7.4 A 55nm Time-Domain Mixed-Signal Neuromorphic Accelerator with Stochastic Synapses and Embedded Reinforcement Learning for Autonomous Micro-Robots

Anvesha Amravati, Saad Bin Nasir, Sivaram Thangadurai, Insik Yoon, Arijit Raychowdhury

Georgia Institute of Technology, Atlanta, GA

Even as rapid advances are being made in the areas of deep neural networks (DNNs) and convolutional neural networks (CNNs) with most hardware demonstrations geared towards inference in vision-based platforms [1-5], we recognize that true autonomy in intelligent agents will only emerge when such bio-mimetic systems can perform continuous learning through interactions with the environment. Reinforcement learning (RL) presents one such computational paradigm inspired by behaviorist psychology, where autonomous agents take actions in an environment to maximize a notion of cumulative reward. This concept is deeply rooted in the human brain where dopamine mediated neurotransmitters (in the cortex, striatum and thalamus of the brain) have been shown to encourage reward-motivated behavior in all our social interactions (Fig. 7.4.1). In this paper, we present a 690μW (V_{CC}=1.2V) neuromorphic accelerator fabricated in 55nm CMOS, which: (1) inherits unique properties of stochastic neural networks, (2) leverages recent advances in Q-learning as an implementation of RL, and (3) demonstrates energy-efficient time-domain mixed-signal (TD-MS) circuit architectures, to provide autonomy to a mobile, self-driving micro-robot at the edge of the cloud, with possible applications in disaster relief, reconnaissance and personal robotics.

The feed-forward path is implemented in a three-layered neural network (input, hidden and output) and the network sizes and bitwidths are optimized for minimum power for the target application. Fig. 7.4.1 illustrates the system diagram where three ultra-sonic (US) sensors feed pulses (depth information) directly to a layer of 84 TD neurons through an array of stochastic synapses, thus avoiding time-to-digital conversion at the sensor interface. These hidden layer neurons perform a weighted sum of the inputs and using an activation function (rectified linear unit, ReLU) each neuron produces pulses that are retransmitted via stochastic synapses to the output layer of neurons. The output layer, after a winner take all (WTA) comparator produces an action that leads the robot to move straight, left or right (action). Each action is associated with changing sensor data, which is re-evaluated for continuous RL. Using backpropagation and gradient descent, the feedback circuit predicts the reward (i.e., avoid obstacles and cover maximum distance for area mapping), computes the loss function and updates the model (synaptic weights) for further exploration. The test-chip enables full-scan, embedded timing and memory controllers, debug features, direct interfaces to US sensors and can interface with a microcontroller board for motor and sensor control.

Motivated by the fact that edge devices need to operate at ultra-low power, and the observation that such systems require low effective-number of bits (4-8b) in the feed-forward data-paths, we employ analog computing blocks. However, voltage-domain analog circuits require high V_{CC} to accommodate the dynamic range and data-conversion is typically expensive. Hence, we explore TD–MS circuits, where operands are represented by pulse-widths, thereby enabling large dynamic ranges even close to V_T. As a trade-off, the architectures are slower, which is perfectly acceptable for the applications in hand. Fig. 7.4.2 captures the components and simulations for the TD-MS neuron and the stochastic synapse. A TD-MS multiply-and-accumulate (MAC) is implemented by a 21b counter (Fig. 7.4.2) which multiplies the 6b input (x) from a pre-synaptic neuron to the 6b weight (w) of the synapse. The input is a pulse of width $T = x \bullet T_0$ (generated by a digital to pulse converter, DPC) where T_0 is the unit delay of the DPC. A local DCO with embedded memory (stores the model weights for the fanin synapses, w) generates a frequency $F = w \bullet F_0$ and clocks an up/down counter. The DCO is kept ON for the period, T, thus enabling in-situ computation of $x \bullet w$ and accumulation in the counter. The up/down counter easily enables negative values of w (downcount) in the signed magnitude system. The ReLU activation function of the neuron is implemented by a DPC (Fig. 7.4.2) and feeds into the following stage. Such a TD-MS MAC shows unique properties: (1) the energy to perform a MAC is proportional to the magnitude of the operands and hence the importance of the computation in the neural network – a feature inherent in the brain, but missing in digital logic (shown in the color-map), (2) 45% lower system area (largely contributed by lower routing overhead) than a digital implementation, (3) 47% lower interconnect power, since each synaptic connection is one buffer-chain and goes through one 0→1 and one 1→0 transition irrespective of the operand value and (4) 16% lower leakage power. Stochastic synapses with drop-connect prevents data overfitting and are implemented here with a buffer-chain whose delay comprises of (1) a scan-programmable fixed part, and (2) a stochastic part where the 0→1 and 1→0 transition delays are randomly altered (or the pulses dropped randomly for drop-connect) by a local high-speed LFSR. At every cycle of a reference clock, one neuron in a layer produces pulses (Fig. 7.4.2) that are simultaneously captured by all the neurons in the next layer. The controller then activates the next neuron of the layer and so on. In each cycle, one neuron of every layer fires in a pipelined fashion, increasing system throughput. Fig. 7.4.3 illustrates the data-flow for RL, key circuit components, including reuse of compute blocks with the feedforward path. After an action, A_t, Q-learning provides reward prediction for the loss function evaluator. We implement a support vector machine (SVM) hinge loss function in 2 cycles. 84 parallel units compute the hidden layer gradient using TD macros and reusing the DCOs and DPCs in the feed-forward path. Next, the gradients of the hidden layer (Δ) are calculated. Hidden layer outputs are multiplied by the predicted reward to produce weight updates (w^{ho}). In the input layer, the gradients Δ pass through a ReLU gradient estimator and multiply with the outputs of the input neurons to update the weights of the input layer (w^{ih}). In the final cycle, all the updated weights are written back into the local memory of the neurons and the system is readied for the next feedforward exploration. Unique characteristics of the system architecture include (1) memory-in-logic to reduce data movement, and (2) TD analog for computation and communication seamlessly interfacing with digital (counters and memory) for storage.

