Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Object Segmentation and Reconstruction via an Oscillatory Neural Network: Interaction among Learning, Memory, Topological Organization and γ-band Synchronization

E. Magosso, C. Cuppini and M. Ursino
Department of Electronics, Computer Science and Systems, University of Bologna, Cesena, Italy
emagosso@deis.unibo.it

Abstract—**Synchronization of neuronal activity in the γ-band has been shown to play an important role in higher cognitive functions, by grouping together the necessary information in different cortical areas to achieve a coherent perception. In the present work, we used a neural network of Wilson-Cowan oscillators to analyze the problem of binding and segmentation of high-level objects. Binding is achieved by implementing in the network the similarity and prior knowledge Gestalt rules. Similarity law is realized via topological maps within the network. Prior knowledge originates by means of a hebbian rule of synaptic change; objects are memorized in the network with different strengths. Segmentation is realized via a global inhibitor which allows desynchronitation among multiple objects avoiding interference. Simulation results performed with a 40x40 neural grid, using three simultaneous input objects, show that the network is able to recognize and segment objects in several different conditions (different degrees of incompleteness or distortion of input patterns), exhibiting the higher reconstruction performances the higher the strength of object memory. The presented model represents an integrated approach for investigating the relationships among learning, memory, topological organization and γ-band synchronization.**

I. Introduction

HUMAN perception shows the capability of simultaneously combining various sensory inputs (such as size, color, smell, taste, etc), which involve different cortical areas, to form a coherent high-level representation of an object (binding). Of course, attributes of different objects must be separated and independently processed (segmentation) to avoid erroneous grouping and confusion. The question still remains open on how the brain may solve these problems, integrating highly parallel and distributed information into a unified, consistent percept. A recent hypothesis proposes that binding and segmentation may be solved in the temporal domain, through synchronization of neuronal oscillatory activity in the γ range (40-60 Hz) [1]-[7]. In such scheme, one perceptual object is represented by a synchronous group of oscillators while different objects are represented by desynchronized groups; thus, within a period of oscillation, multiple patterns alternate .

With regard to the plausibility of this theory, a fundamental question has to be assessed: what is the mechanism which generates synchronism among oscillatory activities of different neuronal groups? A possibility is that synchronization among oscillatory units may arise from synaptic connections reflecting Gestalt principles [4], [5]. Gestalt laws [8] may be of low-level, corresponding to local and geometrical properties of sensory stimuli (such as proximity, continuity, common fate, symmetry) or of high-level, involving semantic and cognitive interpretation of objects, such as similarity and prior knowledge.

It is amply documented that prior knowledge strongly affects the ability of human subjects to organize and segment a scene [9]. Synaptic connections encoding memory may originate from Hebbian learning rules, in which synaptic weights change in presence of conjunctive activation of the pre-synaptic and post-synaptic neurons. The similarity principle states that, all else being equal, the most similar elements (in color, size, taste, tone, etc.) are grouped together [9]. This principle may reflect the existence of topological maps in the brain, in which proximal neurons signal similar values for the features and connections among neurons are characterized by short-range excitation and long-range inhibition [10].

Several models of oscillating neural networks, expressing sensory segmentation as signal synchronization, have been presented in past years [1], [11]-[17]. In these models, the rules used for segmentation are inscribed in the synaptic connections linking the oscillators. Most of these models utilize low-level Gestalt rules to segment a visual scene at an early processing stage. On the contrary, only few works rely on high level Gestalt rules for representation of objects at higher mental levels and for their storing in memory.

In addition to the segmentation problem, neural modeling faces with the problem of recalling previously memorized objects even in the presence of partial and/or shifted input information [15]. In classical associative memories, such as the Hopfield net, a previously memorized object can be recalled perfectly starting from a fragment of it, but recognition falls in case of a shift in the input properties.

Here, we present a two-dimensional network of oscillating neurons which realizes segmentation and recognition of high-level objects, by combining γ-band synchronization with the similarity and prior knowledge Gestalt rules. First, the network learns objects according to a Hebbian rule; objects are memorized with different synaptic strengths to simulate the presence in memory of both familiar and unusual objects or objects with different emotional impacts. Then, the ability of the network to segment and recognize the previously memorized objects (starting from partial and shifted input information) is analyzed during the retrieval phase in relation to the different force of their memorization.

1-4244-0033-3/06/$20.00 ©2006 IEEE.

