Some researchers such as [2] even suggest that we are missing something fundamental in the currently available models.

The new paradigm presented in this paper aims at filling a small part of this void. The paradigm consists in constructing a model able to perform perception based on the simultaneous activity of several assemblies of neurons.

2. DESCRIPTION OF THE MODEL

The model is based on some of the principles observed in biological nervous systems :

- Specialization of cerebral areas .
- Somatotopic maps.
- Hebbian Learning [6].
- Nonlinear and Chaos dynamics. [4-8]

This section first presents the principles of representation and memorizing. The second part describes the principles of recognition/perception.

2.1 Representation.

The first two characteristics of biological nervous systems are used for the representation of input data. They are as follows:

- Specialization of cerebral areas. There is a functional specialization of the different parts of a nervous system. For instance, in human beings, the brain is divided in many parts. One of them is the 'visual cortex' which is itself divided in several areas named MT, V1, V2, V3 etc.. Each of these area is specialized in a particular function, as processing speed, or edges, or colour or the position or..
- Somatotopic maps: A large part of these brain areas are somatotopic, i.e.
 they performs a topology-preserving mapping of sensory organs onto the
 cortex. This mapping is not linear, i.e. most the brain has to discriminate
 between inputs from the sense organs, most it use a great number of
 neurons.

In order to simulate this, the new model presented here uses "Multiple" "Kohonen Self Organizing Maps" (KSOM) [7]. "Multiple" to address specialization and KSOMs to address somatotopic properties.

For example, to discriminate between different objects in a video stream using this model, it is possible to use one specialized KSOM per attribute used to represent these objects. Then, each object or type of object will be represented and memorized by one activity configuration of the whole KSOM. The number of objects between which the system is able to discriminate follows a geometric series, function of the number of neurons per map and the number of attributes/maps. So if we use for example only three mono-dimensional KSOM of 10 neurons each (one for color, one for height, one for width), one is able, in theory, to discriminate between 10³ objects with only 30 'neurons' (Fig. 2).

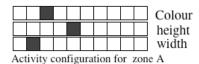




Figure 2. Examples of activity configuration.

4 Michel Collobert

Each KSOM is trained with the different values of the associated attribute for all objects. Because an important feature of KSOM is its non-linearity, we are sure to discriminate as well as possible between objects using this attribute. For example if the set of objects to identify/discriminate is composed of faces, the KSOM associated with the color attribute is composed only with neurons responding each at a particular hue of skin.

Another important feature of a KSOM is the somatotopic property. So in the case when the value of an attribute varies for an object between two instants, the two corresponding configurations of activity will be very close.

In addition, memorizing, not a particular state of one object at a particular time, but its prototype, can be done by learning the statistics of its representation. The more a particular value of an attribute is observed, the more the corresponding neuron is active. So we obtain the following configuration of activity of this kind (Fig. 3):



Fig. 3. Example of activity configuration of KSOM

Note that this representation presents similarities with parts of brain images obtained with scans, such as in fig. 4.

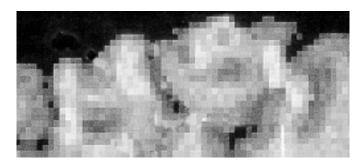


Figure 4. Example of activity configuration of brain maps

2.2 Memorizing

To memorizing these configurations of activity, we use the well know Hebbian Learning: "When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some grown process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased "[6].

We suppose there is an existing connection between each neuron of each map and we apply this Hebbian rule during learning with normalization. Unlike to the number of items the system is able to memorize, the number of connections grows according to an arithmetic series law.

2.3 Recognizing / Perception

The brain is known for its chaotic behavior [4] [8]. Because of this chaotic character, the brain can be seen as a huge system of oscillators, each oscillator moving from one memorized configuration of activity to another in the absence of external stimuli. A new configuration of activity resulting from sensorial inputs can force the system to be stabilized in a previously memorized state. This stabilization can be called recognition/perception. This is consistent with the findings of many psychologists which attest that the brain is stabilized by the external world (problem of sensory deprivation). Some neurobiologists even say that the brain is trying to construct a coherence when faced by the external world[1]. Unfortunately mathematics which handle chaotic systems are of no help on this.

However if we use another property of cortical areas, namely the diffusion of the excitation between neurons, we can resolve this problem. Recognition then simply becomes a matter of searching the nearest configuration of activity:

- 1. If the input configuration to recognize is strictly identical to one memorized, the problem is solved.
- 2. Otherwise, a pairing of the input configuration to one of the previously memorized configurations needs to be performed. Let us suppose that we have to pair the input configuration of fig. 5 with one of the two previously memorized configurations of fig. 2:



Figure 5. Input configuration

To carry out pairing in this case, the principle of diffusion of the activity towards the adjacent neurons, (preferably with attenuation) can be used:

6 Michel Collobert



Figure 6. Example of the first stage of diffusion

By superimposing this new configuration onto the two initial configurations 'A' and 'B', it appears that it is the 'B' one which "agrees" best to the required form. The input form and the stored form 'B' enter in "resonance" while the form 'A' is inhibited. Thanks to the KSOM nonlinear property, a step of diffusion is normalized for each map' space. So a step of diffusion is equivalent for each attribute.

3. CONCLUSION

The system has been tested on a relatively simple problem of tracking object in a video flow. However, this model is still more theoretical than operational at the moment. First experiments in object tracking show that the model has some promising properties. These properties are due to its 'biological-like' features:

- The approach is compatible with true multi-sensorialities and data fusion: This is because there is no a priori problem in mixing different maps from different sensor modalities. For example, the paradigm could be used for speech recognition using both sound and lips movements.
- The model could be tuned in order to operate (to resonate), even if there is missing data. Such a feature may be useful in vision to treat occlusion problems. But it may be also useful to simulate logical inference if we see this problem as an 'tuning' problem between facts.
- In a hardware implementation of this model, no one single neuron is important by itself (there is no 'grandmother cell'). So a chip design of this model would be more resistant to local failure than more traditional models. In our model, the failure of one component is not catastrophic for the system and the sole consequence is a decrease in the quality of the discrimination.

References

- [1] Berthoz A. Le sens du mouvement O. Jacob 1997 (The brain's sense of movement . Harvard University Press 2000)
- [2] Brooks R., The relationship between matter and life, Nature (2001), vol 409.
- [3] Collobert R., Large Scale Machine Learning, Thesis Paris VI, (2004)

- [4] Faure P., Korn H., Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. C. R. Acad. Sci. Paris, Ser. III 324 (2001) 773-793
- [5] Franceschini N, Pichon J.M. and Blanes C., From insect vision to robot vision, Phil. Trans. R. Soc. Lond. B (1992) 337, 283-294
- [6] Hebb D.O., The Organization of Behavior: A Neurophysiological Model, Wiley, New York 1949.
- [7] Kohonen T., Self-Organizing Maps, Springer, Berlin, 1995
- [8] Korn H., Faure P., Is there chaos in the brain? II. Experimental evidence and related models, C. R. Biologies 326, 787-840 (2003) Elsevier
- [9] Zhang G.P. Neural Networks for Classification : A Survey, IEEE Trans on Systems, Man and Cybernetics Part C, vol 30, No 4, nov 2000, pp 451-462