Insect-inspired body size learning model on a humanoid robot

Paolo Arena, Luca Patané, Dario Sanalitro and Alessandra Vitanza

***Abstract*— In this paper an insect-inspired body size learning algorithm is adopted in a humanoid robot and a control system, mainly developed with spiking neurons, is proposed. It implements an evaluation of distances by using the typical parallax method performed by different insect species, such as *Drosophila melanogaster*. A Darwin-OP robot was used as test-bed to demonstrate the potential application of the learning method on a humanoid structure. The robot, equipped with a hand extension, was free to move in an environment to discover objects. As consequence, it was able to learn, using an operant conditioning, which objects can be reached, via the estimation of their distance on varying the length of the equipped tool. The learning scheme was tested both in a dynamical simulation environment and with the Darwin-OP robot.**

I. Introduction

The research field in robotics is in continuous evolution and there are strong relations with the new achievements in understanding living beings from mammals to insects and plants. The multisensory representation of our body is a key element in the development of different higher level cognitive functions [1]. The capability to acquire knowledge about the peripersonal space is an important skill needed by autonomous systems to interact with objects in unstructured environments [2]. The body model can be defined as a sensorimotor representation of the body that can be used to accomplish selected tasks [3]. The body size learning is therefore the ability, shown by living beings, to learn the relations between objects and its own body. For instance, the system can acquire knowledge about either the reachability or the traversability of a point of interest by analyzing the effects of its own motion on the environment. This capability is directly connected to the vision system and is related to the acquisition of parallax motion.

In literature there are several studies that investigate the presence of this form of learning on different animal species including: pigeon and praying mantis [4], *Drosophila melanogaster* [5] and mammals like monkeys [6] and humans [7]. In a previous work we investigated the body size formation in insects with particular attention to *Drosophila melanogaster* and we developed a computational model inspired to the main neural centres devoted to the acquisition and formation of this knowledge [8]. That preliminary work is here extended by reformulating the developed model, including different processing layers and re-adapting the learning system to a humanoid robot, demonstrating that the proposed strategy can be applied for different bodies and scenarios. In particular, the considered task consists in learning the correct distance from which the robot is able to touch an object of interest using its arm. The robot can learn different body size models depending on the presence of a tool (e.g., an arm extension) that can change its reachability space. It may be worth mentioning that the possibility to learn its own reachability space (i.e., workspace) is an interesting ability to autonomously build internal representations arising from experience. Moreover, being a difficult task, learning techniques give a promising approach compared to the analytical and geometric methods. In latest years there has been an increasing attention on computational neuroscience methods focused on formulating bio-inspired models, algorithms and related structures such as spiking neural networks. These models enhanced the biological realism of control systems, alongside the possibility to use spatial and temporal information in the processes of computing, just as real neurons do. Indeed, we propose a bio-inspired solution by adapting a well-tested and investigated control system [8]. The control architecture is based on a spiking neural network developed to implement a distance estimation through parallax motion. The network consists of an ensemble of neurons able to process the visual information acquired by the robot in terms of angular positions and to determine a winning neuron that is representative of the estimated distance between the robot and the target of interest. The body size model is therefore conceived by the system using a learning process based on a threshold adaptation rule. This mechanism induces hyperpolarization or depolarization into the output neuron, to make it responsive only to the correct distances according to the knowledge acquired from its experience. A similar learning mechanism was already successfully applied in a different context, to neurons devoted to visual processing with the result to induce specialization in a group of robots [9], [10]. The threshold updating rule was determined using the data stored in a memory where events related to the visual sensory system were collected. Differently, in this work, the threshold adaptation rule is applied using the information acquired by the robot while performing actions in the environment in order to acquire the corresponding reward or punishment, following an operant conditioning schema [11].