II. Model Description

The network (Fig. 1) is composed of a grid of 40x40 Wilson-Cowan oscillators, subdivided into 4 distinct cortical areas. Each single oscillator consists of a feedback loop between an excitatory unit and an inhibitory unit, while the output of the network is the activity of all excitatory units. Each neural group may be silent, if it does not receive enough excitation or may oscillate in the γ frequency band if sufficiently excited. Excitation of each neuron depends on both the external input and on the connections from other neurons. In the present study, the input to each neuron is described as a scalar quantity, ranging between 0 and 1.

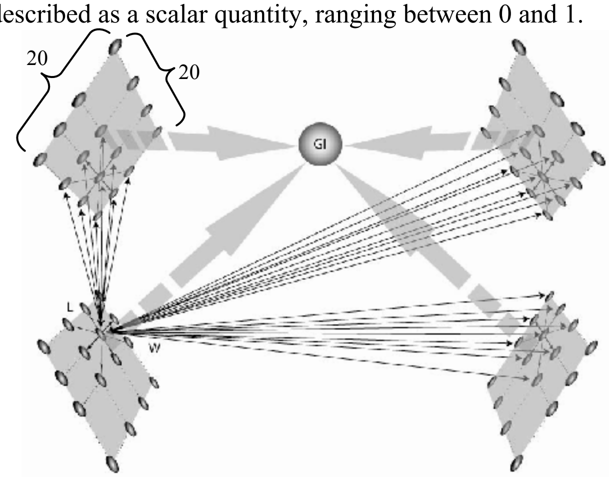

Fig. 1. Schematic diagram describing the network structure. GI: Global Inhibitor; L: lateral intra-area synapses; W: inter-area synapses.

Each area is devoted to the representation of a specific feature of an object (color, geometrical form, tone, smell, taste, etc). Hence, each object is represented as a collection of four features. Neural groups within each area represent the value of that particular feature according to a topological organization, that is proximal neurons in the area signal the presence of similar values for the attribute, while distant neurons signal the presence of different values. Oscillators within each area are connected via lateral excitatory and inhibitory synapses, arranged according to a Mexican hat disposition. This means that a neuron excites (and it is excited by) its proximal neurons in the area, and it inhibits (and it is inhibited by) more distal neurons. As a consequence, the presence of a specific value for the feature (signaled by a non-zero input for the coding neuron) does not produce the activation of the single neuron, but causes the occurrence of an entire excitation bubble, involving also neurons which code similar values for the attribute.

The lateral intra-area connections are established "a priori" in the model, to reflect the similarity principle, and they do not undergo any learning process. On the contrary, all inter-area synapses, linking neural groups belonging to different areas, are initially set to zero; then, during the learning phase, the weights of these connections vary according to a Hebbian learning rule.

In particular, we memorized three objects within the network. During the learning phase, the network receives one object at a time: only the four neurons signaling the features of one object receive a non-zero input, while all other neurons do not receive any external stimulus. The external input produces a pattern of activity within the network that encodes the information to be stored. Synaptic weights change during this period on the basis of the resulting pattern of internal activity, following a Hebbian learning rule. In particular, we assumed that during the exposure time to the pattern, the strength of the synapses increases until an upper saturation is reached. Hence, the synapses linking two oscillators (in position *ij* and *hk*, in different areas) are modified as follows during learning:

$$W_{ij,hk}(t+\Delta t) = W_{ij,hk}(t) + \varepsilon_{ij,hk} \cdot x_{ij} \cdot x_{hk} \qquad (1)$$

$W_{ij,hk}$ is the weight of synaptic connections, x is the activity of the oscillator and $\varepsilon_{ij,hk}$ is the learning rate which obeys to the equation

$$\varepsilon_{ij,hk} = \varepsilon_0 \cdot \left(W_{max} - W_{ij,hk}\right) \qquad (2)$$

W_{max} is the maximum value allowed for the synapses, and ε_0 is the maximum learning rate (when synapses are zero).

Then, synapses so generated are "frozen", to preserve prior knowledge within the associative inter-area connections. It is worth noticing that due to the occurrence of an excitation bubble within each area, during the learning phase inter-area synapses originate not only among the neurons signaling the exact attributes of the object but also among neurons which lie inside the activation bubble, signaling similar values of the features. The network has been exposed to each object for a different time during the learning phase; therefore, inter-area synapses assume a different strength for each object. In particular, object 1 is the strongest (that is it has been shown to the network for the longest time), while object 3 is the weakest.

