A Spiking Neural Network Emulating the Structure of the Oculomotor System Requires No Learning to Control a Biomimetic Robotic Head

Praveenram Balachandar and Konstantinos P. Michmizos, Member IEEE

Abstract—Robotic vision introduces requirements for real-time processing of fast-varying, noisy information in a continuously changing environment. In a real-world environment, convenient assumptions, such as static camera systems and deep learning algorithms devouring high volumes of ideally slightly-varying data are hard to survive. Leveraging on recent studies on the neural connectome associated with eye movements, we designed a neuromorphic oculomotor controller and placed it at the heart of our in-house biomimetic robotic head prototype. The controller is unique in the sense that (1) all data are encoded and processed by a spiking neural network (SNN), and (2) by mimicking the associated brain areas’ connectivity, the SNN required no training to operate. A biologically-constrained Hebbian learning further improved the SNN performance in tracking a moving target. Here, we report the tracking performance of the robotic head and show that the robotic eye kinematics are similar to those reported in human eye studies. This work contributes to our ongoing effort to develop energy-efficient neuromorphic SNN and harness their emerging intelligence to control biomimetic robots with versatility and robustness.

I. INTRODUCTION

Covering all ranges of robotics, from structure [1], [2] and mechanics [3], [4] to perception [5], [6], actuation [7], [8] and autonomy [9], [10], biomimetic robots imitate design [2] and movement [8] principles found in nature to perform desired tasks in unstructured environments [6]. An orthogonal direction towards biomimesis is to imitate the most advanced biological controller: the brain. This particular type of brain-mimesis often entails the use of neural networks [11], of varying degrees of complexity [12], [13] and biological plausibility [14], [15], [16], that promise advances to Robotics [17] and insights to Brain Science [18], [19].

With robots being arguably the sweet spot for neuromorphic artificial intelligence, the emergence of neuromorphic chips [20], [21] has spurred interest for a bottom-up rethinking of biomimetic controllers in the form of spiking neural networks (SNN) that seamlessly integrate to non-Von Neumann architectures. We and others have recently proposed brain-inspired SNNs that embed into such chips and solve robotic problems with unparalleled energy-efficiency [22] while promising a robust yet versatile alternative to the brittle inference-based machine learning solutions [23]. The main criticism to neuromorphic computing is that, in the absence of a strong learning algorithm, the current state of the art does not share the same scaling abilities with the mainstream deep learning methods.

Alongside efforts to implement backpropagation algorithms in SNN [24], [25], most neuromorphic algorithms are indeed simple enough to be trained via variations of STDP, a Hebbian-type local learning rule [26], [27]. An alternative direction, that we have started to explore, entails the SNNs to be dictated by the underlying neural connectome associated with the targeted function [22], [28]. In this paper, we extend this direction by presenting a neuro-mimetic SNN that draws from the connectome of the human oculomotor system and enables our in-house robotic head to track a laser target. We demonstrate how the robotic prototype achieved real-time tracking of the visual target, by coordinating saccadic and pursuit eye movements with neck movements. Given that the structure of any network, be it biological or artificial, serves its function, the SNN’s behavior emerged out of this bottom up neural-level representation, without any training.

II. METHODS

A. The Robotic Head Prototype

The biomimetic robotic head prototype comprised of two cameras (eyes) and a neck (Fig. 1a). Each camera was mounted on a pan-tilt mechanism, which allowed horizontal and vertical movement of the eyes. The eyes were fixed to a base plate mounted on top of another pan-tilt system, to replicate neck movements. Overall, the robot had 6 degrees of freedom (DOF). The three pan-tilt systems were controlled by Dynamixel AX-12A digital servos driven by an Arduino-based Arbotix-M controller. The controller relayed servo position deltas to the Arbotix-M robocontroller through a USB serial interface. The deltas were computed based on the location of the laser target in a foveated field of view. The range-of-motion (ROM) for the servos controlling the eyes was restricted to 100 (70) degrees horizontally (vertically) from the center, to keep the eye kinematics within the biologically plausible range. Similar restrictions for biologically plausible ROM were imposed on neck movement.