The F_0-T_0 design-space is shown in Fig. 7.4.4 and measured results illustrate that the system remains bounded from 0.4-1.0V avoiding both counter overflow and loss of resolution in the neuron's MAC. The measured F_{DCO} (of the hidden layer) shows peak performance of 780MHz (at 1V), INL of 1.2LSB (1.6LSB) at 1V (0.4V) and DNL of 1.4LSB (1.5LSB) at 1.0V (0.4V). Measured T_{DPC} also shows good linearity with INL of 1.3LSB (1.6LSB) at 1V (0.4V) and DNL of 1.5LSB (at 1V) and 1.8LSB (at 0.4V). We note that: (1) the DCO and the DPC are both composed of programmable buffer chains and their delays track each other across V_{CC}, and (2) the programmability of the DPC and the DCO depends on the number of stages of buffer delay, which results in high linearity. Measured stochasticity of test synapses demonstrate a uniform distribution with variation of ±40%. The benefit of stochasticity is seen in the system emulation (Fig. 7.4.5) where a stochastic network reduces the loss function by more than 30% for 10^4 training samples. Fig. 7.4.5 illustrates measured F_{MAX} and power demonstrating operation down to 0.4V and peak power of only 690μW (at 1.2V). The peak throughput shows a wide dynamic range catering to a variety of RL tasks. Peak energy efficiency is obtained at 0.8V (V_T~0.5V), where we note 690pJ/inference and 1.5nJ/training, demonstrating a 1.25pJ/MAC (worst-case). Comparison with the existing literature demonstrates ultra-low power (690μW at peak performance), an average of 3.12TOPS/W and enables unique neuromorphic functionality. The testchip is mounted on a mobile nano-robot for autonomous exploration and learning and the overall distance moved by the robot as a function of the reference clock is shown in Fig. 7.4.6. The die-shot and chip micrograph are shown in Fig. 7.4.7.

Acknowledgements:
This work was sponsored by Qualcomm Inc. and National Science Foundation under grant 1640081, and the Nanoelectronics Research Corporation (NERC), a wholly-owned subsidiary of the Semiconductor Research Corporation (SRC), through Extremely Energy Efficient Collective Electronics (EXCEL), an SRC-NRI Nanoelectronics Research Initiative under Research Task ID 2698.002.

References:
[1] J. Park, et al., "Online RL NoC for Portable HD Object Recognition Processor," *CICC,* 2012.
[2] D. Shin, et al., "DNPU: An 8.1TOPS/W Reconfigurable CNN-RNN Processor for General-Purpose Deep Neural Networks," *ISSCC*, pp. 240-241, 2017.
[3] B. Moons. et al., "ENVISION: A 0.26-to-10TOPS/W Subword-Parallel Dynamic-Voltage-Accuracy-Frequency-Scalable Convolutional Neural Network Processor in 28nm FDSOI," *ISSCC*, pp. 246-247, 2017.
[4] J. Sim, et al., "A 1.42TOPS/W Deep Convolutional Neural Network Recognition Processor for Intelligent IoT Systems," *ISSCC*, pp. 264-265, 2016.
[5] Y-H. Chen, et al., "Eyeriss: An Energy-Efficient Reconfigurable Accelerator for Deep Convolutional Neural Networks," *ISSCC*, pp. 262-263, 2016.

