Modeling attentional loop in the insect Mushroom Bodies

Paolo Arena, Luca Patané and Pietro Savio Termini

Abstract—Insects show advanced capabilities and a rich behavioral repertoire in task solving and thus they are becoming a reference point in Neuroscience for studying simple cognitive structures. In particular, thanks to many neurogenetic tools, the fruit fly Drosophila melanogaster became a relevant source of inspiration for Robotics. Mushroom Bodies (MBs) are very interesting neural structures involved in the regulation of behaviors in fruit fly, even if their main role regards olfactory conditioning. In this paper a novel bio-inspired neural architecture is presented, where the MBs role in attention tasks is focused. The model is a multi-layer spiking neural network where the MBs and their direct and or indirect interactions to other key elements of the insect brain, the Central Complex and the Lateral Horns, are modeled. The biological background of the proposed model is presented together with a detailed description of the architecture; simulation results and remarks on the biological counterpart are also reported.

I. INTRODUCTION

Insects are a very interesting case of study regarding their behavioral repertoire and learning capabilities. In spite of the tiny number of neurons in their brain, insects are able to process a variety of sensory modalities and to solve an impressive number of even non elementary tasks [1], [2]. In the last years, thanks to new tools from neurogenetics, Neuroscience provided interesting information regarding the neural mechanisms involving behavioral choices and sensorial processing in insect brain. As a consequence, insects are becoming a very interesting case of study also in robotics. Models and algorithms inspired by insects have been designed and implemented, with the main aim of making autonomous intelligent machines [3], [4], [6], [5], [7], [8]. Certainly the fruit fly Drosophila melanogaster is, up to now, one of the most studied insect, and a lot of work has been done in order to model and analyze the functionalities of the different parts of its brain [9]. The Central Complex (CX) and the Mushroom Bodies (MBs) are the most analyzed structures of the Drosophila melanogaster brain. The CX resides at the midline of the brain between the protocerebral hemispheres. In Drosophila, the CX is composed of about 2000 to 5000 neurons. The MBs are a paired structure of the insect protocerebral hemispheres. The most important constituents of the MBs are the 2500 Kenyon cells (KCs) per side which run in parallel from the calyx through the peduncle and to the lobes. In flies, there is a prominent olfactory input from the antennal lobes into the caliciles [10], [11], [13], [14]. Input from other sensory modalities is not topologically identified in Drosophila, but the role of MBs in tasks related to vision, decision making and behavioral adaptation have been reported [15], [16]. In the insect brain, MBs interact with Lateral Horns (LHs) and Antennal Lobes (ALs). MBs are responsible for odors representation. In our work we consider a particular case of behavioral alteration, supposedly linked to MBs in Drosophila, that regards the filtering of the sensorial inputs coherently with the attention towards a visual stimulus. In fact, recent works have shown that mutations affecting olfactory memory formation in Drosophila also produce distinct defects in visual attention-like behaviors [17], [18], [19]. In this paper a model inspired by the Drosophila brain is presented. The model is a multilayer spiking neural network. The first layer is a visual sensorial layer and it has been modeled as a lattice of neurons, in which each neuron codifies a particular feature of the visual input. A locally competitive topology is here implemented. It can be noticed that biological evidences indicate the CX as the place of high visual processing and visual memory in Drosophila brain, and that no direct connection is found between CX and MBs. However, indirect connections that can involve other neural assemblies are commonly considered plausible in Neurobiology [20]. The first layer of our model can be considered as a simplification of such interactions. The input layer randomly projects connections to the MBs model, which is composed of two distinct subnetworks, namely the $\alpha - /\beta -$ lobes layer and the $\alpha' - /\beta' -$ lobes layer. Feedback connections are present from the MBs layers to the input layer. The LHs model, as suggested from neurobiology [21], has been thought as an input-triggered system that provides a delayed global inhibition to the MBs network. The actual model is an extension of a previous work, successfully used to model expectation-based processes in the insect olfactory system [22]. The paper is divided into five sections: Section II presents the biological background of the work. In Section III the proposed architecture is introduced and described. Section IV presents some simulation results and also provides comparisons with the biological counterpart. Finally, Section V draws the conclusions.

