



HAL
open science

Using Neural Dynamics to Switch Attention

Julien Vitay, Nicolas P. Rougier

► **To cite this version:**

Julien Vitay, Nicolas P. Rougier. Using Neural Dynamics to Switch Attention. International Joint Conference on Neural Networks - IJCNN'05, Jul 2005, Montreal, Canada. inria-00000144

HAL Id: inria-00000144

<https://inria.hal.science/inria-00000144>

Submitted on 5 Jul 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Using Neural Dynamics to Switch Attention

Julien Vitay

Loria, Campus Scientifique, BP 239
54506 Vandoeuvre-les-Nancy, France
E-mail: vitay@loria.fr

Nicolas Rougier

Loria, Campus Scientifique, BP 239
54506 Vandoeuvre-les-Nancy, France
E-mail: rougier@loria.fr

Abstract—We present a distributed and dynamic model of visual attention based on the Continuum Neural Field Theory that allows to sequentially focus salient locations in an image. A working memory system ensures that the corresponding objects are only focused once, even if they are moving around, such that the visual search is efficient. The model has been implemented on a robotic platform in order to search for natural objects such as fruits.

I. INTRODUCTION

Despite its massively parallel architecture, the brain has to cope with high-dimensional and temporal sensory information that exceeds its processing capacities. A solution would be to multiply the number of neurons to adequately represent these information flows, but for evident reasons of brain volume relative to the body and energy consuming, the evolution has led to the emergence of serial mechanisms somehow "emulating" a parallel functioning. For example, in the visual perception domain, the fundamental experiment by Treisman and Gelade [1] has drawn the distinction between two modes of visual search: when an object has characteristics sufficiently different from other objects in the scene, it literally "pops-out" from the scene and the search for it is very quick and independent from the number of other objects; oppositely, when this object shares some features with distracting objects or when it does not differ enough from its background, the search is very difficult and the time needed for it increases linearly in average with the number of distractors, as if every object were sequentially scanned until the target is found. These two search behaviours are then respectively called "parallel search" and "serial search".

The purpose of this paper is to show an example of how a serial behaviour can emerge from a completely distributed neural substrate. The task we chose is to sequentially and uniquely focus salient targets on the image seen by a robot. What we understand here by salient targets is targets whose visual characteristics are not sufficient to produce a "pop-out" effect (like an orange among green apples) but are a conjunction of basic features partly shared by distractors (e.g. a small green lemon among big green apples and small yellow lemons). The idea there is that the small green lemon has no particular advantage in the task compared to the other fruits, because of the lack of "small and green" conjunction filters: the robot has to sequentially scan each fruit until the correct target has been recognized by another mechanism.

Our view is that visual attention is a mechanism enhancing

the processing of interesting (understood as task-relevant) locations and darkening the rest [2], [3], so that fine recognition (or disambiguation) can be processed only at these locations. The first neural correlate of that phenomenon has been discovered by Moran and Desimone [4] in V4 where neurons respond preferentially for a given feature in their receptive field. When a preferred and a non-preferred stimulus for a neuron are presented at the same time in its receptive field, the response becomes an average between the strong response to the preferred feature and the weak response to the non-preferred one. But when one of the two stimulus is attended, the response of the neuron represents the attended stimulus alone (strong or poor), as if the non-attended were ignored.

It appears that attention is an integrated mechanism distributed over sensorimotor structures, whose purpose is to allow increased processing on a small number of regions in the input space in order to achieve relevant motor behaviours (see [5] for a more detailed review). Therefore, virtually all structures involved in behaviour have to deal with attention: for example the link between working memory and attention has been established in [6] and [7]. Attention is a motivated and integrated process.

We suppose here that the focus of attention is the only part of the visual information that efficiently enters the inferotemporal pathway to be recognized, where the progressively overlapping receptive fields allow the recognition of an object independently of its retinal location. The goal of the model we present here is to sequentially switch this focus of attention on the different salient objects by the means of a widely distributed neural architecture.