 978-1-5090-4940-0/18/$31.00 ©2018 IEEE

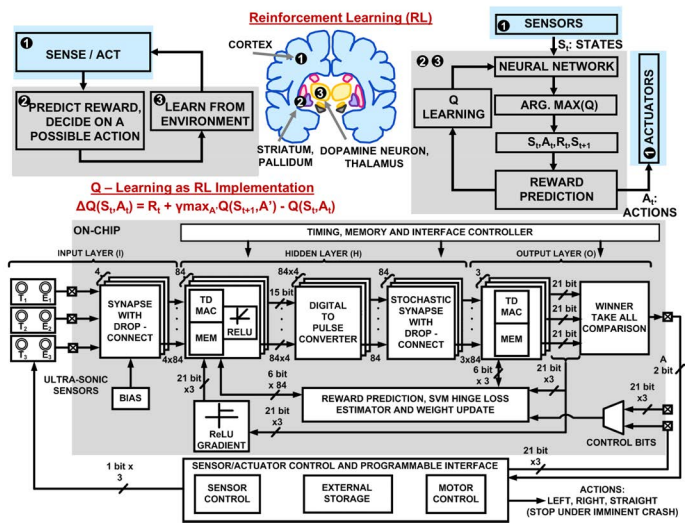


Figure 7.4.1: Motivation and system diagram for reinforcement learning as a neuromorphic computational model in autonomous mobile micro-robots.

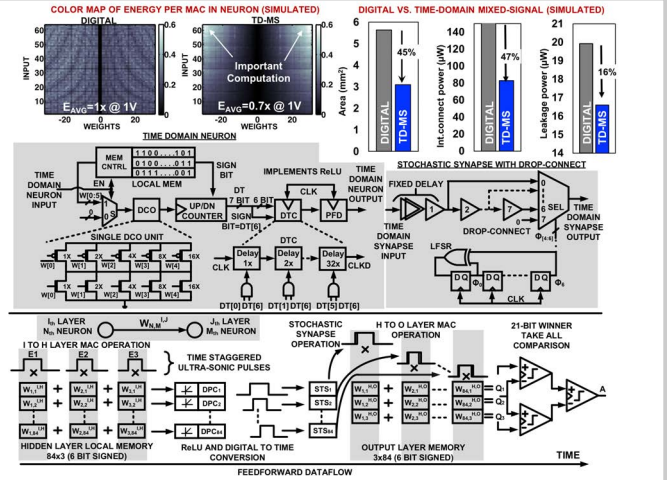


Figure 7.4.2: Comparison of time-domain mixed-signal (TD-MS) design vs. digital implementation, design of TD-MS neuron with ReLU, stochastic synapses and feedforward timing diagram illustrating pipelined data-propagation.

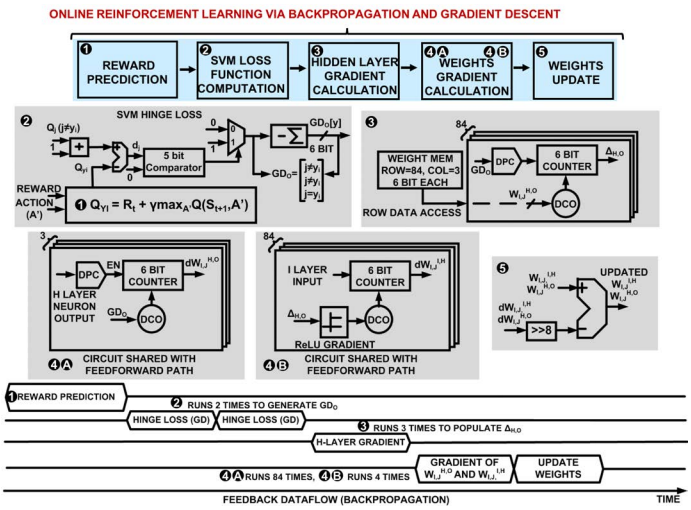


Figure 7.4.3: Flow-chart and corresponding circuit diagrams for online RL via backpropagation and gradient descent.

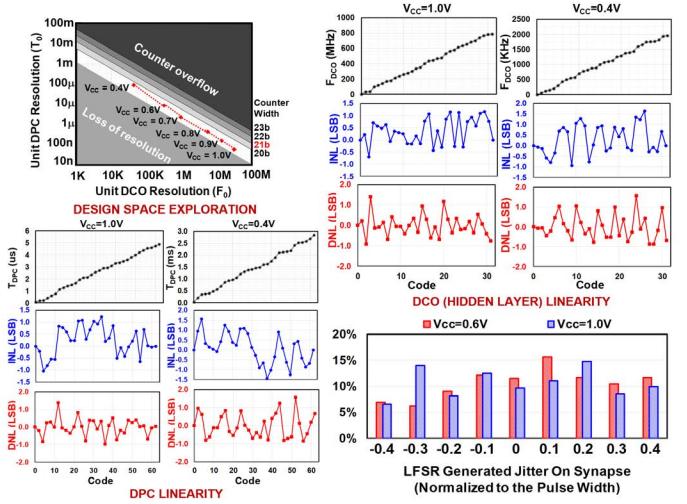


Figure 7.4.4: Design space illustrating measured operating points, measured DCO characteristics, measured DPC characteristics and measured stochasticity generated by the synapses.

