Visual Selection and Shifting Mechanisms Based on a Network of Chaotic Wilson-Cowan Oscillators

Liang Zhao, Fabricio A. Breve, Marcos G. Quiles, and Roseli A. F. Romero
Department of Computer Science
Institute of Mathematics and Computer Science, University of São Paulo
Caixa Postal: 668, São Carlos - SP, 13560-970, Brazil
{zhao, fabricio, quiles, rafrance}@icmc.usp.br.

Abstract

In this paper, a Visual Selection and a Shifting Mechanisms based on a lattice of coupled chaotic Wilson-Cowan oscillators is proposed. The oscillators representing each object in a given visual scene are synchronized to produce a chaotic trajectory. A cooperation and competition mechanisms are also introduced to accelerate oscillating frequency of the salient object as well as to slow down other objects in the same scene. The model can not only discriminate each object among others in a given visual scene, but also deliver the focus of attention to a salient object in an instant of time. In comparison to other visual selection approaches, this model presents at least three new features. First, it is able to highlight objects in complex forms, including those that are non-linear separable. Second, a shifting mechanism used to change the focus of attention among the objects. Lastly, oscillators representing the salient object will jump from chaotic phase to periodic phase. This behavior matches well to biological experiments on pattern recognition of rabbit. Computer simulations are performed and the results show that the proposed model is promising as a Selection Mechanism with Shifting embedded in a Visual Attention System.

1. Introduction

Due to the limited processing capacity of biological systems, some mechanisms have evolved in order to permit these systems to perform tasks, such as scene interpretation, in real time. Visual attention is an efficient mechanism that biological systems have developed to address the reduction of provided visual information. Attention can be related to the individual capacity of discriminating one significant stimulus among others. This process appears to optimize the search procedure by selecting a number of possible candidate image and feature subsets which can be used in tasks such as recognition [21]. It is also responsible to break down complex tasks into a series of small localized computational tasks [9]. According to Tsotsos et al. [18], intermediate and higher visual processes seem to select part of the sensory information received from the world and use just these selected data in a further processing. Visual attention is also responsible for reducing the combinatorial explosion resulting from the analysis of all incoming sensory information and possible image relationships [14, 17] and for identifying the part of the visual input where the processing is performed at the same time which irrelevant visual information is suppressed [5].

Visual attention is generated by a combination of information from the retina and early visual cortical areas (bottom-up attention - scene dependent) as well as feedback signals from areas outside of the visual cortex (top-down attention - task dependent) [10, 11]. Bottom-up attention is a feedforward process formed by simple features extracted from the input image, such as intensity, stereo disparity, color, orientation, and others [10]. All this information is combined to create a saliency map which represents the conspicuity points in the visual input. The top-down attention is responsible for modulating the competition of all stimulus within the visual input. This last process entails keeping in a short-term memory the information about location or an object which is used as a target of attention to influence the earlier visual processes [6].

In this work we are mainly concerned with bottom-up attention. Most of the bottom-up visual attention models are related to the concept of a Saliency Map [10]. In these models, the first stage of processing is responsible to decompose the input image into a set of feature maps. After that, a saliency map is generated by a combination of those feature maps. The saliency map is a topographical map which represents, by a scalar quantity, all salient points over the entire input visual stimulus [10, 11]. The main purpose of

this map is to guide a selection mechanism, which is responsible for delivering the focus of attention in a specific (or most conspicuity) region of the image.