II. CONTINUUM NEURAL FIELD THEORY

Even if the whole neural networks domain often draws (more or less tightly) on biological inspiration, core mechanisms like the activation function or learning rules often neglect the inner temporal nature of neurons. They are usually designed with no reference to time while it is perfectly known that a biological neuron is a complex dynamic system that evolves over time together with incoming information. If such artificial neurons can be easily manipulated and used in classical networks such as the Multi-Layer Perceptron (MLP), Kohonen networks or Hopfields maps, they can hardly pretend to take time into account. At the same time, the Continuum Neural Field Theory (CNFT) has been extensively analyzed

both for the one-dimensional case [8], [9], [10] and for the two-dimensional case [11] where much of the analysis is extendable to higher dimensions. These theories explain the dynamic of pattern formation for lateral-inhibition type homogeneous neural fields with general connections. They show specifically that, in some conditions, continuous attractor neural networks are able to maintain a localised bubble of activity in direct relation with the excitation provided by the stimulation.

A. The Dynamic Equation of the CNFT

We will use the notations introduced by [11] where a neuronal position is labelled by the vector \mathbf{x} which represents a two-component quantity designing a position on a manifold M in bijection with $[-0.5, 0.5]^2$. The membrane potential of a neuron at the point \mathbf{x} and time t is denoted by $u(\mathbf{x}, t)$. It is assumed that there is a lateral connection weight function $w(\mathbf{x} - \mathbf{x}')$ which is in our case a difference of Gaussian functions (DoG) as a function of the distance $|\mathbf{x} - \mathbf{x}'|$. There also exists an afferent connection weight function $s(\mathbf{x}, \mathbf{y})$ from the position \mathbf{y} in the input manifold M' to the point \mathbf{x} in M . The membrane potential $u(\mathbf{x}, t)$ satisfies the following equation (1):

$$\tau \frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + \int_M w_M(\mathbf{x} - \mathbf{x}') f[u(\mathbf{x}', t)] d\mathbf{x}' + \int_{M'} s(\mathbf{x}, \mathbf{y}) I(\mathbf{y}, t) d\mathbf{y} + h \quad (1)$$

where f represents the mean firing rate as a function of the membrane potential u , $I(\mathbf{y}, t)$ is the input at time t at the position \mathbf{y} in M' and h is the neuron threshold. w_M is given by the equation (2).

$$w_M(\mathbf{x} - \mathbf{x}') = Ae^{\frac{|\mathbf{x} - \mathbf{x}'|^2}{a^2}} - Be^{\frac{|\mathbf{x} - \mathbf{x}'|^2}{b^2}} \text{ with } A, B, a, b \in \mathfrak{R}^{*+} \quad (2)$$

Furthermore, we use a Gaussian function for afferent connections as in equation (3).

$$s(\mathbf{x}, \mathbf{y}) = Ce^{\frac{|\mathbf{x} - \mathbf{y}|^2}{c^2}} \text{ with } C, c \in \mathfrak{R}^{*+} \quad (3)$$

Finally, and depending on the nature of the manifold M we consider (respectively a plane or a sphere surface), we can respectively use the Euclidean distance or the curve distance (which is defined as the shortest length of the geodesic between two points).

B. Discretization

In order to be able to perform numerical simulations using neural network models, we have to discretize these equations. We denote by n the discretization level which represents the regular segmentation of the interval $[-.5, .5]$ into n segments of size $1/n$. A manifold M can consequently be discretized as a set of $n \times n$ units and previous neuronal position \mathbf{x} can be denoted \mathbf{x}_{ij} with $i, j \in [0, n-1]^2$. The corresponding neuronal position is now given by equation (4)

$$\mathbf{x}_{ij} = \left(\frac{i}{n} - 0.5, \frac{j}{n} - 0.5 \right) \quad (4)$$

and equation (1) now becomes:

$$\tau \frac{\partial u(\mathbf{x}_{ij}, t)}{\partial t} = -u(\mathbf{x}_{ij}, t) + \sum_{\mathbf{x}'} w_M(\mathbf{x}_{ij} - \mathbf{x}') f[u(\mathbf{x}', t)] d\mathbf{x}' + \sum_{\mathbf{y}} s(\mathbf{x}_{ij}, \mathbf{y}) I(\mathbf{y}, t) d\mathbf{y} + h \quad (5)$$

One can observe on Figure 1 the impact of projecting the map onto a sphere surface using the curve distance versus projecting onto a plane using the Euclidean distance.