The learning process was performed using a virtual simulation environment and subsequently, the learned controller was embedded on the Darwin-OP robot for a testing on a real arena. This humanoid robot was already successfully adopted, as a simple and efficient testbed, to evaluate a push recovery controller developed through reinforcement

Paolo Arena, Luca Patané, Dario Sanalitro are with the Dipartimento di Ingegneria Elettrica, Elettronica e Informatica, Universitá degli Studi di Catania, Italy (email: {parena, lpatane}@dieei.unict.it). Alessandra Vitanza is with the ICT, Innovation and Research Unit, IRCCS - Casa Sollievo della Sofferenza, San Giovanni Rotondo, Foggia, Italy. (email: avitanza@dieei.unict.it)

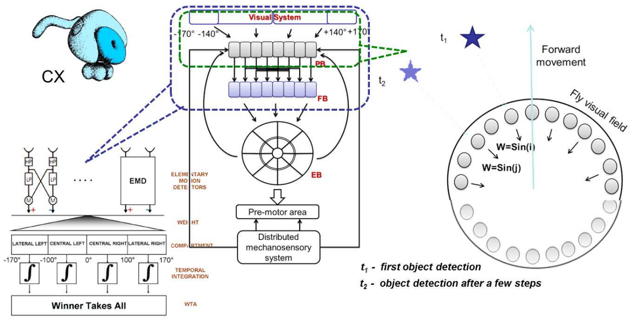


Fig. 1. Block scheme showing the different elements considered while modeling the CX. The visual system transfers information to the PB structure involved in body size learning. The FB participates to visual learning and orientation control whereas the EB is responsible for the formation of spatial memory.

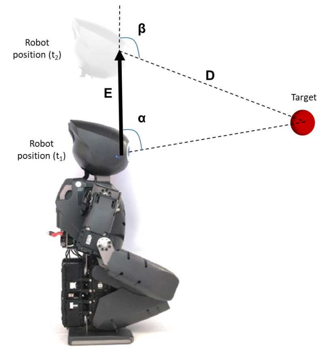


Fig. 2. Procedure for the acquisition of the parallax motion in the humanoid robot.

learning [12]. Herein, we are endowing the robot with a controller for visuo-motor coordination based on a learned body size model, which allows the robot to catch balls randomly distributed in the environment.

II. Biological background

The learning processes devoted to guide the formation of a body model involve an adaptive calibration of the proprioceptive information on the basis of which a representation of the own body can be created. This calibration is obtained by detecting the relations between the self-produced actions and the induced sensory information. The importance of this form of learning is demonstrated by the fact that even insects like *Drosophila melanogaster* with a number of neurons in the central brain $\simeq 10^5$ (compared with $\simeq 10^{11}$ in humans) possess this skill. The learning process requires the presence of visual feedback through parallax motion generated by self-motion for the body size formation [5]. In gap climbing scenarios with flies characterized by different body sizes, the number of unsuccessful attempts is maximum at the largest just surmountable gap width. This evidence can be explained only with the presence of a body size model available to be used for behavioral decisions otherwise the number of attempts should be the same independently from the gap width. The presence of parallax motion, as a fundamental stimulus for the formation of the body size model, was demonstrated by hatching flies in the dark. Under these conditions, flies are not able to develop a body model and the body size is no longer taken into account for behavioural decisions [13].

The central complex (CX) is a crucial neural structure for the body size knowledge formation [14], [15]. A plausible model for CX consists of a distributed structure, spatially associated to the visual field, able to acquire and process spatial information in terms of orientation angle between relevant targets and the insect.

In Fig. 1 the different elements identified in this neural assembly are reported. The visual system acquires information on relevant objects in the scene, transferring the *where* and *what* to the Protocerebral bridge (PB) and the Fanshaped body (FB) that are responsible for heading control and visual learning, whereas the Ellipsoid body (EB) is involved in spatial memory formation. The designed model, initially developed for direction control, spatial memory and other capabilities [16], [17], was extended to include body-size learning processes.

III. Body size model

Operatively, to create a body size model we need a distance estimation procedure based on parallax motion. Fig. 2 shows the movements performed by the robot to acquire the α and β angles needed for the distance estimation. Although in *Drosophila melanogaster* the parallax motion is acquired when walking in the environment; on the humanoid robot, a sufficiently reliable straight forward trajectory between two consecutive sensory acquisitions cannot be guaranteed due to a series of stability problems, thus, we adapted a different procedure. To achieve the same result, the robot creates the displacement needed for the parallax motion moving the head position but, maintaining a stationary position of the torso and shifting the considered angle from the azimuth to the altitude.

