Investigation of cerebro-cerebellar reward-based mechanisms on a virtual neurorobot

Francesco Jamal Sheiban
Dept. of Bioengineering (DEIB)
Politecnico di Milano
Milan, Italy
francescojamal.sheiban@mail.polimi.it

Alessandra Maria Trapani
Dept. of Bioengineering (DEIB)
Politecnico di Milano
Milan, Italy
alessandramaria.trapani@polimi.it

Alice Geminiani
Dept. of Brain and Behavioral Sciences
University of Pavia
Pavia, Italy
alice.geminiani@unipv.it

Egidio Ugo D'Angelo
Dept. of Brain and Behavioral Sciences
University of Pavia
IRCCS Fondazione Mondino
Pavia, Italy
egidiougo.dangelo@unipv.it

Alessandra Pedrocchi
Dept. of Bioengineering (DEIB)
Politecnico di Milano
Milan, Italy
alessandra.pedrocchi@polimi.it

***Abstract*—Combining computational neuroscience tools with robotic devices allows to reproduce and investigate the interactions between embodied functional brain models and sensory-rich environments, integrating neural activity and behavioral data. As these models seldom include faithful biological features, their development can be used to suggest and test hypotheses on neurophysiological or pathological mechanisms, eventually impacting on challenging clinical applications. This project implemented a virtual neurorobotic experiment aimed at investigating the functional role of cerebro-cerebellar interactions in motor learning tasks, specifically designed to clarify the contribution of different brain areas in motor preparation and execution during rewarded goal-oriented actions.**

***Index Terms*—Neurorobotics, reward-based learning, computational neuroscience**

I. Introduction

Adaptive behavior in biological organisms results from interactions among brains, bodies, and environments [1]. Neurorobotics allows to incorporate features of neuroanatomy and neurophysiology within robotic devices to generate biologically-comparable experimental data to study such mechanisms through supervised protocols. In particular, a neurorobotic device is a device that engages in a behavioral task, is situated in a structured environment and whose behavior is controlled by a simulated nervous system having a design that reflects, at some level, the brain's architecture and dynamics [1]. Thus, neurorobotic models allow to develop and test theories of brain-environment interactions with devices either implemented with hardware solutions or reconstructed via software. This project was developed using the Neurobotics Platform (NRP), an integrated software toolkit developed within the Human Brain Project (HBP) specifically aimed at allowing researchers to design and execute neurorobotic experiments with simulated robots using customized brain models [2].

II. Materials and Methods

A. The Neurorobotics Platform

To simulate behaviors, the NRP combines 3D physical environment reconstructions with realistic brain models based on spiking neural networks (SNN), whose information-processing mechanisms mimic the action potentials of biological neurons. Connections between these two components are implemented with transfer functions translating either the robot sensory information to brain model inputs (robot to neuron (R2N) functions) or neural network outputs to robotic motor commands (neuron to robot (N2R) functions). Therefore, this project required the implementation of both a virtual environment and a neural network model, with their respective transfer functions, to simulate the execution of a reward-based behavioural task by the robotic subject.

B. Behavioural task and virtual environment

The behavioural task that the robotic subject engages in is a reach-to-grasp associative task: in standing position, the robot places its hands on a resting bar and waits for a directional somatosensory stimulus, modelled as a rotating bar that touches the robot left/right shoulder. After this anticipatory signal, the robot waits without moving for a visual go-cue (given by a color-changing screen) after which it is required to reach one of the two cylindrical objects in front of it, mirroring the direction of the somatosensory stimulus: if no reaching contact is detected after a certain time delay, the task fails and the robot has to place its arms back on the resting bar to start a new trial, otherwise the task is completed and the subject receives a reward input signal. Among the available NRP robotic models, the iCub humanoid robot was chosen [3], being able to perform forelimb movements involved in the protocol. The virtual environment setup (Fig. 1) included all the objects required for the task execution, embedding

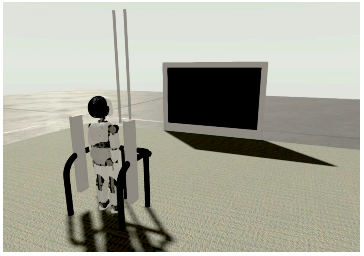

Fig. 1. **Virtual environment.** The virtual environment in which the neurorobot is situated includes a color-changing screen to deliver a visual go-cue, two floating cylinders representing the reaching movement targets and a custom-designed 3D model of a sensorized handrail (resting bar) with one rotating bar for each side, delivering a directional somatosensory stimulus on the neurorobot's shoulders.

different sensors to monitor the experiment phases completion; R2N transfer functions were implemented to carry the virtual stimuli from the robotic sensors to the brain model as well as N2R transfer functions to drive the robot through the SNN.