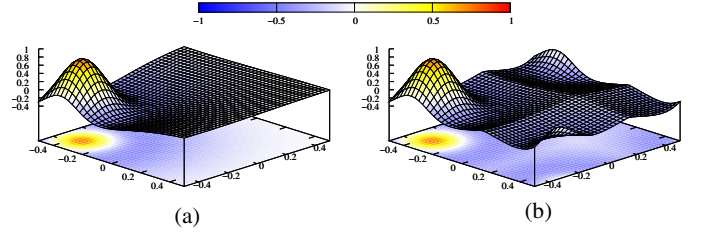


Fig. 1. Lateral connectivity pattern is a simple difference of Gaussian functions (DoG) between a sharp positive Gaussian function and a wider negative one with different intensity and same center. The profile of the DoG is the same for every unit in a map and drives the global activity profile of the whole map. The distance used (Euclidean or curve) depends on the type of projection of the manifold M . On (a) the projection has been made onto a plane and the Euclidean distance has been used whereas on (b), the projection has been made onto a sphere surface and the curve distance has been used.

C. Some Properties

There are several models using population codes focusing on noise clean-up such as in [12], [13] or more general types of computation such as sensorimotor transformations, feature extraction in sensory systems or multisensory integration [14], [15], [16]. Deneve et al. [13] were able to show through analysis and simulations that it is indeed possible to implement an ideal observer using biologically plausible models of cortical circuitry and it comes as no surprise that this model relies heavily on lateral interactions. We also designed a model [17] that uses lateral interactions, as proposed by the CNFT, and fall into the more general case of *recurrent network whose activity relaxes to a smooth curve peaking at a position that depends on the encoded variable* that was analyzed as being a good implementation of a Maximum Likelihood approximator [13]. This dynamic model of attention has been described using the Continuum Neural Field Theory that explains attention as being an emergent property of a neural population. Using distributed and iterative computation, this model has been proven very robust and able to track one static or moving target in the presence of noise with very high intensity or in the presence of a lot of distractors, possibly more salient than the target. The main hypothesis concerning target stimulus is that it possesses a spatio-temporal continuity that should be observable by the model, i.e. if the movement of the target stimulus is too fast, then the model can possibly loose its focus. Nonetheless, this hypothesis makes sense when considering *real world* robotic applications.

III. A COMPUTATIONAL MODEL OF ATTENTION SWITCHING

In [17] we experimentally showed how a single map of neurons using lateral interactions according to the Continuum Neural Field Theory was able to track a stimulus on an input map despite the presence of huge levels of noise, but also despite the presence of other stimuli considered as distractors. As soon as this stimulus is focused, the appearance of potentially more salient stimuli in the input space does not disturb the system, even when the focused stimulus is moving. We can draw a parallel between this interesting property and the "spotlight" metaphor of attention where attended locations are preferentially processed independently of what can happen elsewhere. The question that remains is to determinate how this focus of attention can be moved to another location, especially when the currently attended place has no behavioral relevance. A solution used in the "Bottom-up Visual Attention" model by L. Itti [18] is to locally inhibit the attended location to allow the system to switch to another salient location. This has been done in reference to the "Inhibition-of-Return" phenomenon discovered by Posner and Cohen [19] who showed that previously attended locations have decreased processing abilities, as if they were partially inhibited. The drawbacks of this mechanism is that the switch of attention is automatic (each location is attended a fixed amount of time depending on the neural dynamics) and that nothing ensures that each potentially interesting location will be attended. For example, if the inhibition is too short, the focus of attention can switch back and forth between the two most salient locations only. The purpose of the present model is to deal with these issues: building a system that can explore all the salient locations in a scene without exploring twice the same place and that is able to stop switching whenever a satisfying target is found. Thus we need two different mechanisms: a mechanism able to change the focus of attention when required; a mechanism ensuring that a previously visited location can not be chosen again. The first mechanism deals with Inhibition-of-Return whereas the second is linked with active working memory. As Inhibition-of-return can follow moving targets [20], the two mechanisms have to be updated by perception. In the following paragraph, we will describe the architecture of our model which, even if biologically inspired, does not pretend to model the real attentional mechanisms in the brain.

A. Architecture

As one can see on Figure 2, this model is composed of ten different maps of 40×40 or 20×20 units, what makes the system difficult to analyse, but each unit is governed by the same equation as in Equation 5. The connections between maps have a Gaussian extent like in Equation 3, in a "Receptive-Field"-like manner. The connections inside a map are a difference of Gaussian (Equation 2). As a consequence, all the maps share the same topography. Parameters can be found in [5].

