Online Reward-Based Training of Spiking Central Pattern Generator for Hexapod Locomotion

Ashwin Sanjay Lele, Yan Fang, Justin Ting, Arijit Raychowdhury
School of Electrical and Computer Engineering, Georgia Institute of Technology

2020 IFIP/IEEE 28th International Conference on Very Large Scale Integration (VLSI-SOC) | 978-1-7281-5409-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/VLSI-SOC46417.2020.9344100

***Abstract*—Online learning in legged robot under stringent performance and energy constraints thwarts the application of conventional reinforcement learning and optimization algorithms. The integration of complex sensors and data pre-processing required in using these algorithms makes this more challenging. Spiking neural networks allow local learning and low computing power opening new possibilities neuromorphic paradigm to such tasks. Central pattern generation based learning to walk in hexapod robots perfectly matches the temporal learning in SNNs allowing end-to-end learning. We propose a stochastic reinforcement-based algorithm allowing the hexapod to learn using the reward generated by the gyro sensors and camera-based visual inputs. The system is implemented on a Raspberry pi to demonstrate convergence to bio-observed gait patterns.**

I. Introduction

Central pattern generators are neural circuits in the brain generating temporally correlated spiking patterns actuating rhythmic muscle movements of limbs. Their activity is modulated by the spiking generated by various sensory inputs like the vestibular system in cockroaches. The CPG is also triggered by the visual inputs required in approaching and tracking prey. This biological closed-loop spiking system presents an ideal inspiration for a low-power autonomous robot. [1]

Spiking neural networks promise high energy efficiency and decentralized processing coupled with the ability to accommodate RL based learning perfectly suited for online reward-based tasks. Electronic implementations of spiking-CPG (SCPG) typically use a fully connected spiking neural network model shown in Fig. 1(a). Linear equation solving based reverse engineering approach has been demonstrated for hexapod CPGs capable of generating multiple gaits [1]. Another offline training using an evolutionary algorithm has been demonstrated in [2]. However, none of these approaches uses the autonomous learning capability of SNNs to learn to produce walking gaits without any prior knowledge in online reward-based end-to-end learning.

In this work, we demonstrate a stochastic reward-based weight update algorithm allowing a hexapod robot to learn to walk without any prior knowledge using the online rewards generated by the gyro-sensor and visual input. The rewards alter the weights exploring the correct pattern of leg movement Fig. 1(e). The gait converges to bio-observed tripod gait in most cases with some occurrences of sub-optimal slower gaits still enabling the forward motion. To the best of our knowledge, this is the first autonomous gait-learning demonstration using SNNs published in [3].

[1] 978-1-7281-5409-1/20/31.002020*IEEE*

II. Methodology: Algorithm and Hardware

The SCPG network consists of six fully connected neurons, firing of a neuron causes movement of the corresponding leg. The neurons follow leaky-integrate and fire (LIF) behaviour. Two neurons namely, input neuron (N_{in}) and gyro-driven neuron (N_{gyro}) trigger the SCPG. The ideal gait pattern requires alternate neurons fire in one step causing tripod gait Fig. 1(e).