C. Brain model

The reconstructed brain model, designed following biological findings from literature, is composed by two identical modules to discriminate the directional stimuli involved in the protocol. Feedback loops between premotor and frontal cortices, motor thalamus and cerebellum implement short-term memory and temporal decisions mechanisms. More specifically, the directional stimulus is conveyed through the sensory cortex to the premotor cortex of the corresponding module. This area acts as an integrator [4], producing movements through signals to the motor cortex when reaching a spiking threshold, and engages in a feedback loop with the thalamus to store the directional stimulus information before the go-cue [5]. The spiking activity in this loop is prevented from letting the premotor cortex reach the threshold thanks to sustained inhibitory inputs from the medial prefrontal cortex, which effectively blocks impulsive actions [6]. The timing of this silencing mechanism is regulated by the cerebellum with thalamic projections that influence cortical areas [7]: by learning the association between the go-cue and the reward availability, it suppresses the medial prefrontal cortex upon the go-cue stimulus, allowing the premotor cortex to surpass the threshold, causing the robot to perform the reach.

III. Results

With all the necessary NRP components implemented, experiments were carried out controlling the correct flow of sensory information and monitoring the spiking activity of the brain simulation. We assessed that the proposed brain model was able to sustain preparatory activity and that the neurorobotic agent was able to carry out the behavioural protocol, once provided with the cerebellar contribution expected at the end of reward-based learning as direct input to the cortical loops in the SNN. Figure 2 shows a temporal plot of the spiking activity recorded from the network during the

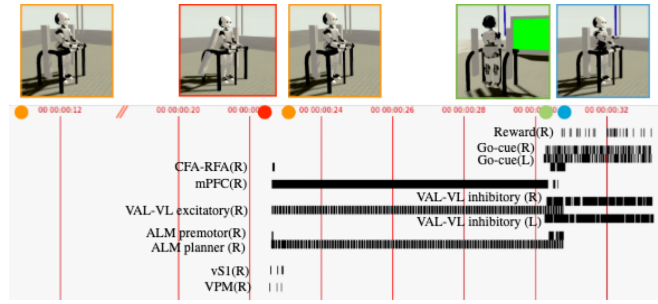

Fig. 2. **Spiking neural network activity during task execution.** The observed spiking activity demonstrates the proper functioning of the brain model, showing sustained working memory activity in the motor cortex (ALM planner) / thalamus (VL-VLA excitatory) loop and the medial prefrontal cortex (mPFC) suppressing the movement during waiting phase (yellow to green dot). The go-cue (green dot) silences the medial prefrontal cortex through thalamic inhibitory connections (VL-VLA excitatory), causing subsequent premotor and motor cortex (ALM premotor, CFA-RFA) firing. Spikes encoding reward signals are observed after a successful grasping movement performed by the neurorobot towards the water-delivering object on the right side (blue dot), mirroring the preparatory stimulus direction.

different task execution phases, depicted in the colored boxes and temporally marked by the colored dots.

IV. Discussion

The neurorobot correctly executing the protocol suggests that, although representing an initial simplified implementation, the proposed brain model describes a good scheme of the neural activity underlying the behavioral task in the different areas involved. The resulting network represents a novelty with respect to available neurorobots driven by single-module cerebellar networks [8] as it is driven by a multi-area brain model including cerebro-cerebellar loops, it critically involves thalamic contribution in preparatory activity and it is composed by two identical modules to take into account the directional-selectivity nature of the protocol. These results pose a solid basis on which to refine the neural network model (e.g. scaling up the network, including other brain areas such as the basal ganglia, and embedding distributed plasticity) and simulate the full learning protocol, eventually exploiting high-performance computing resources.

References

[1] A. K. Seth, O. Sporns, and J. L. Krichmar, "Neurorobotic Models in Neuroscience and Neuroinformatics," in Neuroinformatics, vol. 3, 2005.
[2] E. Falotico et al., "Connecting Artificial Brains to Robots in a Comprehensive Simulation Framework: The Neurorobotics Platform," in Frontiers in Neurorobotics, vol. 11, 2017.
[3] G. Sandini, G. Metta, and D. Vernon, "The iCub Cognitive Humanoid Robot: An Open-System Research Platform for Enactive Cognition," in 50 Years of AI, LNAI 4850, 2007.
[4] M. Murakami, H. Shteingart, Y. Loewenstein, and Z. F. Mainen, "Distinct Sources of Deterministic and Stochastic Components of Action Timing Decisions in Rodent Frontal Cortex," in Neuron, vol. 94, 2017.
[5] Z. V. Guo et al., "Maintenance of persistent activity in a frontal thalamocortical loop", in Nature, vol. 545, 2017.
[6] S. Killcross, and E. Coutureau, "Coordination of Actions and Habits in the Medial Prefrontal Cortex of Rats," in Cerebral Cortex, vol. 13, 2003.
[7] M. J. Wagner, and L. Luo, "Neocortex – Cerebellum Circuits for Cognitive Processing," in Trends in Neurosciences, vol. 43, 2020.
[8] C. Casellato et al., "Adaptive Robotic Control Driven by a Versatile Spiking Cerebellar Network," in PLoS ONE 9(11), 2014.