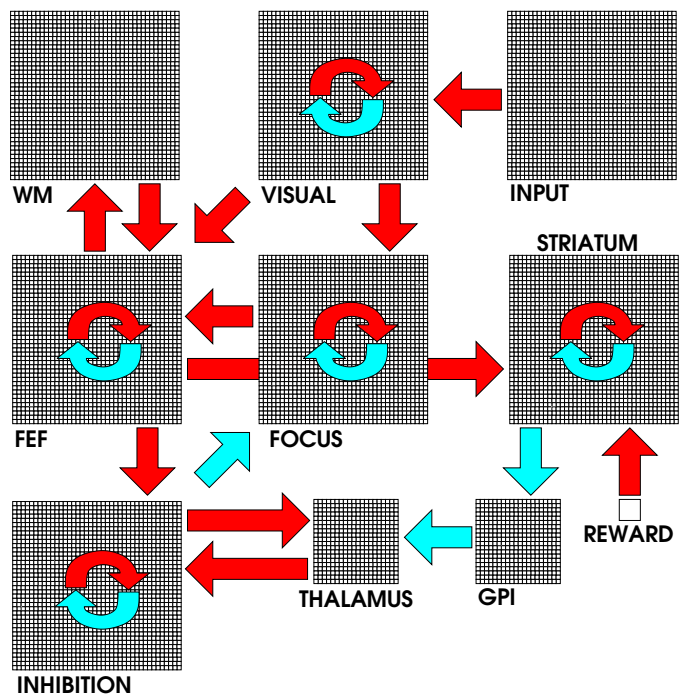


Fig. 2. The different maps of the model, with schematic connections. Red (dark) arrows represent excitatory connections, blue (light) arrows represent inhibitory connections, circular arrows represent lateral connections. See text for details.

1) *Input map*: The INPUT map in the model (cf. Figure 2) is a pre-processed representation of the visual input. Basically, it has to show localized bubbles of activity to mimic saliency in the visual scene. In the simulation, we will generate noisy Gaussian bubbles into that map of 40×40 units. Contrary to the rest of the network, this map has no dynamic behaviour, it just represents visual information.

2) *Visual map*: The VISUAL map receives excitatory inputs from the INPUT map in a receptive-field manner. The lateral connectivity in the VISUAL map ensures that only a limited number of bubbles of activity can emerge anytime. As a consequence, the activity of the VISUAL map is virtually noiseless and expresses only the most salient stimuli present within the input. If too many stimuli are presented in the same time, then the dynamic interactions within the map will reduce this number to the most salient stimuli only. Roughly, in the present architecture, this number is around seven stimuli which can be presented simultaneously (this is mainly due to the size of the map compared to the lateral extent of the inhibitory lateral connections).

3) *Focus map*: The FOCUS map receives excitatory inputs from the VISUAL map. The inhibitory extent of the lateral connectivity is wider than in the VISUAL map so that only one bubble of activity can emerge anytime. When no stimulus is present within the input, no activity appears in the FOCUS map. With these three maps (INPUT, VISUAL and FOCUS), the system can track one stimulus in the input map which will be represented by only one bubble of activation in FOCUS,

which we suppose to represent the currently attended location. In [17] we demonstrated that this simple system had interesting denoising and stability properties. Now, to implement a coherent attention-switching mechanism, we need to add a switching mechanism coupled with a working memory system. The switching mechanism will be done by adding an inhibitory connection pattern from a map later labelled INHIBITION. Let's first describe the working memory system.

4) *FEF and WM maps*: FEF and WM maps implement a dynamic working memory system that is able to memorize stimuli that have already been focused in the past together with the currently focused stimulus. These maps are reciprocally connected so that the WM map reflects the activity of the FEF map and sends it back to FEF which in turn increases its activity. Outside this coupled system, the FEF map receives excitatory connections from both the VISUAL and FOCUS maps. Activity in the VISUAL map alone is not sufficient to generate activity in FEF; it needs a consistent conjunction of activity of both VISUAL and FOCUS to trigger some activity in FEF map. Since there is only one bubble of activity in the focus map, the joint activation of VISUAL and FOCUS only happens at the location of the currently focused stimulus. So, when the system starts, several bubbles of activation appear in VISUAL map, only one emerges in FOCUS, what allows the appearance of the same bubble in FEF map. As soon as this bubble appears, it is transmitted to WM which starts to show activity at the location of that bubble which in turn excites the FEF map. This is a kind of reverberatory loop, where mutual excitation leads to sustained activity. One critical property of this working memory system is that once this activity has been produced, WM and FEF map are able to maintain this activity even when the original activation from FOCUS disappears. For example, when the system focuses on another stimulus, previous activation originating from the FOCUS map vanishes to create a bubble of activity somewhere else. Yet the previous coupled activity still remains, and a new one can be generated at the location of the new focus of attention. Importantly, the system is also sensitive to the visual input and thus allows memorized stimuli to have a very dynamic behaviour since a bubble of activity within FEF and WM tends to track the corresponding bubble of activity within the VISUAL map. In other words, once a stimulus has been focused, it starts reverberating through the working memory system which can keep track of this stimulus, even if another one is focused.