Intra-area and inter-area synapses are not able to warrant desynchronization among multiple objects during the retrieval phase. Hence, in order to solve the segmentation problem, we introduced a desynchronizing mechanism in the network, which consists of a global inhibitor. The global inhibitor is connected to all excitatory units of the network: it computes the overall excitatory activity of the network and sends back an inhibitory signal when this activity overcomes a threshold.

III. Results

First, the behavior of the network has been analyzed by assuming that all three objects are given as input to the network but with different degrees of incompleteness.

Simulation shown in Fig. 2 has been performed assuming the absence of one input property in each object. The figure displays network activity in all neural groups at different snapshots during the simulation, starting from a random initial state. After a short transient, the inter-area synapses, which incorporate prior knowledge, allow restoration of lacking information, recreating the property which is not given as input, while the desynchronization mechanism avoids interference among properties of different objects. Results of further simulations are summarized in Table I, which reports the percentage of success obtained in 10 different trials for each considered task (with random initial values for the network). For each object, the number of properties given as input is specified. The network exhibits high performances even in very difficult conditions such as task 9, where object 1 lacks two properties and the other two objects lack one property: even in this case, after an initial transient, the three objects are perfectly reconstructed and segmented. Regarding to this, it is worth noticing that object 1 can be reconstructed even when only two properties are given as input (tasks 8 and 9), thanks to its strong inter-area synapses; however, only one input property falls to reconstruct the entire object. On the contrary, only two properties are not sufficient to recover objects 2 and 3, which are characterized by weaker inter-area synapses. Nevertheless, the missed reconstruction of an object, because of the insufficient number of input properties, does not interfere with the reconstruction of the other two objects. An example of this network behaviour is shown in Fig. 3; in this case, object 1 is not reconstructed since it receives only one input property, while the other two objects are perfectly reconstructed starting from three input properties.

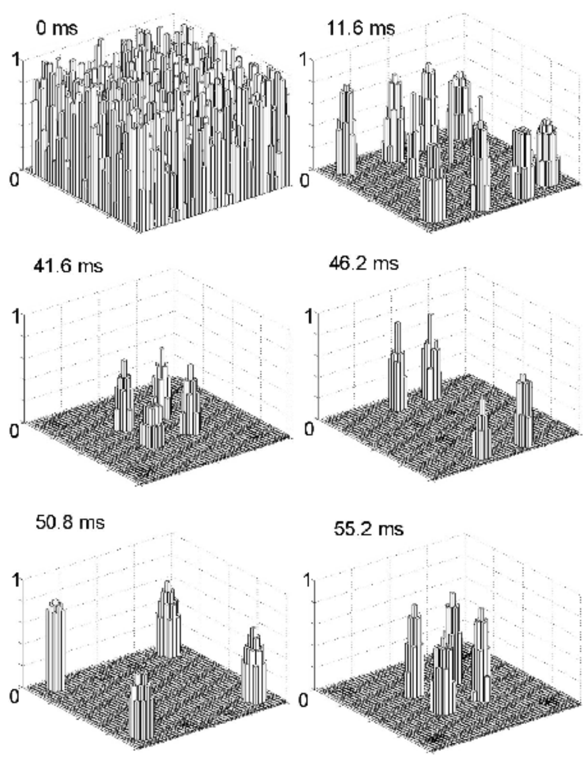

Fig. 2. Network activity at different snapshots during the simulation. Each pixel represents an oscillator. The emerging height is proportional to the corresponding oscillator's activity. Each of the three objects receives three input properties: after an initial transient (starting from a random initial state) the objects are perfectly reconstructed.

TABLE I

task	**object / number of input properties**	**percentage of success**
1	obj 1/4 prop – obj 2/4 prop – obj 3/3 prop	100%
2	obj 1/4 prop – obj 2/3 prop – obj 3/4 prop	100%
3	obj 1/3 prop – obj 2/4 prop – obj 3/4 prop	100%
4	obj 1/4 prop – obj 2/3 prop – obj 3/3 prop	100%
5	obj 1/3 prop – obj 2/4 prop – obj 3/3 prop	90%
6	obj 1/3 prop – obj 2/3 prop – obj 3/4 prop	100%
7	obj 1/3 prop – obj 2/3 prop – obj 3/3 prop	80%
8	obj 1/2 prop – obj 2/4 prop – obj 3/4 prop	100%
9	obj 1/2 prop – obj 2/3 prop – obj 3/3 prop	100%