Although several models of bottom-up visual attention have been proposed, the Selection mechanism used by these models are implemented by a Winner-Take-All (WTA) neural network where just one neuron is activated but not the entire object becomes salient. For example, in the model proposed in [11], when a neuron receives the focus of attention, a circle with a fixed radius is considered to be the region of attention of the model. In this case, it is not possible to deliver the attention to complex objects that are non-linear separable. To deal with this limitation, new object selection mechanism should be developed.

von der Malsburg [19] proposed a mechanism of temporal correlation as a representational framework. This theory suggested that objects are represented by the temporal correlation of the firing activities of spatially distributed neurons coding different features of an object. A natural way of encoding temporal correlation is by using synchronization of oscillators where each oscillator encodes some features of an object [16, 20, 22]. Inspired from the biological findings and von der Masburg's brain correlation theory, Wang and his collaborators have developed oscillatory correlation theory for scene segmentation [3, 4, 16, 23], which can be described by the following rule: the neurons which process different features of the same object are synchronized, while neurons which code different objects are desynchronized. There are two basic mechanisms working simultaneously in each oscillatory correlation model: synchronization and desynchronization. The former serves to group neurons into objects while the latter serves to distinguish one group of synchronized neurons (an object) from another. Oscillatory correlation theory has been extended and successfully applied to various tasks of scene analysis, such as image segmentation, motion determination, auditory signal segregation, and perception ([22] and references there in). Another way to model oscillatory correlation and to achieve unlimited capacity of segmentation (number of objects can be segmented in a given scene) is to utilize the properties of chaos and chaotic synchronization [8, 24, 25]. In the model proposed by Zhao et. al. [25], a large number of locally coupled chaotic oscillators can be synchronized, so that each object in a given scene is represented by a synchronized chaotic trajectory in the corresponding network. Consequently, all such chaotic trajectories can be easily separated by the high sensitivity to initial conditions, which is the hallmark of chaos, and the fact that a chaotic trajectory is dense in its invariant set. With this procedure, the authors claim that the model has unbounded capacity of object segmentation.

The use of synchronization has also received additional support from neurobiological findings which have demonstrated that visual attention is strongly linked with synchronization among neurons. Biological experiments have shown that visual attention increases the coherence among neurons responding to the same stimulus suggesting that synchronization is an important mechanism for visual selection [2, 7, 12, 13].

In this paper, we construct a chaotic oscillatory correlation network for object selection with a shifting mechanism. Our model is based on a network of coupled chaotic Wilson-Cowan oscillators [25]. Such a network is used to create a selection mechanism where one of several objects is highlighted (receives the focus of attention). As the system runs, each group of neurons representing an object of a visual input is synchronized and produce a unique chaotic trajectory. At the same time, a competition mechanism is also introduced, where synchronized neurons cooperate each other to accelerate their firing frequencies and slow down other neurons with different oscillating activities. Finally, the most salient object will jump to a high frequency periodic oscillating phase, while all other objects will be quite silent. After receiving attention, this object is inhibited in order to permit others objects to become salient, it means, to receive the focus of attention. This later process can be seen as a habituation system where the presence of a fixed winner stimulus must be followed by a progressive diminution of its response and allowing other stimulus to become active (focus of attention shifting). In this paper, we consider pixel intensity as visual attention clue, i.e., the brightest object is considered as the most salient object.

The rest of the paper is organized as follows. In section 2, the segmentation process performed by a network of Wilson-Cowan oscillators is described and the proposed model are presented. In Section 3, the results obtained through simulation of the proposed model applied to synthetic images are showed. Finally, in Section 4, conclusions are drawn.

2. Model Description

In this section, we first present a network of coupled chaotic Wilson-Cowan oscillators for scene segmentation. Then we introduce new selection mechanism, embedded in the segmentation network, to accomplish visual attention task.

2.1. Scene Segmentation using Coupled Wilson-Cowan Oscillators

The model is a two dimensional network and governed by the following equations:

$$\dot{x}_{i,j} = -ax_{i,j} + G(cx_{i,j} + ey_{i,j} + I_{i,j} - \theta_x) + k\Delta x_{i,j}$$
$$\dot{y}_{i,j} = -by_{i,j} + G(dx_{i,j} + fy_{i,j} - \theta_y) + k\Delta y_{i,j} \quad (1)$$
$$G(v) = \frac{1}{1 + e^{-(v/T)}}$$

where (i,j) is a lattice point with $1 \leq i \leq N, 1 \leq j \leq M$. k is the coupling strength. $\Delta x_{i,j}$ and $\Delta y_{i,j}$ are coupling terms among excitatory units and inhibitory units, respectively. They are defined by