5) *Switching Sub-Architecture*: The mechanism for switching the focus in the FOCUS map is composed of several maps (REWARD, STRIATUM, GPI, THALAMUS and INHIBITION) grossly inspired by the architecture of the cortico-basal-thalamo-cortical loop [21]. The general idea is to actively inhibit locations within the focus map to prevent a bubble of activity from emerging at these locations. This can be performed in cooperation with the working memory system which is able to provide the information on which locations have already been visited.

The STRIATUM map receives weak excitatory connections from the FEF map, which means that in the normal case no

activity appears on STRIATUM map. But when the REWARD neuron (which sends a connection to each neuron in the STRIATUM) fires, it allows bubbles to emerge at the location they are potentiated by FEF. The REWARD activity is a kind of "gating" signal which allows the STRIATUM to reproduce or not the FEF activity. The STRIATUM map sends inhibitory connections to the GPI, which has the property to be tonically active: if the GPI neurons receive no input, they will show a great activity. They have to be inhibited by the STRIATUM to quiet down. In turn, the GPI map sends strong inhibitory connections to the THAL map, which means that when there is no reward activity, the THAL map is tonically inhibited and can not show any activity. It is only when the REWARD neuron allows the STRIATUM map to be active that the GPI map can be inhibited and therefore the THAL map can be "disinhibited". Note that this is not a reason for the THAL to show activity, but it allows it to respond to excitatory signals coming from somewhere else. This disinhibition mechanism is very roughly inspired by the structure of the basal ganglia, which are known as mediating selection of action [21]. It allows more stability than direct excitation of the THAL map by FEF. The INHIBITION map is reciprocally and excitatorily connected with the THAL map, in the same way as FEF and WM are. But the reverberatory mechanism is gated by the tonic inhibition of GPI on THAL. It is only when the REWARD neuron fires that this reverberation can appear. INHIBITION receives weak excitatory connections from FEF (not enough to generate activity) and sends inhibitory connections to FOCUS. The result is that when there is no reward, the inhibitory influence of the INHIBITION map is not sufficient to change the focus of attention in FOCUS, but when the REWARD neuron fires, INHIBITION interacts with THAL and shows high activity where FEF has stored previously focused locations, what prevents the competition in FOCUS to create a bubble at a previously focused location, but rather encourages it to focus on a new location.

B. Simulated Behaviour

As detailed in Figure 3, the dynamic of the behavior is essentially ruled by both the existing pathways between different maps (either excitatory or inhibitory) and the inner dynamic of neurons. For example, consider the case where the INPUT map is clamped such that it reflects the activity of three noisy bubbles at three different locations in the visual field. In Figure 3-a), the three noisy bubbles in map INPUT are somehow filtered out in the VISUAL map (by virtue of the lateral interactions), allowing only one bubble to emerge in the FOCUS map which is immediately stored in FEF and WM. In Figure 3-b), a switch signal is explicitly sent to the network via the REWARD unit, allowing the STRIATUM to be excited at the location corresponding to the unique memorized location in the working memory system. This striatal excitation inhibits in turn the corresponding location within the GPI map. In Figure 3-c), the localized destabilization of the GPI prevents it from inhibiting the thalamus at this same location and allow the inhibition map to activate itself, still at the same location. In