II. BIOLOGICAL BACKGROUND

The CX resides at the midline of the insect brain, between the protocerebral hemispheres. It is composed of four main regions [23]. The fan-shaped body (FB) is the largest neuropil which possesses eight fans in the latero-lateral extent, six layers in the dorso-ventral direction and four shells in the anteriorposterior direction. The ellipsoid body (EB) has a perfect toroid shape and possesses 16 segments and it has two radial zones. The paired noduli reside ventral to the FB just dorsal to the oesophagus. The protocerebral bridge (PB) is a rod-like neuropil composed of a chain of 16 glomeruli just posterior to the FB. Because there are no
known projections from primary sensory areas to the CX, the input into the CX seems to be highly preprocessed. Experimental evidences from many different insect species demonstrate that primarily visual information is processed in the CX [24], [25]. More details can be found in [26]. The MBs are a paired structure of the insect protocerebral hemispheres. The MBs "frame" the CX without known direct connections. In flies, there is a prominent olfactory input from the antennal lobes into the calicai. In fact, MBs and LHs are neural structures of the insect brain able to codify the spatio-temporal information coming from the glomeruli of ALs. Connections from LHs and MBs have been found; the entity of these connections in Drosophila are not well known, but in locusts, LHs offer an inhibitory effect to the MBs neurons [21]. Input from other sensory modalities is not obvious in Drosophila, but tasks related to vision, where MBs are involved, were identified for fly experiments. In honeybees, MBs receive prominent visual [11], gustatory, and mechanosensory [12] input. In flies and bees, the MBs lobe region receives information on sugar reward or electric shock. There is an output of the MBs to pre-motor areas of the brain. Inside the MBs the flow of information is through the MBs cells from the calyx towards the lobes. Very recently, recurrent connections between MBs and ALs have been found [27]. The presence of this functional feedback from the MBs to the ALs suggests top-down modulation of olfactory information processing in Drosophila. The presence of dynamically changing conditions in the environment leads animal to develop attention process. Attention facilitates to focus on the attended events, while filtering out irrelevant information. These interesting processes have been studied in Drosophila [17], [18], [19], [20], where mutations affecting olfactory memory formation also affects the visual attention-like behavior. In particular, results in [20] suggest that MB in flies behaves like a sensory gain controller, allowing the processing of salient signals, filtering out the background noise and distracting signals.

III. THE PROPOSED NEURAL ARCHITECTURE

Taking inspiration from the insect brain and from the different behavioral aspects and learning capabilities of the MBs, here a neural architecture mainly based on spiking neurons is proposed. This architecture is mainly constituted by three interacting lattices. The input lattice is derived from the capabilities of the CX. It is able to decompose and represent visual stimuli through features. The input layer randomly projects to the MBs $\alpha - /\beta -$ and $\alpha' - /\beta'$ - lobes. Each lobe is modeled as a spherical lattice with clustering capabilities. Delayed feedback synapses are present from the $\alpha - /\beta -$ lobes to the input layer. These connections act as a filter for the sensory inputs on the basis of the previously followed object. A block diagram of the architecture is shown in Fig. 1, where the different networks, their interactions and the main role are summarized.

A. The Neural Model and Learning Algorithm

The spiking networks used to model the insect brain blocks are based on the Izhikevich spiking neuron, proposed in [28]. Izhikevich’ neural model is well known in literature and offers many advantages from the computational point of view. The model is represented by the following differential equations:

$$
\begin{align*}
\dot{v} &= 0.04v^2 + 5v + 140 - u + I \\
\dot{u} &= a(bv - u)
\end{align*}
$$

with the spike-resetting

$$
\begin{align*}
&\text{if } v \geq 0.03, \text{ then } \\
&v \leftarrow c \\
u &\leftarrow u + d
\end{align*}
$$

where $v$ is the membrane potential of the neuron, $u$ is a recovery variable and $I$ is the synaptic current. The value used for the parameters are different between the ALs and MBs. In the first case the Tonic Spiking model has been used whereas to model the KCs the parameters have been optimized to guarantee an efficient and robust clustering formation capability as better explained in the next sections. Neurons are connected through synapses. The synaptic model transforms the spiking dynamics of the pre-synaptic neuron into a current that excites the post-synaptic one. The mathematical response of the synapses to a pre-synaptic spike is ruled by the following equation:

$$
\epsilon(t) = \begin{cases} 
W(t)/\tau \exp\left(-t/\tau\right), & \text{if } t > 0 \\
0, & \text{if } t \leq 0
\end{cases}
$$

where $t$ is the time lasted from the emitted spike, $\tau$ is the time constant and $W$ is the efficiency of the synapse. This last parameter can be modulated through experience. The Spike-Timing Dependent Plasticity (STDP) can reproduce Hebbian learning in biological neural networks [29], [30].
The algorithm works on the synaptic weights, modifying them according to the temporal sequence of occurring spikes. The updating rule can be expressed by the following formula:

$$\Delta W = \begin{cases} A^+ \exp (\Delta t/\tau^+), & i f \ \Delta t < 0 \\ -A^- \exp (-\Delta t/\tau^-), & i f \ \Delta t > 0 \end{cases}$$

(4)

where $\Delta t$ is the time delay between pre and post synaptic spikes. In this way the synapse is reinforced if the pre-synaptic spike happens before the post-synaptic one, it is weakened in the opposite situation. Parameters $\tau_+$ and $\tau_-$ represent the slope of exponential functions, while positive constants $A_+$ and $A_-$ represent the maximum variation of the synaptic weight. Interesting applications of this learning paradigm to biorobotics, together with details on the parameters, are reported in [31].

B. The Sensory Layer

The CX resides in all insect brains and it is composed of four main regions [23]. In our work, the functionalities of one of these areas, the FB body, have been focused. Experimental evidences from Drosophila and many other different insect species demonstrate that primarily visual information is processed in the FB. In our model the Drosophila FB directly inspired the input layer. Even though direct connections between CX and MBs are not documented in biology, indirect connections are plausible [20]. Each neuron in the sensorial layer encodes a particular aspect related to a detected visual stimulus. Neurons within the input layer are organized in groups. Each group codifies a kind of feature, and neurons in the same group codify different values or intensity of that feature. Neurons in the same group are linked together through inhibitory synapses. This topology implies that, when the layer is stimulated, after a short transient time, only one neuron in each group remains excited, according to a winner-takes-all topology. Therefore, when an object is presented to the network, it is decomposed in its relevant features, and the corresponding neuron of each feature group remains excited. Neurons in different groups are linked together through plastic synapses, instantly reinforced for neurons encoding different features of the same object. The input layer projects excitatory direct connections to the MBs. Each neuron of this layer is connected to each KC with a given probability of connection, recalling the sparseness of connections in the corresponding areas of the insect brain.