$$
\begin{aligned}
\Delta v_{i,j} =\ & \gamma_{i-1,j-1;i,j}(v_{i-1,j-1} - v_{i,j}) + \\
& \gamma_{i-1,j;i,j}(v_{i-1,j} - v_{i,j}) + \\
& \gamma_{i-1,j+1;i,j}(v_{i-1,j+1} - v_{i,j}) + \\
& \gamma_{i,j-1;i,j}(v_{i,j-1} - v_{i,j}) + \\
& \gamma_{i,j+1;i,j}(v_{i,j+1} - v_{i,j}) + \quad (2)\\
& \gamma_{i+1,j-1;i,j}(v_{i+1,j-1} - v_{i,j}) + \\
& \gamma_{i+1,j;i,j}(v_{i+1,j} - v_{i,j}) + \\
& \gamma_{i+1,j+1;i,j}(v_{i+1,j+1} - v_{i,j})
\end{aligned}
$$

where

$$\gamma_{i,j;p,q} = \begin{cases} 1, & \text{if element } (i,j) \text{ is coupled to } (p,q), \\ 0, & \text{otherwise.} \end{cases}$$
$$(3)$$

Without consider the coupling terms, eqn. (1) represents a Wilson-Cowan neural oscillator [1], which has been widely used in neural network modeling. It is a feedback loop between an excitatory unit (x) and an inhibitory unit (y). $I_{i,j}$ is an external stimulation received by oscillator (i,j). If $I_{i,j}$ is a constant, no chaos can appear since it is a two-dimensional continuous flow. In order to get a chaotic oscillator, the external stimulation is defined as a periodic function: $I_{i,j}(t) = A_{i,j}cos(t)$, where $A_{i,j}$ is the amplitude of external stimulation. In all simulations of this paper, $A_{i,j}$ is considered as a bifurcation parameter, which receives gray level of an input pixel.

The segmentation strategy is described below. Considering a scene image containing several objects. The network is organized that each element corresponds to a pixel of the image and a proper parameter of each oscillator is chosen to encode the gray level of the corresponding pixel. As the system runs, the neurons self-organize according to a predefined similarity criterion, such that the connections between pairs of neighboring oscillators with similar gray level will be maintained, while those connections between oscillators of very different gray level will be cut. Consequently, all neurons belonging to the same segment will be synchronized to form a unique trajectory, then each object is represented by a synchronized chaotic orbit. Following the

high sensitivity to initial conditions and the dense properties of chaos, all such synchronized chaotic trajectories will be mutually different in time. In this way, objects in a given scene can be segmented.

2.2. Network of Wilson-Cowan Oscillators as a Visual Selective System with Shifting

The proposed model can be seen as an adaptation of the model described above with three new major features. First, a mechanism to select a salient object by increasing the firing frequency of its corresponding oscillators, and by decreasing the firing frequency of the oscillators corresponding to the other objects is implemented. Second, a shifting mechanism is used to change the focus of attention among the objects. And finally, the neurons representing the object in the focus of attention change their behavior from a chaotic to a periodic phase. These features satisfy the following essential requirements of a Visual Selection system:

- considering as input one or a combination of features (Saliency Map), the neural network must highlight (select) the region of the image where the focus of attention should be directed;

- all other locations of the visual input must be suppressed by the system in order to keep the focus of attention on just one of the active objects.

- the focus of attention must be shifted to other remaining active locations. It means that the neural network must implement a type of habituation system where the presence of a fixed winner stimulus must be followed by a progressive diminution of its response and allowing other stimulus to become active.

The spiking frequency of Wilson-Cowan oscillators can be controlled by changing the parameter b in equation (1). So, we first analyzed the bifurcation diagram of periodically driven Wilson-Cowan oscillators by varying the parameter b, as shown in Fig. 1. From this figure, we see that when $b \leq 0.005$ there is nearly no oscillation. As b increases we see sessions of periodic windows and chaotic behavior. When b is small the chaotic behavior is predominant, but as b increases the periodic behavior becomes more frequent.