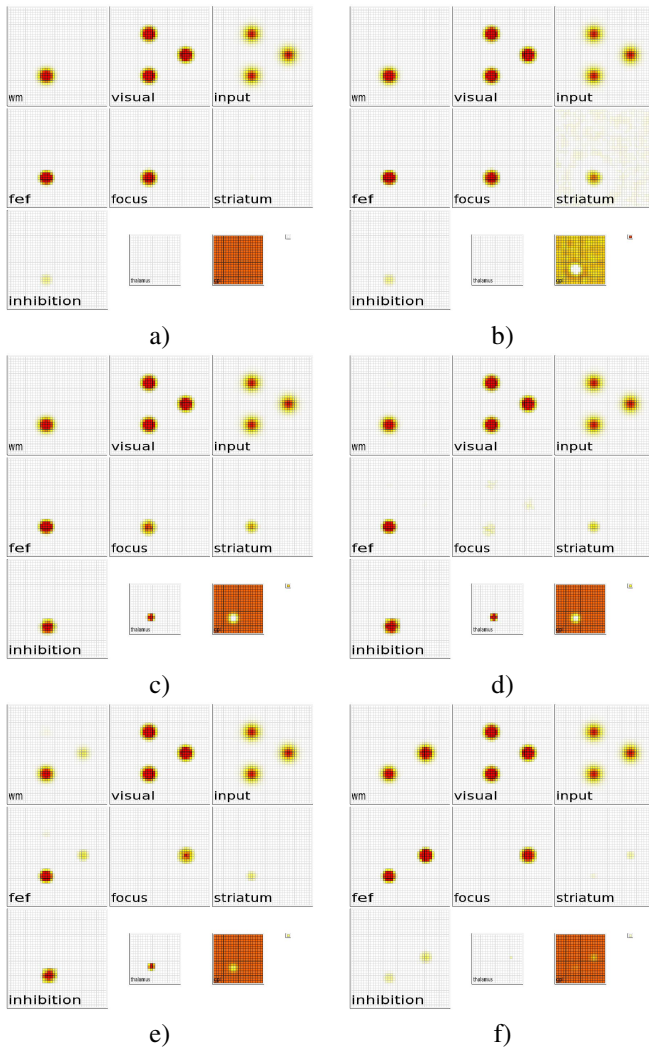


Fig. 3. A simulated sequence of focus switching. See text for details.

Figure 3-d), the INHIBITION map is now actively inhibiting the FOCUS map at the currently focused location. In Figure 3-e), the inhibition is now complete and another bubble of activity starts to emerge within the FOCUS map (precise location of the next bubble is unknown, it is only ensured that it can not be the previously visited stimulus). In Figure 3-f), once the focus is fully activated, it triggers the memorization of the new location while the previous one is kept in memory.

C. Experimental Results on a Robotic Platform

This work is part of the FET MirrorBot project (Biomimetic multimodal learning in a mirror neuron-based robot) which aims at studying emerging embodied representations based on mirror neurons (discovered by Rizzolatti et al. [22] in the monkey premotor cortex which fires both when the monkey looks at and performs an action) and implementing them in a robot to investigate the task of searching for objects. Therefore this model is not stated to be a generic model of visual attention, but rather an example of mechanism that can be used in visual search to ensure that attention is not attracted

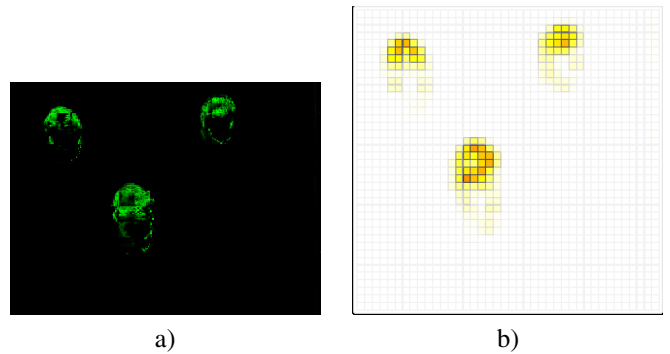


Fig. 4. a) A gaussian filter around the green colour (H=80 S=50 in HSV coordinates) is applied to the image to simulate the fact that green objects are attended. b) Activation in INPUT map.

twice to the same location when exploring a visual scene. As a consequence, we did not implement here any visual recognition, nor top-down influence on visual processing. We supposed that the input to our model is a kind of "saliency map" as in [18] representing the saliency of the visual scene, regardless of whether this saliency is purely "bottom-up" (i.e. due to the intrinsic properties of the objects) or "top-down" (i.e. influenced by task requirements). Such a map may be the equivalent of the area LIP as discovered by Gottlieb et al. [23], but this is still controversial.