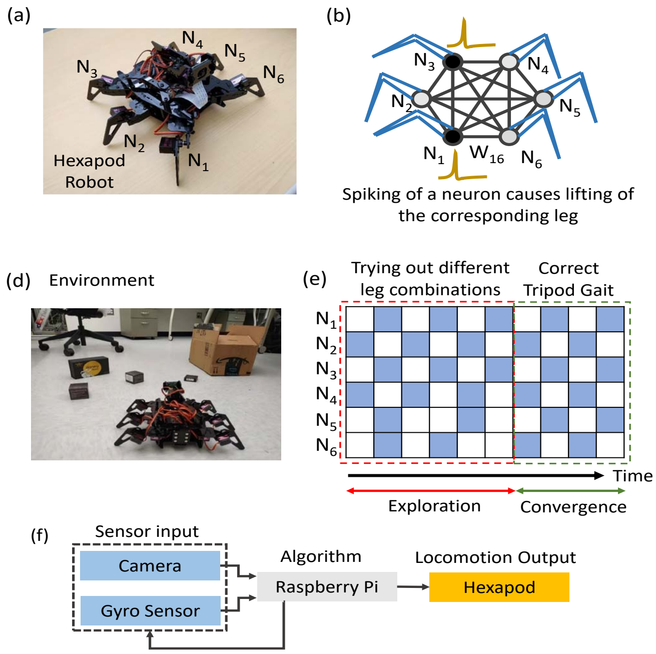

Fig. 1. (a,b) Hexapod robot with labeled neurons for mimicking the CPG. Spiking of a neuron causes motion of the corresponding leg (c) Office environment for demonstration (d) SCPG spiking as the algorithm progresses. The random spiking in the exploration phase gets latched to the correct tripod gait showing convergence (f) Hardware implementation of the system

The full system consists of an SCPG, gyro-sensor and optical camera. In every exploration step, the system is initialized by taking the gyroscope reading and capturing the image. Input neuron triggers a combination of SCPG neurons and corresponding legs to move. During the step, gyro-sensor evaluates the balance and camera captures an image after the step. Higher stability generates higher gyro-reward. The vision-based reward is calculated by applying the light-weight odometry method described in [4] to the initial and final images to estimate the magnitude and direction of the movement. The total reward is modulated by a random number and the weights are stochastically updated.

Gy-521 MPU-6050 gyro sensor and piCamera provide sensory inputs to the Raspberry pi 3 Model B+ processing unit. Adeept RaspClaws Hexapod Spider Robot is used as the locomotion platform.

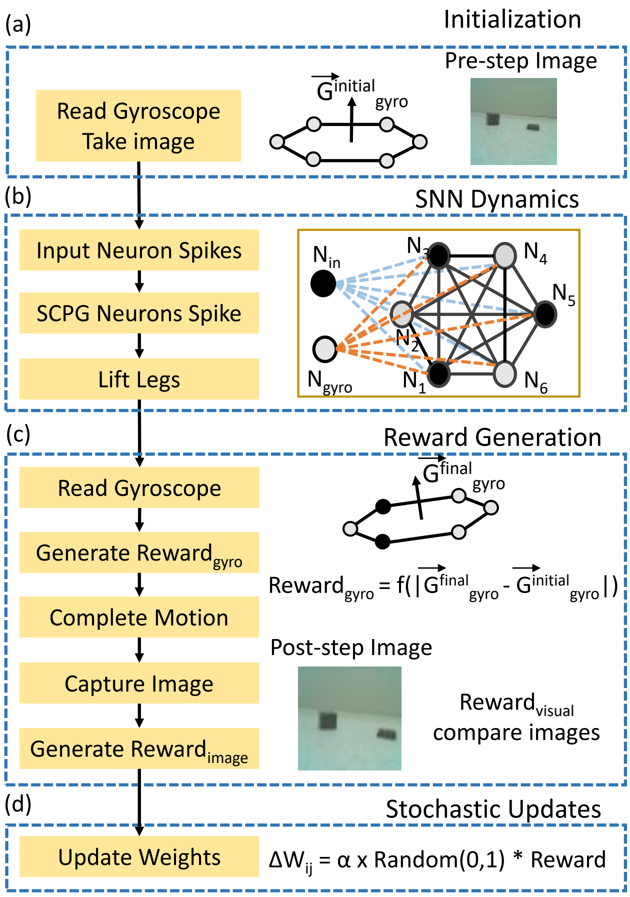

Fig. 2. Block diagram of the algorithm (a) Every step is initialized with a gyroscope reading and an image of the surroundings (b) LIF neuronal dynamics compute the spiking of the neurons (c) Sensory inputs are used to calculate the awards (e) Weight update using calculated rewards