C. Model of the Mushroom Body Lobes and their Interaction with the Lateral Horn

Like in the biological counterpart, the KCs of the MBs project, through the peduncle, to specific appendices, called lobes. They possess roughly the same topology, but were shown to serve different functions. In our model, one network represents the $\alpha - \beta$ – lobes, whereas another one represents the $\alpha' - \beta'$ – lobes. The two networks are connected randomly with the input layer. Each network was designed to show a cooperative-competitive dynamics: if excited, the neurons in this layer begin a competition and, after a transient, only one cluster of neurons will remain active and stable. The LH model resets both the networks each 120 ms. The lobes are connected to each other through two sets of plastic synapses, one from the $\alpha - \beta$ – lobes to the $\alpha' - \beta'$ – lobes and the other set from the $\alpha' - \beta'$ – to the $\alpha - \beta$ – lobes. It is known from Neurobiology that the neurons belonging to the two clusters are morphologically different. Moreover, whereas $\alpha - \beta$ – neurons give information back to the ALs, generating a feedback at the sensory stage, the $\alpha - \beta$ – neurons were found to provide no output signals back into the system, but they just receive sensory input in the last place at the level of the calices. So it was assumed in [32] that the information which arrives to these lobes is retained there and used as a kind of backup copy for memory purposes. The synapses between the lobe systems are reinforced when two clusters in different lobes are concurrently active. This structure was used in [32] to detect whether the object presented to the input layer remains the same in two subsequent steps. In fact, under these conditions, the plastic synapses between the lobe systems create a positive loop between the clusters in the two lobe systems: this allows the networks to increase the spiking rate of the active neurons. We assumed also to have a neuron sensitive to the firing activity of the $\alpha - \beta$ – lobes network to detect this situation. Two subsequent presentations of the same object generate a positive loop between the two lobe systems and a corresponding increase of spiking rate, whereas the successive presentation of a different object is identified due to the missing loop that does not boost the spiking neural activity in the lobes. We will assume that the feedback synapses are activated by the increase in the synaptic activity in the lobes. A biologically plausible scheme of the architecture
is shown in Fig. 2; the MBs are responsible for the gating of preprocessed visual inputs coming from CX even if the identification of direct connections between the two centers is still missing in Neurobiology. The dynamical evolution of the neural architecture is shown in Fig. 3. The features presented are acquired by the first layer and used as input for the $\alpha-\beta-$ lobes in the MBs, where a cluster arises. The second presentation of the same object leads to the detection of the sameness through the positive feedback among the $\alpha-\beta-$ and $\alpha'-\beta'-$ lobes. In the third step the presentation of multiple concurrent objects is managed by the system that, through the descending feedback connections between the $\alpha-\beta-$ lobes and the sensory layer is able to devote its attention to the previous detected object, filtering out the disturbing input. The main part of the proposed architecture was carefully studied together with Neurobiologists in [32] and was found able to show emergent capabilities to cope with the “delayed match-to-sample” issue, well known in insects. This network did not take into account the presence of feedback synapses from the lobes to the input layer. The introduction of this loop enables to elicit attentional capabilities, the main theme of this paper.

**D. The role of the feedback synapses**

A set of feedback synapses, from the $\alpha-\beta-$ lobes lattice to the input layer has been implemented in this augmented version of the MB network. These connections produce a pre-polarization of the input layer and act as a filter for the sensorial input, based on an attention-like paradigm. Feedback synapses are updated according to the STDP learning algorithm. The role of feedback connections in the insect olfactory system model has been analyzed in [22], on the basis of the biological evidences in [27]. In this paper two major important differences are reported with respect to [22]: the first one is that the spatial temporal dynamics that was assumed to link the different object internal representations in time is modeled as a specific function of the lobes, and the second aspect is that the output of the feedback synapses is delayed; the postsynaptic current influences the input layer only after the lobes resetting via LHs. This is equivalent to assume that the action of these feedback connections, being focused to enhance attentional loops, is able to persist also after the inhibition from LHs.

**IV. Simulation Results**

In this section the parameters of the network and the simulation results are presented.

**A. Network Parameters**

In the proposed simulations, the input layer consists of a 4x4 lattice of neurons. In particular, there are four groups of neurons ($f_1, f_2, f_3$ and $f_4$), and each group is made of four neurons (e.g. the neuron $j$ in the group $f_i$ will be called $f_{ij}$). Neurons in the same group are connected to each other through inhibitory synapses, with a synaptic efficiency $W_{FB} = -3$. Neurons in different groups that, after the presentation of an object, fire together, instantly increase the efficiency of the synaptic connection with a $\Delta W_{FB} = 7$. The initial value of such synapses is zero, and the synapses are reset each step. When an object is detected, the neurons of the first layer encoding the features of that object are excited with an input current $I_{FB} = 70pA$. Each neuron in the input layer has a probability $P = 0.25$ of being connected to KCs. The choice of the sparse connections between the first and the second layer is directly drawn from the sparse connections in the biological counterpart [21]. The main issue is related to the design of connections within each layer. The synapses between the first and second layer have a fixed efficiency of $W_{FB,KC} = 10$. In our architecture each lobe is a 9x9 lattice of neurons with a toroidal topology. The neurons in this layer are all connected to each other according to the paradigm of local excitation and global inhibition. A neighborhood of radius $r = 1$ is defined. In this way, each neuron is connected to the neurons in its neighborhood and with itself through excitatory synapses with an efficiency of $W_{KC} = 50$. It is also connected with the other neurons of the lattice through inhibitory synapses with an efficiency of $W_{KCF} = -30$. This set of connections gives the network its clustering capabilities [22]. The time constant used for fast excitatory and inhibitory synapses in the network is $\tau_{fast} = 1ms$. All the equations are solved using the Euler method with an integration time of 0.08 ms (i.e. one simulation step).