The experiment we chose to validate our model is a task of sequential scanning of identical visual targets, for example green lemons, with the mobile camera device available on our Peoplebot platform. According to the premotor theory of attention, the mechanism involved in covert attention (without eye movement) should be the same as in overt attention (with eye movement). The fact is that this model works as well in covert orienting (like in the simulated behaviour) as in overt orienting. We therefore put the robot in front of three green lemons lying on a table. The task for the robot is to successively gaze at the three lemons without ever looking twice the same fruit. To simulate the saliency of the fruits on the image depending on the task requirements, we just applied a Gaussian filter centered on the HSV coordinates of the green lemons. The result is then fed into the INPUT map as shown in Figure 4. This filtering is very noisy but the lateral interactions in the different maps of the network suppress that noise.

Then, the only difference with the simulated sequence is that at each timestep we extract the position (relative to the image) of the unique bubble in the FOCUS map and transform it into a motor command to the camera so that the system progressively centers the attended fruit on the image. One important thing to notice here is that this command is differential, i.e. just a little percentage of the displacement needed to go to the target is actuated, then the network is updated with a new image and so on. We will discuss this limitation later.

An example of behaviour of the model is given in Figure 5. The center of gaze of the camera is first directed somewhere on the table. The model randomly decides to focus its attention on the bottom-right fruit (let's understand "randomly" as

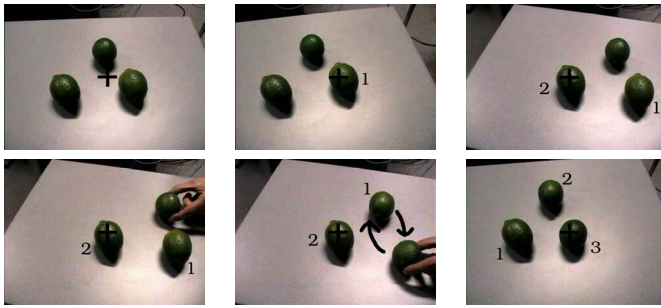


Fig. 5. Some snapshots of the sequence executed by the robot when trying to sequentially gaze at three green lemons. First, the robot initially looks at somewhere on the table. Then it gazes successively at fruit 1 and fruit 2. While fixating fruit 2, even if someone exchanges fruit 1 and the third not previously focused fruit, the robot will fixate the third “novel” fruit.

“depending on the noise in the input image, the initial state of the network and so on”) and step-by-step moves the camera to it. When the camera is on it, the user can decide whenever he wants to focus another fruit by clamping the reward neuron (in a biologically relevant scenario, the system would have to learn that he could obtain more reward by switching its focus and therefore make the reward neuron fire) which inhibits the currently focused object. The focus of attention then moves to one of the two remaining fruits (here the bottom-left one), what makes the camera gaze at it. At this point, the “working memory” system contains the current and the past focused fruits. If the user clamps again the reward unit, the new focused location will necessarily be on the third fruit, even if one slowly exchanges the locations of the first and the third fruit, because the representations in the working memory are updated by perception.

IV. CONCLUSION

Through localized, asynchronous and parallel computations, this model shows the emergence of a purely sequential function, here the sequential scanning of the salient locations in a real image despite noise, target positions and movements, lightening conditions etc. This emergence is only the consequence of the inner dynamics of the units and of the architecture of the system (the links between the units). This architecture is not meant to model precisely the actual structure of the brain (even if some names are not randomly chosen, especially for the switching mechanism) but rather to show that a unique substrate (a map of units with the same dynamics) can be involved in a given function without ever being explicitly specialized to a certain sub-problem of the task. The major problem encountered by this model is the fact that the motor commands have to be differential to allow sensory processing during the movement, what is inconsistent with physiological findings. Furthermore, if two salient moving objects cut each other, the model has no means to decide which one has to be attended after the occlusion. As a consequence, our current research relies on better coupling of this architecture with a feature-extraction system to bind recognition and localization of visual targets using synchronized neural assemblies [24].