The target was projected on a wall at 55 cm away from the eyes, using a laser diode mounted on a separate Arduino-controlled pan-tilt system. The movement profiles of the laser and the eyes were calibrated with respect to a fixed reference position for estimating the tracking accuracy of the proposed controller, by using linear regression analysis libraries to determine a higher order polynomial relation between the target angle and servo position. The laser and servo positions were recorded and mapped to the eye and neck kinematics, as well as target position.
Fig. 1: (a) The robotic head prototype fixating at a moving laser target projected on a wall. A pan-tilt system controlled the position of the laser (not shown). (b) The Receptive Field (RF) of the superior colliculus (SC) neurons (artificial retina) shown as squares in the input image. An SC neuron was activated when the input (laser dot) was within their receptive field. (c) The variation of SC neurons’ weights with respect to their retinotopic pixel position. The same weight distribution applied to both horizontal and vertical distances. (d) The SNN connectivity for moving the eyes and the neck in the horizontal direction. The shown connectivity represents innervation of the eyes’ muscles for moving the eyes to the left and the neck to the right. The opposite movements were controlled by a sub-network of symmetric connectivity to the one shown here (omitted for illustrative purposes). (e) The SNN connectivity for moving the eyes and the neck in the vertical direction. The input SC neurons correspond to the neurons in the top half above the fovea and the bottom half above the fovea for both eyes. The input from the two eyes are used together to generate coordinated movement of both eyes upward or downward.

B. Spiking Neural Network: The Oculomotor Controller

Mimicking biology, we embedded into the eyes of the controller a foveated structure (Fig. 1b). The goal of the SNN was to keep the moving laser target within the “fovea” of each eye through coordinated movements of both eyes and the neck. It did so by emulating the distinct neurons and neural areas involved in pursuit and saccade eye movements as well as neck movement, described below. In that sense, the SNN’s target tracking behavior emerged from the structure of the network and, thereby, required no learning to operate.

The design of the SNN structure was based mostly on the connectome of the saccade system in the brain. Saccade control structures are grouped into two regions for vertical and horizontal movement control, namely the horizontal gaze center (PPRF) and vertical gaze center (riMLF), found in the pons region of the brainstem [29], [30], [31]. Taking inspiration from these structures, we designed the oculomotor SNN having two separate sub-networks, one serving as the controller for horizontal eye and neck movements (PPRF) and another being the controller for the vertical movements (riMLF) [32] (Fig. 1d, e). The structure of the SNN for vertical control was simpler than the horizontal one under the assumption that both eyes moved together and by the same amount.

Each camera frame was mapped to an input neuronal layer, so that a neuron fired when the laser activated any of the pixels within that neuron’s receptive field. Such neurons have been found in the superior colliculus (SC) [33]. The SC innervated the gaze centers which in turn innervated the extra-ocular muscles to shift the gaze towards the target. When the target was significantly away from the fovea, the controller also engaged the neck through the SC and the vestibular system, following experimental findings [34].

The SC neurons’ firing rate encoded the distance of saccadic eye movements. Specifically, stimulating neurons at the periphery resulted to larger saccades compared to stimulating neurons with receptive fields close to the fovea. In accordance with experimental findings [35], the weights of the synaptic connections between the SC neurons and the brainstem regions (Fig. 1c) were modeled as an increasing function of the distance (number of pixels) from the fovea (center of the frame).

For the horizontal saccade control, the SNN included several bursting neurons, namely long lead bursting neurons (LLBNs), excitatory bursting neurons (EBNs) and inhibitory bursting neurons (IBNs). Upon activation of SC neurons,
LLBNs and EBNs were triggered, initiating the oculomotor response to the target. Another class of neurons, omni-pause neurons (OPNs), controlled fixation and maintained the fixation on a target, by inhibiting the activity of EBNs and IBNs. The combined effect of excitation from the SC neurons and the selective inhibition from the OPNs and IBNs led to the movement of both eyes in the same direction towards the target. The conjugate eye movement is facilitated in the brain through the abducens nuclei and interneurons that trigger the response of the contralateral oculomotor nuclei [29], [30]. The latter neurons were modeled in the SNN as motor neurons (MNs).

Independent movement of the eyes, or dis-conjugate (vergence) movement, was also possible through the same structure. Several bursting neurons are found to be direction selective [36], contributing to the oculomotor response only in a specific direction. We modeled these neurons, denoted by S in Fig. 1, that received inputs from both the ipsilateral and contralateral sides and inhibited the activity that promoted conjugate eye movements. This mechanism for the independent control of the robotic eyes is also in alignment with experimental findings [30].

The effective output of the sub-networks for horizontal and vertical movement was translated into equivalent servo position deltas using a firing rate encoding scheme. The firing rate was computed over windows of 20 ms and then scaled to a servo position value between 0 to 1023, which was the range of valid positions for the AX-12A servos.