**B. Experimental Results**

In this section some simulation results are illustrated, in order to demonstrate how the proposed model works. In particular, three tests have been conducted. The first one aims to show the capabilities of the model in filtering the input stimuli according to the attention paradigm. The other experiments aims at showing the arousal of expectation-like effects in this model. This kind of feature was already presented in [22]. The same test is here presented to show that the upgraded model introduced in this paper retain the old capabilities, and extend its performance including new ones, based on a more strict biological background.
1) **Attention test:** This experiment shows the global capabilities of the network in reproducing an attentional filtering based on expectational feedback. After two subsequent presentations of the same object, two objects are contemporarily presented at the third step. In particular, we assume that one of the objects is the expected one, whereas the other one is a distractor. During the first step the neurons of the first layer representing the features of the presented object are excited and project to the $\alpha - /\beta -$ and $\alpha - /\beta -$ lobes. The $\alpha - /\beta -$ lobes elicit a winning cluster, while the $\alpha' - /\beta' -$ lobe will take the current from the first layer only during the next step, because of the delayed synapses. During the second step, the $\alpha - /\beta -$ are again excited by the current from the first layer and elicit the same cluster arose at the previous step, while the $\alpha' - /\beta' -$ lobes are excited by the current layer coming from the first layer at the first step. The loop-mechanism is then active [32] at the second step and the feedback synapses from the $\alpha - /\beta -$ lobes to the first layer are activated. At the third step a competition between the neurons representing the features of the two presented objects occurs in the first layer. The feedback synapses act as a bias on the features of the expected object. The evolution of the membrane potential of neurons representing expected and unexpected features are shown in Fig. 4. After a transient, the winner-takes-all topology allows the network to filter out the unexpected object; after a short competition during the transient, the activity of the input neurons representing the distractor is suppressed.

2) **Expectation-like test:** This simulation could be considered a preliminary test, in which the role of the feedback connections is underlined. After two subsequent presentations of the same object, no input is presented at the third step. During the initial two steps, the behavior of the $\alpha - /\beta -$ and $\alpha' - /\beta' -$ lobes can be reconducted to the previous experiment. During the third step, the sensory layer is not excited by any external current, but only by the $\alpha - /\beta -$ lobes through the feedback synapses. As shown in Fig. 5, the feedback connections allow the reconstruction of the features of the object expected in the input layer. In this case no competition happens in the first layer during the third step, because the only neurons that receive an excitatory current are the ones that represent the expected object. The spiking rate of the active neurons in this layer is, however lower than the spiking rate in the previous experiment at the third step.

3) **Reinforcing the same object:** After two subsequent presentations of the same object, again the same object is presented at the third step. In this case the feedback synapses reinforce an already present stimulus; in fact, the currents that excites the input neurons during the third step are the sum between the feedback current from the $\alpha - /\beta -$ lobes at the previous step and the currents from the sensory layer, that represent the actual sensorial stimuli. The spiking rate of the neurons in the first layer representing the features of the reinforced object is increased (Fig. 6) with respect to the situations shown in Fig. 5 and 4. Also in this case the role of the synapses in the first layer is not underlined by the simulation.
supposed from behavioral evidence that even the presence of an indirect link would suffice to reproduce some of the interesting capabilities involving multimodal integration. In this paper, the proposed model is able to reproduce an attention-like filtering of the sensorial input, which results to be of primary interest in mobile robotics, if visual inputs are taken into account. In fact simulation results show the robustness of the proposed architecture in presence of competing objects at the input layer, which could be used to enhance the capabilities of roving robots in the analysis of specific objects in cluttered environments, minimizing the effect of noise. The integration of this structure within the main blocks already designed in [9], will allow the realization an insect brain inspired neural architecture for applications in perceptual robotics.

ACKNOWLEDGEMENT

This work was supported by EU Project EMICAB, grant no. 270182. The authors thank Prof. R. Strauss and Dr. B. Zäpf, Institut für Zoologie III (Neurobiologie), University of Mainz, for useful discussions and comments.

REFERENCES