REFERENCES

- [1] A. Treisman and G. Gelade, “A feature-integration theory of attention,” *Cognitive Psychology*, vol. 12, pp. 97–136, 1980.
- [2] M. I. Posner, “Orienting of attention,” *Quarterly Journal of Experimental Psychology*, vol. 32, pp. 3–25, 1980.
- [3] A. Treisman, “Features and objects: The bartlett memorial lecture,” *The Quarterly Journal of Experimental Psychology*, vol. 40, pp. 201–237, 1988.
- [4] J. Moran and R. Desimone, “Selective attention gates visual processing in the extrastriate cortex,” *Science*, vol. 229, pp. 782–784, 1985.
- [5] J. Vitay, N. P. Rougier, and F. Alexandre, “A distributed model of spatial visual attention,” in *Neural learning for intelligent robotics*, S. Wernter and G. Palm, Eds. Springer, 2005.
- [6] J. W. DeFockert, G. Rees, C. D. Frith, and N. Lavie, “The role of working memory in visual selective attention,” *Science*, vol. 291, pp. 1803–1806, 2001.
- [7] S. M. Courtney, L. Petit, J. M. Maisog, L. G. Ungerleider, and J. V. Haxby, “An area specialized for spatial working memory in human frontal cortex,” *Science*, vol. 279, pp. 1347–1351, 1998.
- [8] H. R. Wilson and J. D. Cowan, “A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue,” *Kybernetik*, vol. 13, pp. 55–80, 1973.
- [9] J. Feldman and J. Cowan, “Large-scale activity in neural nets. i. theory with applications to motoneuron pool responses,” *Biological Cybernetics*, vol. 17, pp. 29–38, 1975.
- [10] S.-I. Amari, “Dynamical study of formation of cortical maps,” *Biological Cybernetics*, vol. 27, pp. 77–87, 1977.
- [11] J. G. Taylor, “Neural bubble dynamics in two dimensions: foundations,” *Biological Cybernetics*, vol. 80, pp. 5167–5174, 1999.
- [12] R. J. Douglas, C. Koch, M. Mahowald, K. A. Martin, and H. H. Suarez, “Recurrent excitation in neocortical circuits,” *Science*, vol. 269, pp. 981–985, 1995.
- [13] S. Deneve, P. Latham, and A. Pouget, “Reading population codes: a neural implementation of ideal observers,” *Nature Neuroscience*, vol. 2, pp. 740–745, 1999.
- [14] K. Zhang, “Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory,” *Journal of Neuroscience*, vol. 16, pp. 2112–2126, 1996.
- [15] S. Deneve, P. Latham, and A. Pouget, “Efficient computation and cue integration with noisy population codes,” *Nature Neuroscience*, vol. 4, no. 8, pp. 826–831, 2001.
- [16] S. M. Stringer, E. T. Rolls, and T. P. Trappenberg, “Self-organising continuous attractor networks with multiple activity packets, and the representation of space,” *Neural Networks*, vol. 17, pp. 5–27, 2004.
- [17] N. Rougier and J. Vitay, “Emergence of attention within a neural population,” *Submitted*, 2004.
- [18] L. Itti, “Visual attention,” in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., M. A. Arbib, Ed. MIT Press, 2003, pp. 1196–1201.
- [19] M. I. Posner and Y. Cohen, “Components of visual orienting,” in *Attention and Performance*, H. Bouma and D. Bouwhuis, Eds. Erlbaum, 1984, vol. X, pp. 531–556.
- [20] S. P. Tipper, J. C. Brehaut, and J. Driver, “Selection of moving and static objects for the control of spatially directed action,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 16, pp. 492–504, 1990.
- [21] O. Hikosaka, Y. Takikawa, and R. Kawagoe, “Role of the basal ganglia in the control of purposive saccadic eye movements,” *Physiological Reviews*, vol. 80, no. 3, pp. 953–978, 2000.
- [22] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi, “Premotor cortex and the recognition of motor actions,” *Cognitive Brain Research*, vol. 3, pp. 131–141, 1996.
- [23] J. P. Gottlieb, M. Kusunoki, and M. E. Goldberg, “The representation of visual salience in monkey parietal cortex,” *Nature*, vol. 391, pp. 481–484, 1998.
- [24] A. K. Seth, J. L. McKinstry, G. M. Edelman, and J. L. Krichmar, “Visual binding through reentrant connectivity and dynamic synchronization in a brain-based device,” *Cerebral Cortex*, vol. 14, pp. 1185–1199, 2004.

ACKNOWLEDGMENT

The authors wish to thank the FET MirrorBot project and the Lorraine Region for their support.