In Fig. 2, we show the time series of a Wilson-Cowan oscillator varying b. From this figure we can notice that as b increases, the frequency of oscillation increases. When b takes a high value (for example, the case of $b = 0.1$ in the figure), the oscillator not only fires more frequently, but also shows periodic resembling behavior.

In our model, we take advantage of these oscillatory changes to determine visual attention, which means that the
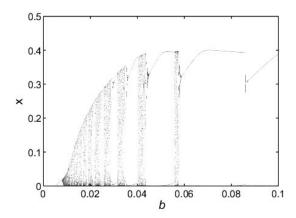
Figure 1. Bifurcation diagram of periodically driven Wilson-Cowan oscillators by varying parameter b. The stepsize $\Delta b = 0.0001$

Figure 2. Temporal activities of oscillators with $b = 0.01$, 0.02, 0.034, 0.05 and 0.1 respectively. Vertical scale of second to fifth oscillators were shifted downwards by 0.5

synchronized oscillators corresponding to the salient object will present a periodic oscillation, while the oscillators corresponding to the other objects will remain chaotic with low frequency.

In order to achieve this behavior, we let the oscillators run with a fixed b parameter until they synchronize and the segmentation task can be performed. After that, whenever any oscillator fires, say oscillator (i, j), it will produce two types of signals: 1) a reinforcement signal to itself and all other oscillators that fire together with it. This reinforcement is weighted by the pixel intensity that feeds this oscillator; 2) an inhibitory signal to all the oscillators that have different activities to (i, j).

The reinformecement/inhibitory behavior is defined by the following equation:

$$b_{p,q}(\tau) = b_{p,q}(\tau-1) + \frac{\alpha}{M(\tau)} \sum_{i,j \epsilon \Delta(\tau)} \omega_{i,j} f(||x_{i,j} - x_{p,q}||) \tag{4}$$

$$b_{min} \leq b_{p,q}(\tau) \leq b_{max} \tag{5}$$

where (p, q) is neuron's index, τ is a time instant with at least one firing oscillator, $M(\tau)$ is the number of oscillators at the firing state in τ time, $\Delta(\tau)$ is the set of oscillators at the firing state in τ time, b_{min} and b_{max} are constant, $||x||$ is the norm of x, and $f(x) = a_1 x + a_2$. By setting $a_1 < 0$, $f(x)$ defines that each firing neuron, say neuron (i, j), may send a positive or a negative signal to another neuron, say (p, q), depending on the difference between them. Thus, each neuron oscillator sends a positive signal to the oscillators of the same group (represent the same object), while it sends a negative signal to the oscillators of other groups. $\omega_{i,j}$ encodes the intensity of pixel (i, j).
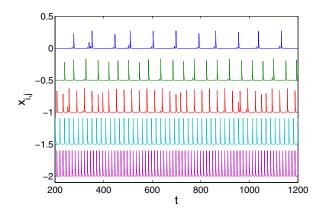
Each oscillator that receives a reinforcement signal will increase the value of parameter b accordingly, while the oscillator receiving an inhibitory signal will decrease the value of parameter b as well. The maximum value of b an oscillator can take is b_{max}, while the minimum value is b_{min}. After some time, only the oscillators corresponding to the salient object will keep firing, increasing their activation frequency until they enter a chaotic state with fixed phase. The oscillators corresponding to the non-salient objects will gradually decrease their activation frequency and maintain chaotic both in amplitude and phase. In our model, this object which remains active is the one that will receive the focus of attention.

A group of neurons representing an object is considered to be salient when the parameter $b = b_{max}$. At this moment, all these neurons are firing with fixed frequency. After that, the habituation mechanism is triggered, which is responsible for inhibiting this object and allowing a shifting of the focus of attention to another salient object. All inhibited objects remain quite silent, it means, with a small chaotic firing frequency. When all salient objects in the input image become inhibited, all neurons representing them are reset and allowed to receive the focus of attention again.