C. Reward-based Hebbian Learning

Although the SNN can control the robotic head with no training, we also included a biologically plausible learning mechanism, to see whether this would further refine the robot’s performance. Interestingly, saccadic amplitude in primates is known to adapt to both a bottom-up visual error signal [37] and a top-down behavioral (goal) signal [38]. Here, we introduced a bottom-up reward-based learning mechanism, based on Sejnowski’s Hebbian learning rule [39].

Reward-based learning relies on maximizing the reward signal associated with the network performing as expected, here when the target was on the fovea. A global reward signal promoted the SNN behavior in bringing the target onto the fovea. The synaptic adaptation was semi-local, i.e. the weight change depended on the global reward signal and the pre-synaptic and post-synaptic neuronal activities at the synapse. The reward value as a function of the position of the target on the frame is shown in Fig. 2. For every pair of pre-synaptic neuron \(j\) and post-synaptic neuron \(i\), reward-based Hebbian learning was defined as:

\[
\tau_e \frac{de_{ij}}{dt} = -e_{ij} + H(p_{pre,j},p_{post,i}) \quad (1)
\]

\[
\frac{dw_{ij}}{dt} = MH(p_{pre,j},p_{post,i})e_{ij} \quad (2)
\]

\[
M(t) = R(t) - <R> \quad (3)
\]

\(e_{ij}\) is the synaptic eligibility trace for the pair of neurons, \(w_{ij}\) is the weight of the synapse between those neurons, \(H\) is the Hebbian learning term and \(M(t)\) is the neuromodulator signal at time \(t\) denoting the difference between given and expected rewards. Here, we empirically estimated the expected reward \(<R>\) as the running average and the time constant \(\tau_e\) was chosen in the range of 1 sec, to bridge the delay between action choice and final reward signal. The Hebbian term \((H(p_{pre,j},p_{post,i}))\), was modelled based on the Sejnowski learning rule. This rule relies on the activities of pre- and post-synaptic neurons, as well as the rate of spikes of these neurons over a window. It is based on the idea that firing rate of the neurons vary around their mean values \(<v_i>\) and \(<v_j>\) and defined by:

\[
\frac{dw_{ij}}{dt} = \gamma(v_i - <v_i>)(v_j - <v_j>) \quad (4)
\]

where \(\gamma\) is the learning rate and \(v_{j,i}\) are the firing rates of the pre- and post-synaptic neuron, respectively.

The laser and the servo positions for each DOF of the robotic head was recorded at 45 Hz and the kinematics of the eyes with respect to that of the laser was studied with and without learning. Both repetitive and random pattern of target positions were used to train the controller and the mean accuracy was used as a measure of behavioral performance.

III. RESULTS

We validated the proposed SNN for its ability to track accurately a moving target, both without and with Hebbian learning. When we did not incorporate learning into the SNN, the vast majority of its weights were adapted by the relative strengths of the connections between neural areas found in experimental studies, and the rest weights were found by trial and error. The C++ code is available in Appendix B in [40]. For the kinematics, the sign of the angles represented position of the target with respect to the origin, defined as the center of the frame. For horizontal movement, the negative (positive) angles represented positions on the left (right) side of the origin. Similarly, for vertical movement, negative (positive) values represented positions below (above) the origin. For calculating the discrepancy between fovea and...
Fig. 3: (a) Kinematics of the eye with respect to the kinematics of the target. The target, shown in blue, is moved around through a sequence of positions on the wall demonstrating sudden changes in both horizontal and vertical directions. The eyes, shown in red and black, follow the target by making multiple saccade like movements. Top panel shows the horizontal kinematics of the eyes with respect to the target and bottom panel shows the vertical component of the movement. (b) Kinematics of the eyes for a sequence of target positions without learning. (c) Kinematics of the eyes for a sequence of target positions with reward-based hebbian learning.

For the no-learning SNN, the mean relative error for the eye position with respect to the target, averaged over both horizontal and vertical direction of both the eyes and neck, was $-0.685^\circ$, averaged over ten 2-minute tracking experiments, for randomly moving targets. An example of the kinematics of the prototype with respect to the kinematics of the laser target on the wall is shown in Fig. 3a. We observed small, jerk-like, eye movements, similar to the experimentally reported miniature versions of voluntary saccades, named microsaccades. We attribute this behavior to the property of the SNN controller, which tries to fixate on the target within the bounds of the fovea: When the dimensions of the fovea were smaller than the target size, the controller constantly readjusted itself, attempting to fit the target within the fovea. This oscillatory activity was particularly evident in the horizontal position, partly because of the larger complexity of the underlying network, compared to the one controlling the vertical movements. Interestingly, the vertical movement exhibited a delayed response (lower bandwidth), leading to lower accuracy for vertical tracking compared to horizontal tracking. This was further corrected by allowing reinforcement learning.