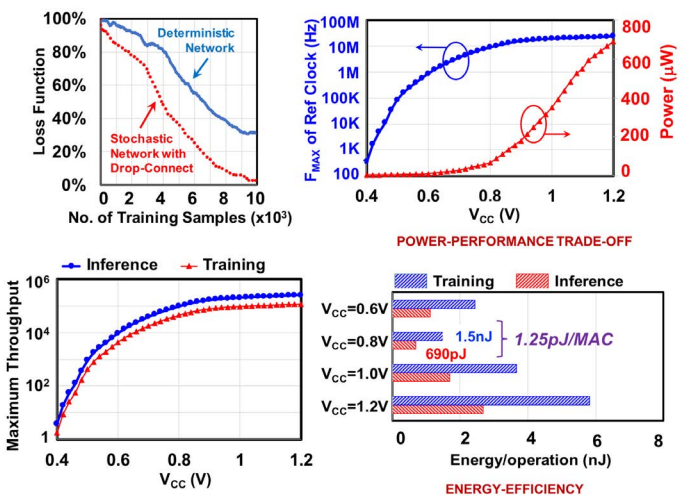


Figure 7.4.5: Role of stochasticity in loss minimization and measured system performance-power trade-offs, illustrating 690μW at peak performance and a worst-case of 1.25pJ/MAC.

	This work	[1]	[2]	[3]	[4]	[5]
ML System	Reinforcement Learning	Object Recognition	CNN-RNN	CNN	DNN	DNN
Technology	55nm	65nm	65nm	180nm	65nm	65nm
Circuit style	Time domain mixed-signal	Digital	Digital	Digital	Digital	Digital
Area	3.4mm^2	4mm^2	16mm2	3.3mm^2	16mm^2	16mm^2
Learning/Training	Online in real time	Offline	Offline	Offline	Offline	Offline
Stochasticity	Present	Absent	Absent	Absent	Absent	Absent
Resolution	6bit MAC/ 21b Counter	16b	16b	4b-16b	16b	16b
Power	690 uW at peak performance	121mW	63mW	7.5-300mW	45mW	278mW
Supply voltage	0.4-1V	1.2V	0.77-1.2V	Unavilable	1.2V	0.82-1.17V
No. inferences/sec	254,000	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported
No. of training/sec	118,000	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported
Performance/Watt	3.12TOPS/W	1.24TOPS/W	2.1TOPS/W	0.26-10TOPS/W	1.42TOPS/W	0.21TOPS/W
Application	Autonomous micro-robotics	Object Recognition	General purpose DNN	Visual recognition	CNN processor	Vision
Min. Energy/inference	690pJ	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported
Min. energy/training	1.5nJ	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported

COMPARISON TABLE

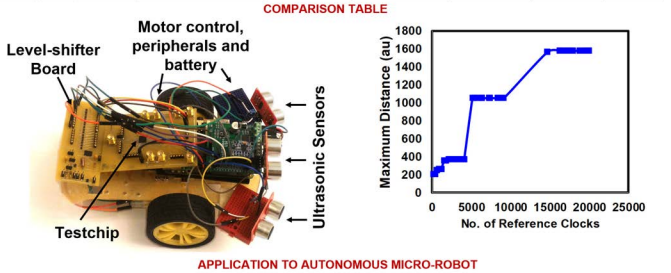
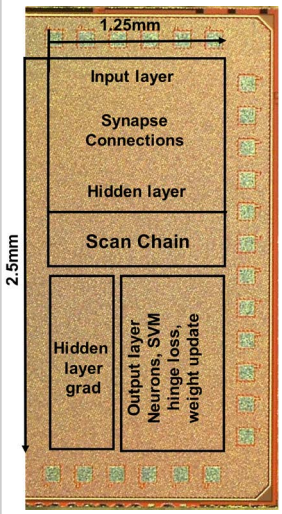


Figure 7.4.6: Comparison table and application to mobile micro-robotics illustrating the use of RL in exploration.



Chip Characteristics	
Technology	55nm 1P8M CMOS
Die area	1.25mm *2.5mm
Testing Interface	QFN package
Pin Count	24

Figure 7.4.7: Die-photo and chip characteristics.

978-1-5090-4940-0/18/$31.00 ©2018 IEEE