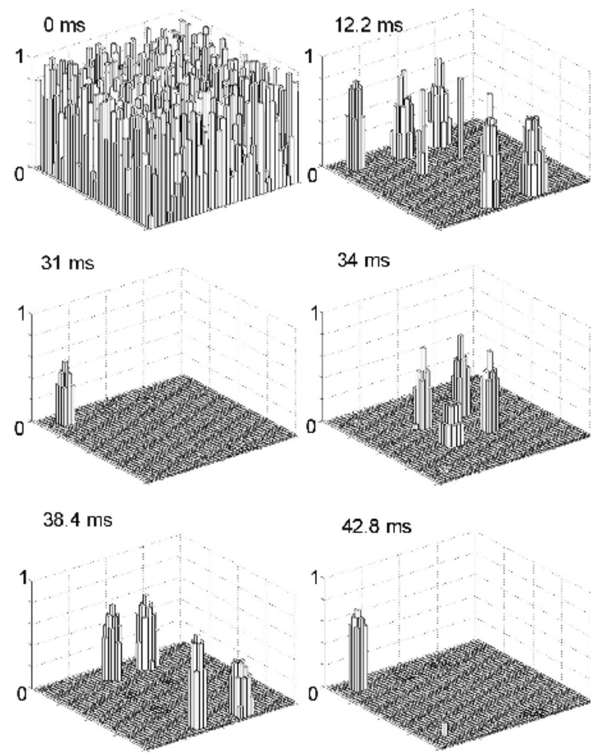

Fig. 3. Network activity at different snapshots during the simulation. Object 1 receives only one input property, which is insufficient for its reconstruction; however, object 2 and 3 are correctly reconstructed starting from three input properties.

In the successive simulations, the network faces partial input information together with a shift in the input properties. The network is able to reconstruct one object with two exact properties, one lacking property and one property shifted from the exact value by one position in each direction (Euclidean distance equal to √2), if the other two objects are complete and perfectly aligned with the stored patterns (100% of success). Object 1 can be successfully reconstructed even when the distance of the corrupted property is as high as √5. The performances of the network decrease when more objects exhibit wrong properties (40% of success when two objects have one lacking property and one property shifted by √2 from the exact value) or in case of excessive corruption of a single object (e.g., the network falls to reconstruct object 1 when three properties are corrupted by √2 and one property is lacking).

IV. Discussion

Aim of the present work is to analyze the solution of the segmentation and recall of complex objects by using a two-dimensional neural network of Wilson-Cowan oscillators, which includes high-level Gestalt laws of perceptual grouping and one mechanism for segmentation.

The basic concept of our model is that the features of an object can be bind together and separated from the attributes of different objects via synchronization of oscillatory groups in the γ-band. The associative properties used in the model to bind different oscillators into a unique pattern reflect two high-level Gestalt laws: similarity and previous knowledge.

The similarity principle ensues in the model from the topological organization of each area, and from the presence of lateral synapses within the same area, arranged according to the classical "Mexican hat" profile. Empirical studies have shown that topological organization exists in the mammalian brain and different topological maps have been isolated in several parts of the cortex, such as the auditory cortex, the visual cortex and the motor cortex [9], [10].

The prior knowledge is incorporated in the inter-area synapses, whose value is established during the learning phase, according to a Hebbian learning rule. Since, during the learning phase, each object is presented to the network for a different time interval, the objects are memorized in the network with a different force. A similar result can be obtained by using the same exposure time but distinct values of the learning rate for each object (parameter ε_0 in (2)). Changes in this parameter might reflect the existence of some additional attentive or emotional mechanisms, which modulate the plastic behaviour of the synapses. Hence, different values of the inter-area synapses from one object to another may correspond to objects with a different familiarity or with a different importance and emotional impact.

The presence of the inter-area and intra-area synapses allows a rapid synchronization among oscillators, thus ensuring a rapid solution of the binding problem. However, the inter- and intra-area connections do not avoid interference among different objects, that is they are insufficient to solve the segmentation problem. Segmentation in the temporal domain requires that the total oscillation period is subdivided into fractions, during which only neurons of a particular object oscillate in phase, with all other neurons remaining silent. In the model, object segmentation ensues from the action of the global inhibitor. The role of the global inhibitor can be explained as follows. When a pattern of oscillators increases its activity in phase, the global inhibitor sends an inhibitory signal to the entire network. This inhibitory signal stops the activity of all oscillators which have not still entered the active phase and excitation remains confined within a single object.