The initial robot position is marked as t_1 and the first angle between the robot head and the target position is evaluated. After the robot stands up, the new angular position t_2 is evaluated and the distance is computed. If a mathematical formulation is adopted, the following formula provides the distance calculation:

$$k = \frac{D}{E} = \frac{sin(\alpha)}{sin(\beta - \alpha)} \quad (1)$$

where α and β are the angles measured respectively at time t_1 and time t_2 as reported in Fig. 2. However, brains do not use this mathematical asset, so, being helped from the particular configuration of the Drosophila CX, a neural network has been designed in order to be able to estimate the distance between the robot and the target (i.e., D). It is important to notice that the distance will be evaluated using as measurement unit the distance between the two angular acquisitions (i.e., E). This method has a solid biological background in fly larva.

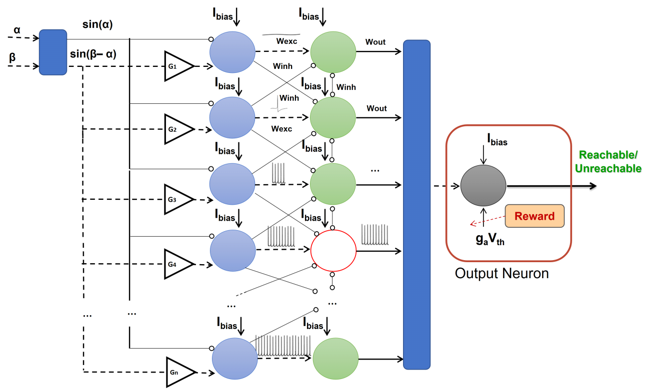


Fig. 3. Block schema of the control structure devoted to learn correlations between the robot body size and the objects in the environment. The robot generates parallax motion by moving the head from one position to another. A suitable visual system permits to detect objects of interest in order to identify the two different angles of view as result of robot head movement, as shown in Fig. 2. The model through a series of processing blocks, is able to estimate the distance from the object. Hence, an output neuron fires if the object is reachable (i.e., its estimated distance is below a threshold). A reward signal guides the learning mechanism according to the success or failure in taking an identified object.

A schema of the neural architecture is reported in Fig. 3. The network is composed of a series of layers developed using spiking neurons. At time t_1, afferent input from the visual layer (i.e., dotted arrow in Fig. 3) inhibits all the neurons of the first layer of the network with a synaptic gain equal to the sinusoidal function of its absolute angular position $\sin(\alpha)$, shown in the picture as a continuous line. This mechanism is not learned, it is considered pre-wired in the neural structures responsible for the topographical mapping of the visual input in the neural centres devoted to the processing.

At time t_2 the neuron in the visual network excites each neuron of the first layer with a current that is obtained multiplying the sin of the difference between the two perceived angular positions $\sin(\beta-\alpha)$ and an incremental gain G_i for $i = 1, ..N$, where N represents the number of neurons used in this layer (in these simulations we selected $N = 90$).

The array of gains was conveniently tuned to match with the second input: excitatory inputs should compensate the inhibitory ones to allow the neuron to fire. The value for these gains was chosen following a simple linear distribution that depends on the position of neuron i:

$$G_i = 0.05 * (i - 1) + 1 \quad (2)$$

This solution allows to discriminate distances from 1 to 5.45 times the distance E travelled between the two angular acquisitions (see Fig. 2), with a precision that linearly increases with the number of neurons.

The first and second layers of spiking neurons have been realized using the Izhichevich's class II neuron model [18], selected for the computational efficiency and dynamical properties. They are connected with direct excitatory synapses and lateral inhibitory ones. All assigned weights are set to $W_{exc} = 2$ and $W_{inh} = -1$, except for the last neuron in the second layer that receives from the $(N-1)$-*th* neuron of the first layer an inhibitory synaptic connection with a weight of $W_{inh} = -3$ to solve balancing problems on the boundary of the network. Adjacent neurons in the second layer are also connected using mutual inhibitory synapses with the same weighs.