3. Computer Simulations

This section presents the simulation results performed on synthetic images in order to check the viability of the proposed model as a Selection Mechanism with shifting of the focus of attention.

In all simulations of this paper, the following parameters are held constant at: $a = 1.0$, $c = 1.0$, $d = 0.6$, $e = -2.5$, $f = 0.0$, $k = 4$, $\theta_x = 0.2$, $\theta_y = 0.15$, $T = 0.025$,

Figure 3. Artificial image with 4 **objects. (**$25 \times$ 25 **pixels).**

$b_{min} = 0.009$, $b_{max} = 0.1$, $\alpha = 0.001$, $a_1 = -4$, $a_2 = 1$, and the initial value of b is 0.02. Such a configuration can guarantee that initially each oscillator is chaotic [25]. With these parameter values, we perform some experiments with synthetic images.

The first experiment was carried by using the artificial image shown in Fig. 3, which has 4 objects with different gray levels. Figure 4 shows the temporal activities of oscillator groups where the first 200 steps are the transient period responsible for the segmentation process. After that, we can see the behavior of the system delivering the attention to a specific object and switching the focus of attention among the objects of the image. As mentioned above, in our simulations only the intensity of pixels are used as input. It means that the object with the highest intensity receives the attention earlier than other objects with lower intensity.
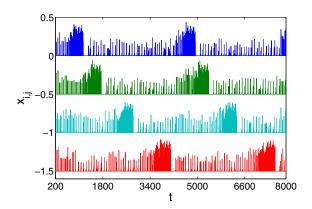

Figure 4. Temporal activities of oscillator blocks. Each trace in the figure corresponds to an object in the input pattern of Fig. 3. Vertical scale of the second objects is shifted downwards by 0.5**.**

Our second experiment is performed by using the artifi-

cial image shown by Fig 5 containing 5 objects of different gray levels. It can be seen from the figure that the five objects are *linearly nonseparable*. Even though, the 5 objects are correctly segmented and Fig. 6 shows the temporal activities of the oscillators corresponding to each object. Again, we see the behavior of the system where the attention is delivery to one of the five objects in an instant of time.

In both experiments it is possible to see that the frequency of non-highlighted objects remain quite silent and chaotic, while the frequency of the object in the focus of attention is increased and the neurons start to fire with fixed frequency as it was expected.
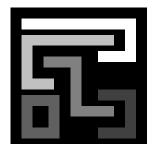

Figure 5. Artificial image with 5 **with** 4 **linearly nonseparable objects. (**25×25 **pixels)**
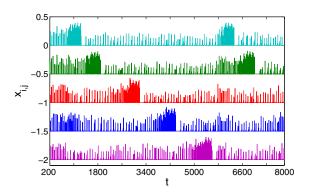

Figure 6. Temporal activities of oscillator blocks. Each trace in the figure corresponds to an object in the input pattern of Fig. 6. Vertical scale of second to fifth objects are shifted downwards by 0.5**.**

4. Conclusions

This paper presents a Visual Selection Mechanism and a Shifting Mechanism based on a network composed of

chaotic Wilson-Cowan oscillators. These mechanisms can be seen as part of a Visual Attention System, which are responsible for selecting one of several regions of interest into the visual input image and also to switch from one previously selected object to another. The proposed model utilize the properties of chaos and chaotic synchronization to discriminate the objects that compose the visual input and also included a inhibition mechanism which is responsible for highlighting the most salient object. Another interesting characteristic of the proposed model is its change of behavior when the object receives the focus of attention. In this case, the former chaotic behavior gives place to a trajectory with fixed phase, which has been observed in the biological experiments of pattern recognition of rabbit [15].

Computer simulations were performed in order check its viability as a selection mechanism and the results show that our model is a promising mechanism for computer vision systems.

As a future work we intend to create a complete saliency map using other features, such as color, saturation, orientation, etc. in order to test our model using real images and compare our results with other models based on saliency maps. In addition, we will also verify the possibility of including some biasing mechanism to emulate top-down factors based on prior knowledge of the visual input, such as a memory holding a specific object.