The introduction of reward-based learning into the SNN resulted to a better tracking ability (Fig. 3c). Learning allowed for a noticeable improvement in tracking for targets with fast vertical components. To quantify the tracking abilities in both cases, we presented a random sequence of target positions to both controllers (Table I.)

### IV. Discussion

Here, we introduced an SNN oculomotor controller and its integration to our in-house robotic head prototype. The kinematics of the robotic eyes were remarkably similar to those of the human gaze in tracking a moving target [38]. In that sense, the goal of this work was to demonstrate that “machine behavior” in general, and robotic function in particular, can emerge naturally from an \textit{a-priori} knowledge that dictates the controllers structure. By drawing from the connectome of the brain areas associated with the targeted behavior, this and other efforts to develop biologically realistic SNNs can help advance Brain Science and Robotics, two fast growing fields that are also converging via biomimicry and neuromorphic computing [41], [42], [22], [28].

In the technical domain, adding biological constraints to an SNN structure removes the need for assuming all-to-all initial connectivity for the trainable network. This may translate to further improvements in training efficiency, as it limits learning to a small number of synaptic connections. In addition, contrary to the typical neuron models in deep networks that can be optimized to perform complex computational tasks [43], spiking neuron models have a non-differentiable output (their all-or-none firing) and therefore are incompatible with standard gradient-descent supervised learning methods [44].

| Eye Kinematics            | RE   | RE-HL | RMSE | RMSE-HL |
|---------------------------|------|-------|------|---------|
| Left Eye - Horizontal     | $-1.88^\circ$ | $-1.487^\circ$ | $3.55^\circ$ | $2.87^\circ$ |
| Right Eye - Horizontal    | $2.115^\circ$ | $2.049^\circ$ | $3.93^\circ$ | $3.44^\circ$ |
| Left Eye - Vertical       | $-1.210^\circ$ | $-0.699^\circ$ | $3.14^\circ$ | $3.02^\circ$ |
| Right Eye - Vertical      | $-0.823^\circ$ | $-0.421^\circ$ | $3.03^\circ$ | $2.95^\circ$ |
In the absence of a strong learning algorithm, the main criticism to neuromorphic solutions is that promising preliminary results [22] cannot share the same scaling abilities with the mainstream deep learning approaches. Here we show the first fruits of our efforts towards scalable SNNs that, by being able to host biological principles of computation known to be critical for intelligence, can give end-to-end neuromorphic solutions towards fully autonomous systems.

Effective as they may have become, robots still cannot duplicate a range of human behaviors, such as dynamically responding to changing environments using error-prone sensors. To operate in a real-world environment, an autonomous robot should 1) be robust to a noisy neural representation, 2) adapt to a fast changing environment, and 3) learn with no or limited supervision or reinforcement. The embodiment of SNNs into robots has been rather sparse and the current approaches aim to give a proof of concept [12], [45], [46], [16], rather than a whole-behavioring robot. While there is definitely value in studying simplified tasks and basic sensory representations [47], there is an ongoing need to propose new controllers capable of naturally handling richer, noisier and more complex scenarios [48]. This work suggests that a promising path towards duplicating a human-like behavior is to draw knowledge from how synergy is achieved in neurons across the implicated brain areas. In addition, contrary to the constantly online processing taking place in the oculomotor system, the traditional learning algorithms rely on separate training and inference phases. That is why we employed unsupervised learning, which is a better fit for lifelong learning of a continuously evolving network that can adapt to new targets and movement patterns.

The re-emergence of neuromorphic computing calls for a bottom-up rethinking of computational algorithms that can seamlessly integrate into non-Von Neumann hardware [22], promising unparalleled energy-efficiency and a robust yet versatile alternative to the brittle inference-based AI solutions [49]. The proposed work brings us closer to realizing this promise by tackling a robotic task where energy-efficiency may become crucial and controllers can scale, in order to exploit spatiotemporal context and commonsense understanding.