This network configuration allows to obtain the following behaviour: depending on the inputs, the first layer of neurons will be divided into two groups in which the first one will not be able to spike due to the low excitatory input gain, whereas the last part of the array will be sufficiently excited and thus will fire. This behavioural activity will be transferred to the second layer where a *winner-takes-all* topology has been implemented: the neuron corresponding to the boundary of activity of the first layer will be the only one able to fire. Thereafter, this information is transferred to the final output neuron that will receive an input current proportional to the index of the firing neurons in the previous layer.

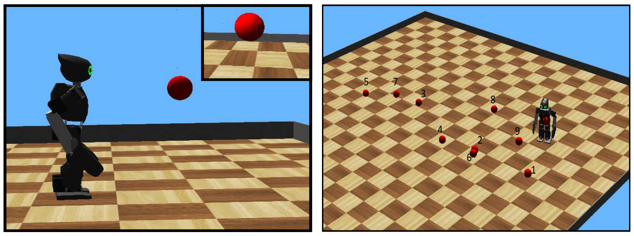
The last part of the network consists of a spiking neuron, whose spiking activity is correlated to the action to be performed by the robot. Namely, only if the output neuron fires, (meaning that the detected object is reachable) the robot starts a catching action. The output neuron is subject to a learning process based on threshold adaptation. When the system receives a reward signal, the threshold level is modified obtaining either a facilitation or a reduction of the spiking response for the output neuron. The threshold adaptation process can be implemented through a voltage-dependent current (I_A) formalized as an additional input to the decision neuron, expressed as I_A = -$g_A * V_{th}$. The system is initially facilitated in trying to perform the catching action by initializing the threshold to low numbers and choosing $g_A = 1$.

The final processing layer can be therefore considered as a gate to determine if an object is reachable or not depending on its estimated distance. The correct decision is therefore learned through an operant conditioning method.

Whenever the robot, guided by the output neuron firing, performs a successful action, this generates a reward used to modify the output neuron threshold depending on the matching between the reinforcement signal and the internal prediction. If the outcome for the selected object is reachable (unreachable) but the performed attempt is unsuccessful, the threshold is increased (decreased) to hyperpolarize (depolarize) the output neuron. If the prediction is correct, the threshold remains unchanged. The learning process is summarized in the following equations:

$$V_{th} = \begin{cases} V_{th} + \Delta V_{th} \text{for incorrect attempt} \\ V_{th} \ \text{for correct prediction} \\ V_{th} - \Delta V_{th} \text{for incorrect give up} \end{cases} \quad (3)$$

The application of Eq. 3 allows the adaptation of the output neuron firing as a function of the incorrect behaviours performed by the robot.



(a) (b)

Fig. 4. Simulation environment used to evaluate the body size learning on the Darwin-OP robot. (a) The robot can detect, using the embedded camera, the presence of red balls (i.e., targets) that randomly appear in the arena (b).

IV. Simulation Results

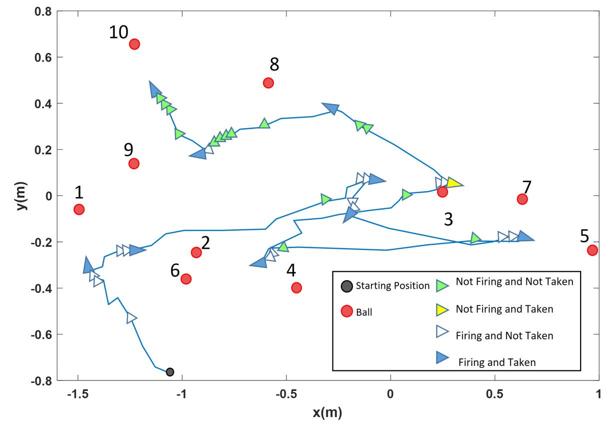
Darwin-OP which stands for Dynamic Anthropomorphic Robot with Intelligence-Open Platform is a miniature-humanoid robot platform with an advanced computational power. It is equipped with multiple sensors and actuated with smart servomotors. Darwin-OP has twenty degrees of freedom each one controlled by a Dynamixel MX-28T servo motor [19].