Acknowledgment

References

[1] B. Baird. Nonlinear dynamics of pattern formation and pattern recognition in the rabbit olfactory bulb. *Physica D*, 22:150–175, 1986.

[2] C. Buia and P. Tiesinga. Attentional modulation of firing rate and synchrony in a model cortical network. *Journal of Computational Neuroscience*, 20:247–264, 2006.

[3] S. R. Campbell and D. Wang. Synchronization and desynchronization in a network of locally coupled wilson-cowan oscillators. *IEEE Transactions on Neural Networks*, 7:541–554, 1996.

[4] S. R. Campbell, D. L. Wang, and C. Jayaprakash. Synchrony and desynchrony in integrate-and-fire oscillators. *Neural Computation*, 11:1595–1619, 1999.

[5] L. Carota, G. Indiveri, and V. Dante. A softwarehardware selective attention system. *Neurocomputing*, 58-60:647–653, 2004.

[6] G. Deco and E. T. Rolls. Attention, short-term memory, and action selection: A unifying theory. *Progress in Neurobiology*, 76:236–256, 2005.

[7] P. Fries, J. H. Reynolds, A. E. Rorie, and R. Desimone. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291:1560–1563, 2001.

[8] D. Hansel. Synchronization and computation in a chaotic neural network. *Physical Review Letters*, 68(5):718–721, 1992.

[9] L. Itti. *Neurobiology of Attention*, chapter 94 - Models of bottom-up attention and saliency, pages 576–582. Elsevier, Oxford, 2005.

[10] L. Itti and C. Koch. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2:194–203, 2001.

[11] L. Itti, C. Koch, and E. Niebur. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11).

[12] Y. J. Kim, M. Grabowecky, K. A. Paller, K. Muthu, and S. Suzuki. Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nature Neuroscience*, 10(1):117–125, 2007.

[13] E. Niebur and C. Koch. A model for neuronal implementation of selective visual attention based on temporal correlation among neurons. *Journal of Computational Neuroscience*, 1:141–158, 1994.

[14] F. Shic and B. Scassellati. A behavioral analysis of computational models of visual attention. *International Journal of Computer Vision*, 73(2), 2007.

[15] C. A. Skarda and W. J. Freeman. How brain make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10:161–195, 1987.

[16] D. Terman and D. Wang. Global competition and local cooperation in a network of neural oscillators. *Physica D*, 81:148–176, 1995.

[17] J. K. Tsotsos. On the relative complexity of active vs. passive visual search. *International Journal of Computer Vision*, 7:127–141, 1992.

[18] J. K. Tsotsos, S. M. Culhane, W. Y. K. Wai, Y. Lai, N. Davis, and F. Nuflo. Modeling visual attention via selective tuning. *Artificial Intelligence*, 78:507–545, 1995.

[19] C. von der Malsburg. The correlation theory of brain function. Technical report, Internal report 81-2: Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany, 1981.

[20] C. von der Malsburg and W. Schneider. A neural cocktail-party processor. *Biological Cybernetics*, 54:29–40, 1986.

[21] D. Walther, U. Rutishauser, C. Cock, and P. Perona. Selective visual attention enables learning and recognition of multiples objects in cluttered scenes. *Computer Vision and Image Understanding*, 100:41–63, 2005.

[22] D. Wang. The time dimension for scene analysis. *IEEE Transactions on Neural Networks*, 16(6):1401–1426, 2005.

[23] D. Wang and D. Terman. Image segmentation based on oscillatory correlation. *Neural Computation*, 9:805–836, 1997.

[24] L. Zhao and E. E. N. Macau. A network of dynamically coupled chaotic maps for scene segmentation. *IEEE Transactions on Neural Networks*, 12(6):1375–1385, 2001.

[25] L. Zhao, E. E. N. Macau, and N. Omar. Scene segmentation of the chaotic oscillator network. *International Journal of Bifurcation and Chaos*, 10(7):1697–1708, 2000.

COMPUTER SOCIETY