Simulation results show that the proposed mechanisms may actually work, producing a high percentage of success even when the network has to face with very difficult conditions (see for example task 9). An interesting result obtained with the model is that the performances of the network in recalling and recognizing previously learned objects depend on the force of their memorization. The object more imprinted in memory (object 1) can be perfectly reconstructed even in case of 50% of incompletion, thanks to the greater strength of its inter-area synapses. On the contrary, at least three features are necessary to recognize the other two objects, which are stored in memory with a lower force.

Moreover, an important characteristic of our network is its ability to recall previously memorized objects, not only starting from partial input information, but also when input information exhibits a shift in the input space. The insensitivity of the network to a moderate input shift arises from the presence of the lateral intra-area synapses, which produce "activation bubbles". This property strongly differentiates the present network from other associative memories, based on Hebbian learning, such as the Hopfield net, which is not able to sustain a shift in the input properties.

Finally, the network avoids reconstruction of excessively incomplete or corrupted objects, hence showing a good compromise between sensitivity (capacity to detect true positives) and selectivity (capacity to reject false positives).

References

[1] C. von der Malsburg and W. Schneider, "A neural cocktail-party processor," *Biol. Cybern.*, vol. 54, pp. 29-40, 1986.

[2] W. Singer, "Synchronization of cortical activity and its putative role in information processing and learning," *Ann. Rev. Physiol*, vol. 55, pp. 349-374, Oct. 1993.

[3] C. Tallon-Baudry and O. Bertrand, "Oscillatory gamma-activity in humans and its role in object representation," *Trends Cogn. Sci.*, vol. 3, pp. 151-162, Apr. 1999.

[4] C. M. Gray, "The temporal correlation hypothesis of visual feature integration: still alive and well, " *Neuron*, vol. 24, pp. 31-47, Sept. 1999.

[5] W. Singer, "Neuronal synchrony: a versatile code for the definition of relations?," *Neuron*, vol. 24, pp. 49-65, Sept. 1999.

[6] J. Bhattacharya, H. Petsche, and E. Pereda, "Long-range synchrony in the γ band: role in music perception," *J. Neurosci.*, vol. 21, pp. 6329-6337, Aug. 2001.

[7] J. Kaiser and W. Lutzenberger, "Induced gamma-band activity and human brain function," *Neuroscientist*, vol. 9, pp. 475-484, Dec. 2003.

[8] K. Koffka, *Principles of Gestalt Psychology*. New York: Harcourt, 1935.

[9] S.E. Palmer, *Vision Science*. Cambridge, MA: MIT press, 1999.

[10] E. T. Rolls and G. Deco, *Computational Neuroscience of Vision.* Oxford: University Press, 2004.

[11] D. Wang and D. Terman, "Image segmentation based on oscillatory correlation," *Neural Comput.*, vol. 9, pp. 805-836, May 1997.

[12] Z. Li, "A neural model of contour integration in the primary visual cortex," *Neural Comput.*, vol. 10, pp. 903-940, May, 1998.

[13] C. Lourenço, A. Babloyantz, and M. Hougardy, "Pattern segmentation in a binary/analog world: unsupervised learning versus memory storing, " *Neural Netw.*, vol. 13, pp. 71-89, Jan. 2000.

[14] Y. Kazanovich and R. Borisyuk, "Object selection by an oscillatory neural network," *BioSystems*, vol. 67, pp. 103-111, Oct.-Dec. 2002.

[15] D. Wang and X. Liu, "Scene analysis by integrating primitive segmentation and associative memory," *IEEE Trans, Man Cybern.-Part B: Cybern.*, vol. 23, pp. 254-268, June 2002.

[16] M. Ursino, G. E. La Cara, and A. Sarti, "Binding and segmentation of multiple objects through neural oscillators inhibited by contour information," *Biol. Cybern.*, vol. 89, pp. 56-70, July 2003.

[17] R. Borisyuk and Y. Kazanovich, "Oscillatory model of attention-guided object selection and novelty detection," *Neural Netw.*, vol. 17, pp. 899-915, Sep. 2004.