To perform the simulations in order to validate and evaluate the capabilities of the control architecture modified for the Darwin-OP robot, the Webots dynamic simulation environment was considered. Webots is a development environment used to model, program and simulate mobile robots [20].

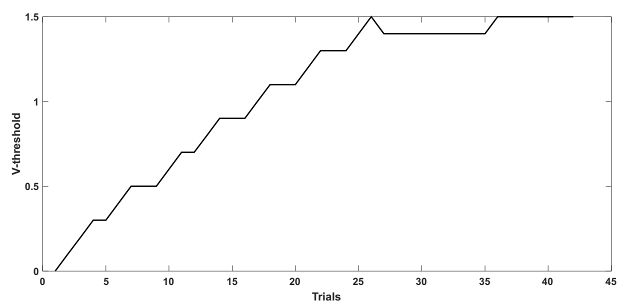
The adopted simulation set-up takes into account the Darwin-OP humanoid robot walking in an arena where targets of interest (i.e., red balls) can appear in different positions. The simulated robot, equipped with a camera, can detect the presence of a target and proceeds to a parallax estimation of the reachability for the selected object. Due to the orientation errors accumulated by the robot during walking, the parallax angles were acquired by changing the robot posture from up to down as discussed in Section III.

The simulation scenario is shown in Fig. 4 where the humanoid robot is walking, guided by the vision system, trying to catch the balls that randomly appear in the arena.

The spiking network processes the information acquired by the vision system and the activity of the output neuron controls the consequent robot behaviour. Thus, if the output neuron fires, the robot tries to catch the ball extending its arms, otherwise an approaching movement is performed and a new acquisition occurs. At the beginning of the learning phase, the output neuron threshold is equal to zero and the neuron is active independently from the parallax estimation. This situation is therefore translated in a continuous execution of the catching behaviour. The robot needs to try this action to evaluate the consequences and to tune its body model accordingly. The success of the procedure is detected analysing if the ball position changes after the catching action. The robot is initially equipped with a tool, an arm extension that changes its reachability space: the length of the tool is about $12cm$ obtaining a total arm extension that increases from $22cm$ to $34cm$. This tool will be taken into account by the learning process that will find the suitable threshold v_{th} associated to this robot configuration.



(a)



(b)

Fig. 5. (a) Trajectory followed by the robot during a learning phase. The robot approaches the visible target performing a series of catching actions whose outcome will affect the threshold update. When a ball is caught a new one randomly appears, the sequence of generated balls is labelled with the order of appearance. (b) The evolution of the output neuron threshold during the 40 trials executed while catching ten red balls. The threshold stabilizes around the value $V_{th} = 1.5$.

An example of the robot behaviour during a learning phase is reported in Fig. 5, when the arm extension is equipped.

To avoid useless trials, if the area of the bounding box around the red ball is below a certain size, the object is considered too far, and an approaching behaviour is implemented. The robot located on a starting position, looks for targets in the arena and detects the red ball labelled with the number 1. Due to the very small area of the detected ball, an approaching behaviour is performed and the robot moves five steps towards the target. After this action the robot evaluates the parallax motion and follows the output neuron commands that considers as reachable the ball. The catching action was not successful therefore a *firing and not taken* event occurs that updates the output neuron threshold with an increment to calibrate the body size model. After a series of unsuccessful actions, the robot finally succeed in catching the ball. The caught ball disappears and a new