In the scientific domain, this work paves the way for a new direction, that of drawing from neuroscience biological knowledge and translating it to computational primitives applicable to SNN. Despite being equipped with biologically realistic models of neurons, current SNN can only offer weak suggestions on the underlying neural mechanisms that give rise to the targeted behavior. Recent applications of gradient-descent alternatives to SNNs [24], [25] promise to introduce SNNs to scalable problems, but they inherit the main limitations that deep networks have. For example, learning through backpropagation will basically match the networks input to its output, much regarding and, thereby, structuring the network as a black-box. Our bottom-up approach can spur the development of neural-controlled robots as test-beds that may be used to inform brain scientists on how the neural system could be structured to function properly. This paper opens up a fascinating possibility for artificial networks to be constrained by the structure of their biological counterparts. Alongside research on the computational mechanisms of biological learning [50], such efforts can introduce new push-pull dynamics between robotics and neuroscience by exploring and exploiting interpretable connections between neurophysiology and behavior.

Finally, this work not only aims to catalyze efforts towards replicating, if not augmenting, human cognitive abilities and engraving them into intelligent robotic assistants. It also gives an educational direction as enabling students across all ages to interact physically with a robot ignites their interest and understanding on its underlying working mechanisms.

V. CONCLUSION

Overall, the paper introduces an alternative approach towards designing robot controllers, that of developing SNNs inspired by the structure of the brain areas associated with the targeted behavior. Here, target tracking was achieved by emulating at a reasonable scale the connectome and the underlying types of the associated neurons, not through learning. This result suggests that building neuro-inspired controllers for this and other types of autonomous robotic behavior is a direction worth pursuing.

REFERENCES

[1] R. S. Fearing, K. H. Chiang, M. H. Dickinson, D. Pick, M. Sitti, and J. Yan, “Wing transmission for a micromechanical flying insect,” in Robotics and Automation, 2000. Proceedings. ICRA’00. IEEE International Conference on, vol. 2. IEEE, Conference Proceedings, pp. 1509–1516.
[2] Y. Sato, Y. Hiratsuka, I. Kawamata, S. Murata, and S. I. M. Nomura, “Micrometer-sized molecular robot changes its shape in response to signal molecules,” Science Robotics, vol. 2, no. 4, 2017.
[3] P. Birkmeyer, K. Peterson, and R. S. Fearing, “Dash: A dynamic 16g hexapedal robot,” in Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on. IEEE, Conference Proceedings, pp. 2683–2690.
[4] J. Zhao, J. Xu, B. Gao, N. Xi, F. J. Cintrn, M. W. Mutka, and L. Xiao, “Msu jumper: A single-motor-actuated miniature steerable jumping robot,” IEEE Transactions on Robotics, vol. 29, no. 3, pp. 602–614, 2013.
[5] X. Deng, L. Schenato, W. C. Wu, and S. S. Sastry, “Flapping flight for biomimetic robotic insects: Part i-system modeling,” IEEE Transactions on Robotics, vol. 22, no. 4, pp. 776–788, 2006.
[6] A. Ramezani, S.-J. Chung, and S. Hutchinson, “A biomimetic robotic platform to study flight specializations of bats,” Science Robotics, vol. 2, no. 3, 2017.
[7] D. W. Haldane, M. M. Plecnik, J. K. Yim, and R. S. Fearing, “Robotic vertical jumping agility via series-elastic power modulation,” Science Robotics, vol. 1, no. 1, 2016.
[8] A. J. Ipsen, “Biorobotics: Using robots to emulate and investigate agile locomotion,” science, vol. 346, no. 6206, pp. 196–203, 2014.
[9] W.-S. Chu, K.-T. Lee, S.-H. Song, M.-W. Han, J.-Y. Lee, H.-S. Kim, M.-S. Kim, Y.-J. Park, K.-J. Cho, and S.-H. Ahn, “Review of biomimetic underwater robots using smart actuators,” International journal of precision engineering and manufacturing, vol. 13, no. 7, pp. 1281–1292, 2012.
[10] V. Kopman and M. Porfiri, “Design, modeling, and characterization of a miniature robotic fish for research and education in biomimetics and bioinspiration,” IEEE/ASME Transactions on mechatronics, vol. 18, no. 2, pp. 471–483, 2013.
[11] G. Bekey and K. Y. Goldberg, Neural Networks in robotics. Springer Science and Business Media, 2012, vol. 202.
[12] E. A. Antonelo, B. Schrauwen, X. Dutoit, D. Stroobandt, and M. Nuttin, “Event detection and localization in mobile robot navigation using reservoir computing,” in International Conference on Artificial Neural Networks. Springer, Conference Proceedings, pp. 660–669.