III. Results

Fig. 3 shows the time evolution of the simulation. The exploration starts and oscillates between different combinations of two and four leg motions. The corresponding cumulative reward is shown in the adjacent figure. After the network latches on to the correct tripod gait, the balance and forward motion are maintained simultaneously generating high reward. Demo-1, shows the convergence to tripod gait in the 66^{th} cycle in hardware. However, the weight updates being stochastic, in some cases, intermediate non-bio-observed gaits are also seen to result in a forward motion with balance preservation. These sub-optimal gaits corresponding to weight parameters getting stuck into local minima are seen in (demo-2). The average number of time steps required for convergence are calculated by careful tuning of rewards and learning rate. Energy cost on Intel's Loihi [5] is estimated to be $\approx 855.1nJ$ using this. After the system has learned the correct gait, the energy consumed in every step is $\approx 9.1nJ$.

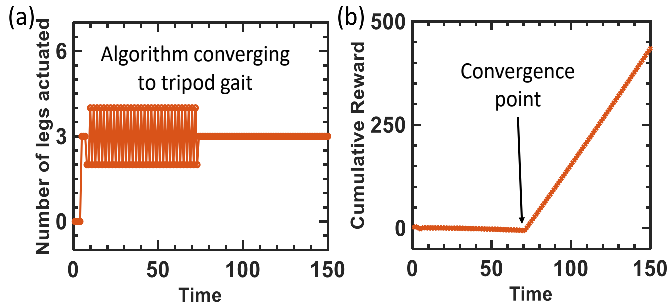

Fig. 3. The number of legs moved at a time instance oscillates around 3 and converges to 3 in a tripod gait resulting in accumulation of high constant reward (b) High positive reward is accummulated with latching to correct gait

TABLE I
Comparison With Prior Work

Reference	Training Approach	Sensory Feedback	Online / Offline
[1]	Linear Programming	None	Offline
[2]	Grammar Evolution	None	Offline
[6]	Reward STDP	Olfactory + Visual	Offline
This Work	Stochastic Reward	balance + visual	Online

IV. Conclusion

We propose a closed-loop learning system in spiking neural network-based CPG with online reward generation to train a hexapod to learn to walk. The low energy consumption validates its potential in edge-robotics. The gait converges to biological tripod gait in most cases while converging to non-bio-observed sub-optimal gaits in some cases.

V. Acknowledgement

This work was supported by CBRIC, a center in JUMP, a Semiconductor Research Corporation program by DARPA.

References

[1] H. Rostro-Gonzalez, P. A. Cerna-Garcia, G. Trejo-Caballero, C. H. Garcia-Capulin, M. A. Ibarra-Manzano, J. G. Avina-Cervantes, and C. Torres-Huitzil, "A cpg system based on spiking neurons for hexapod robot locomotion," *Neurocomputing*, vol. 170, pp. 47–54, 2015.

[2] A. Espinal, H. Rostro-Gonzalez, M. Carpio, E. I. Guerra-Hernandez, M. Ornelas-Rodriguez, and M. Sotelo-Figueroa, "Design of spiking central pattern generators for multiple locomotion gaits in hexapod robots by christiansen grammar evolution," *Frontiers in neurorobotics*, vol. 10, p. 6, 2016.

[3] A. S. Lele, Y. Fang, J. Ting, and A. Raychowdhury, "Learning to walk: Spike based reinforcement learning for hexapod robot central pattern generation," in *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, pp. 208–212, IEEE, 2020.

[4] M. J. Milford and G. F. Wyeth, "Mapping a suburb with a single camera using a biologically inspired slam system," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1038–1053, 2008.

[5] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018.

[6] E. Arena, P. Arena, R. Strauss, and L. Patané, "Motor-skill learning in an insect inspired neuro-computational control system," *Frontiers in Neurorobotics*, vol. 11, p. 12, 2017.