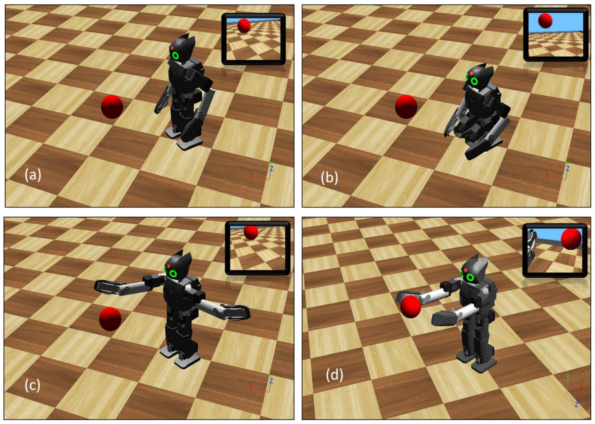
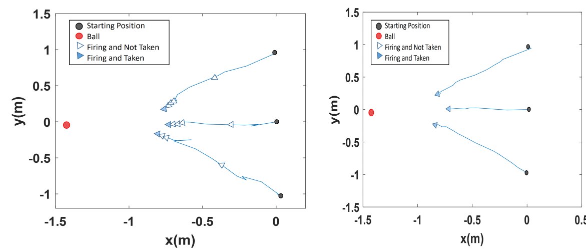


Fig. 6. Sequence of actions performed by the robot during a catching trial: (a) the ball is identified by the vision system and its elevation angle is acquired; (b) the robot modifies its posture (i. e. crouched down) to acquire the new elevation angle of the target ball; (c) the robot tries to catch the ball extending its arms; (d) the arm touching the ball moves the target position acquired from the camera and a rewarding signal is generated.



(a) (b)

Fig. 7. Behaviour of the robot trying to catch the same target from three different starting positions. Trajectories followed and behaviours generated during the learning phase (a) and when the threshold reached its steady state value (b).

one is refreshed in a random position. The robot, trial by trial, learns and after ten balls, for a total of 40 trials the robot succeeds in learning its body size and the threshold is stabilized around the value of 1.5 as reported in Fig. 5 (b).

Fig. 6 shows the sequences of actions performed when the robot successfully catches a ball: the α and β angles are acquired (Fig. 6 (a) and (b)), the robot performs a catching behaviour extending its arms toward the ball (Fig. 6 (c)), the camera detects a change in the ball position rewarding the robot for the successful action (Fig. 6 (d)).

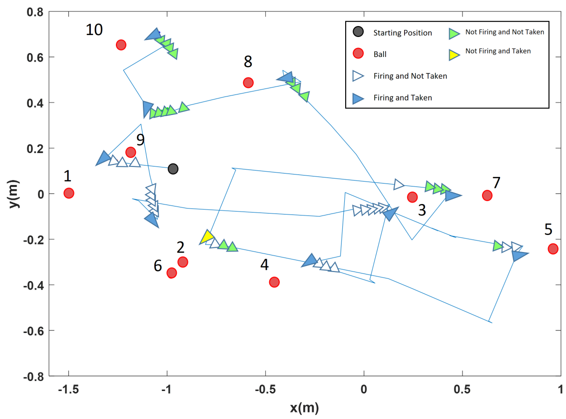
The behaviour performed by the simulated robot to catch the same target, starting from three different positions, is reported in Fig. 7. The *Not taken* event that guides the learning of the v_{th} (see Fig. 7 (a)) disappears when the output threshold reaches the value of $v_{th} = 1.5$ (Fig. 7 (b)).

To evaluate the learning capability of the developed architecture, the behaviour of the robot has been also evaluated by removing the tools from the arms, with a significant reduction of the reachability space. The learning process, performed using the same arena previously introduced, was accomplished using the same sequence of ball presentation used with the arm extension. The robot trajectory marked with the events related to the performed behaviours is shown in Fig. 8 (a). The time evolution of the output neuron threshold is reported in Fig. 8 (b), where a different steady state solution is obtained through the learning process, if compared with the results reported in Fig. 5 (b). The information about the body size of the robot in terms of reachability space of the arm is embedded within the output neuron threshold value that in presence of the arm extension converges towards $v_{th} = 1.5$ whereas without the tool it assesses around $v_{th} = 1.9$.

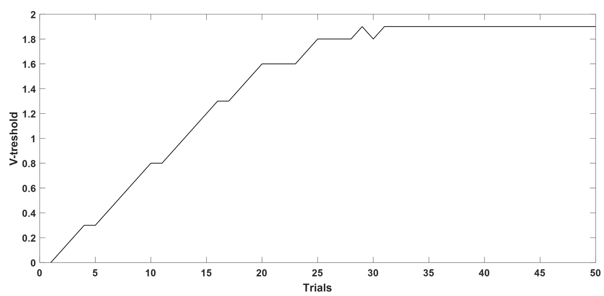
The developed controller was also evaluated in practice, using the Darwin-OP robot. A series of experiments were performed using the body size models tuned in the dynamical simulation environment and tested with the robot in an arena with a red ball used as target. The obtained results are summarized in Fig. 9 where the behaviours performed by the robot with and without the arm extension are shown for comparison. A visual filter able to detect red objects in the scene was used to identify the position of the target. Interestingly, the sequence of actions performed by the robot are equivalent to those obtained in simulation and the learned model in both cases is able to find the correct position from which the ball can be touched extending the arm.

V. Conclusion

This work focuses on the development of a suitable insect-inspired neural network for body size learning. For autonomous robotic systems that need to versatile adapt in complex scenarios, there is a growing interest for techniques allowing a robot to automatically learn its own body schema with only minimal human intervention. Using the experimentally available results, a neural model is here proposed based on the mechanisms used by flies to learn their body capabilities. Although some computationally-oriented hypotheses have been taken into account, they do not compromise the biological coherence of the model. The architecture has been applied both in simulation and in real with the Darwin-OP humanoid robot in a scenario where the robot should learn, through a reward based system, the reachable/unreachable space in the arena. An important point to be focused is that the network is application-independent, in the sense that body size learning is an intrinsic property of the living being and the network should be able to assume the correct behaviour no matter the robot on which it's implemented and the action to execute. The model of the neural network is a complete working prototype, whose functionalities have been tested in a simulated environment with good results, and preliminary experiments were developed on the real robot. A natural extension of the developed model will be the introduction of multiple output neurons that can be separately tuned in order to represent different aspects. For instance, it can be associated to either the reachability of an object within a given number of steps or the possibility to pass through a door depending on its own body occupancy.



(a)



(b)

Fig. 8. (a) Trajectory followed by the robot during a learning phase when the arm extension was removed. (b) The evolution of the output neuron threshold during the 50 trials executed while catching ten red balls. The threshold stabilizes around the value $v_{th} = 1.9$.

ACKNOWLEDGEMENT

This work was partially supported by the MIUR Project CLARA - Cloud platform for LAndslide Risk Assessment (grant SNC_00451). The authors acknowledge the support of Prof. Roland Strauss, Jogu-Mainz (De) for his helpful discussions and R. Crispino, F. Farinella and F. Leggio for their contribution in the initial simulation of the neural model.

REFERENCES

[1] C. Nabeshima, M. Lungarella, and Y. Kuniyoshi, "Timing-based model of body schema adaptation and its role in perception and tool use: A robot case study," in *4th IEEE International Conference on Development and Learning (ICDL-05)*, Osaka, Japan, 2005, pp. 7–12.

[2] J. Sturm, C. Plagemann, and W. Burgard, "Adaptive body scheme models for robust robotic manipulation," in *RSS - Robotics Science and Systems IV*, 2008.

[3] M. Hoffmann, H. Marques, A. Hernandez, H. Sumioka, M. Lungarella, and R. Pfeifer, "Body schema in robotics: a review," *Autonomous Mental Development*, vol. 2, no. 4, pp. 304–324, 2010.

[4] A. Bruckstein, R. J. Holt, I. Katsman, and E. Rivlin, "Head movements for depth perception: Praying mantis versus pigeon," *Autonomous Robots*, vol. 18, no. 1, pp. 21–42, Jan 2005.

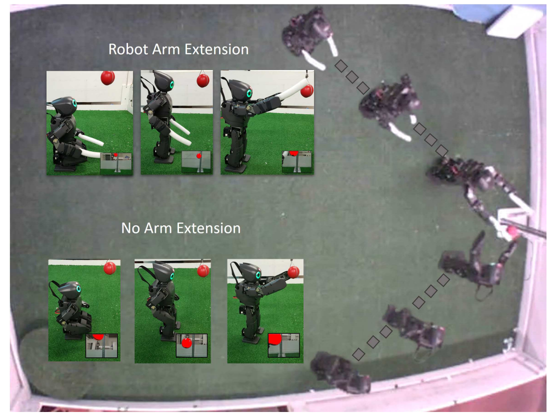


Fig. 9. Trajectories obtained using the Darwin-OP robot to test the body size model learned in simulation. The behaviour obtained (not) using the arm extension is shown in the (lower) upper part of the picture. A series of images shown the final behaviour of the robot that evaluates the parallax motion and catch the ball are reported together with the images acquired from the on-board camera.

[5] T. Krause and R. Strauss, "Mapping the individual body-size representation underlying climbing control in the brain of *Drosophila melanogaster*," in *Proc. 33rd Gottingen Meeting of the German Neuroscience Society*, 2011, pp. T25–11B.

[6] J. Nadler, D. Angelaki, and G. DeAngelis, "A neural representation of depth from motion parallax in macaque visual cortex," *Nature*, vol. 452(7187), pp. 642–645, 2008.

[7] W. P. Medendorp, D. B. Tweed, and J. D. Crawford, "Motion parallax is computed in the updating of human spatial memory," *Journal of Neuroscience*, vol. 23, no. 22, pp. 8135–8142, 2003.

[8] P. Arena, G. D. Mauro, T. Krause, L. Patané, and R. Strauss, "A spiking network for body size learning inspired by the fruit fly," in *Proceedings of International Joint Conference on Neural Networks*, Dallas, TX, 2013, pp. 1251–1257.

[9] A. Vitanza, L. Patané, and P. Arena, "Emergence of diversity in a group of identical bio-robots," *International Journal of Advanced Robotic Systems*, vol. 12, no. 10, p. 148, 2015.

[10] ——, "Spiking neural controllers in multi-agent competitive systems for adaptive targeted motor learning," *Journal of the Franklin Institute*, vol. 352, no. 8, pp. 3122 – 3143, 2015, special Issue on Advances in Nonlinear Dynamics and Control.

[11] B. Skinner, *The Behavior of Organisms:An Experimental Analysis*. New York: Appleton-Century-Crofts, 1938.

[12] S. J. Yi, B. T. Zhang, D. Hong, and D. D. Lee, "Online learning of a full body push recovery controller for omnidirectional walking," in *2011 11th IEEE-RAS International Conference on Humanoid Robots*, Oct 2011, pp. 1–6.

[13] L. Patané, R. Strauss, and P. Arena, *Nonlinear Circuits and Systems for Neuro-inspired Robot Control*. SpringerBriefs in Nonlinear Circuits, 2018.

[14] U. Homberg, *Arthropod brain, its evolution, development, structure and functions*. New York: Wiley, 1987, vol. 347–367, ch. Structure and functions of the central complex in insects.

[15] U. Hanesch, K. Fischbach, and M. Heisenberg, "Neuronal architecture of the central complex in drosophila melanogaster," *Cell Tissue Res.*, vol. 257, pp. 343–366, 1989.

[16] P. Arena and L. Patané, *Spatial Temporal Patterns for Action Oriented Perception in Roving Robots*. Springer, 2008.

[17] ——, *Spatial Temporal Patterns for Action-Oriented Perception in Roving Robots II: an insect brain computational model*. Springer, Series: Cognitive Systems Monographs, Vol. 21, 2014.

[18] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on Neural Networks*, vol. 14, pp. 1569–1572, 2003.

[19] I. Ha, Y. Tamura, and H. Asama, "Development of open platform humanoid robot darwin-op," *Advanced Robotics*, vol. 27, no. 3, pp. 223–232, 2013.

[20] "Webots robot simulation environment - web site." [Online]. Available: https://www.cyberbotics.com/