An FPGA-based cortical and thalamic silicon neuronal network

Takuya Nanami
School of Engineering, The University of Tokyo
Tokyo, Japan

Takashi Kohno
Institute of Industrial Science, The University of Tokyo
Tokyo, Japan

Abstract

A DSSN model is a neuron model which is designed to be implemented efficiently by digital arithmetic circuit. In our previous study, we expanded this model to support the neuronal activities of several cortical and thalamic neurons; Regular spiking, fast spiking, intrinsically bursting and low-threshold spike. In this paper, we report our implementation of this expanded DSSN model and a kinetic-model-based silicon synapse on an FPGA device. Here, synaptic efficacy was stored in block RAMs, and external connection was realized based on a bus that conform to the address event representation. We simulated our circuit by the Xilinx Vivado design suit.

Keywords: silicon neuronal network, neuron model, FPGA, cortex, thalamus.

1. Introduction

Silicon neuronal networks can simulate neuronal activity with low power consumption and in high speed. They are thought to be a promising way to realize an extremely large-scale network comparable to the human brain in future. The Field-Programmable Gate Array (FPGA) devices are commonly used for silicon neuronal networks because they can implement user-designed circuits with low cost. Many silicon neuronal networks on an FPGA can run at a higher speed than the biological real-time [1][2]. For example, a fully-connected 1024-neuron network that is 100 times the real-time speed has been implemented on a single Virtex-5 FPGA [3]. Thomas et al. adopted the Izhikevich (IZH) model that can simulate various types of neuronal activities with their corresponding parameter sets. Li et al. [4] constructed an auto-associative memory with 256 fully connected digital spiking silicon neuron (DSSN) models on an FPGA. The DSSN model [5] is an qualitative neuronal model designed for efficient implementation by efficiently in a digital arithmetic circuit. This model is a non-I&F-based model and can realize several neuronal activities including Class I and II in the Hodgkin's classification [6]. This model can reproduce the gradient response in Class II neurons [7], because it does not abbreviate the calculation of the spiking process. Due to a trade-off between the reproducibility of neuronal activity and computational efficiency, diversified neuronal models are used in silicon neuronal networks. For example, the ionic-conductance models can reproduce a neuronal activity accurately but consumes excessive computational resources in large-scale implementations. On the other hand, neuronal models that approximate a spiking process by resetting of the state variables (integrate-and-fire (I&F) -based models), such as the LIF, exponential I&F model[], adaptive exponential I&F model[8] and IZH models, can be implemented at low computational cost. However, it has reduced reproducibility of complex neuronal activities. For instance, these models assume fixed maximum membrane potentials during the spike process, whereas these potentials are non-uniform in the nervous system.

In our previous research [9], we expanded the DSSN model to cover cortical and thalamic neuron classes, including regular spiking (RS), fast spiking (FS), low-threshold spike (LTS) and intrinsically bursting (IB). We confirmed that the model behaves in the same way as the ionic-conductance model [10] in response to various magnitude of step input in each neuron classes.

In this paper, we designed digital arithmetic circuit for the DSSN model on an FPGA. We simulated our circuit on Xilinx Vivado design suit. The remainder of this paper is organized as follows. Section 2 introduces our neuron model and details of its implementation. The simulation

result is shown in Section 3. Section 4 summarizes the work and suggests ideas for future.

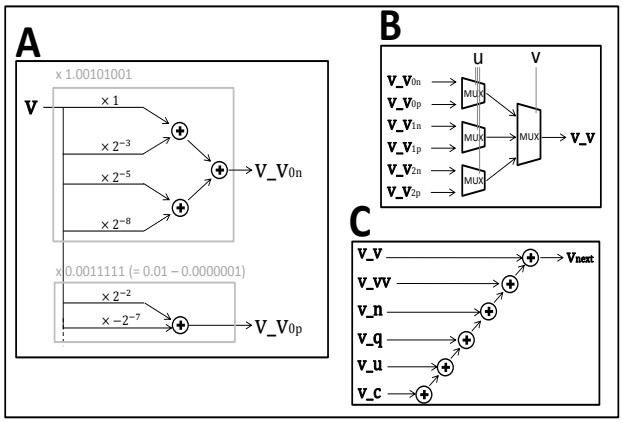

Fig.1 Block diagram of the v-circuit. Symbols ×, + and *MUX* mean a multiplier, adder, multiplexer, respectively. (A) Multiplication realized by shift operation and adder. (B) Add operation of the differential equation of *v*. (C) Correct v_{ij} are chosen by multiplexer depending on *v* and *u*.

2. Methods

2. 1 DSSN model

The DSSN model is a qualitative neuronal model. It is designed so that diverse neuronal activities with fixed point Euler method. Given appropriate parameter sets, this model can reproduce various neuron classes including regular spiking (RS), fast spiking (FS), intrinsically busting (IB), and low-threshold spike (LTS). The model is given by

$$\frac{dv}{dt} = \frac{\emptyset(u)}{\tau}(f(v) - n - q + I_0 + I_{stim}),$$

$$\frac{dn}{dt} = \frac{1}{\tau}(g(v) - n),$$

$$\frac{dq}{dt} = \frac{\varepsilon}{\tau}(h(v) - q),$$

$$\frac{du}{dt} = \frac{\varepsilon_1}{\tau}(v - v_0 - \alpha u),$$

$$\emptyset(\mathrm{u}) = \begin{cases} \emptyset_0 & (u < r_{u0}) \\ \emptyset_1 & (r_{u0} < u < r_{u1}) \\ \emptyset_2 & (r_{u1} < u), \end{cases}$$

$$f(v) \equiv \begin{cases} a_{fn}(v - b_{fn})^2 + c_{fn} & (v < 0) \\ a_{fp}(v - b_{fp})^2 + c_{fp} & (v \geq 0), \end{cases}$$

$$g(v) \equiv \begin{cases} a_{gn}(v - b_{gn})^2 + \mathrm{c_{gn}} & (v < r_g) \\ a_{gp}(v - b_{gp})^2 + c_{gp} & (v \geq r_g), \end{cases}$$

$$h(v) \equiv \begin{cases} a_{hn}(v - b_{hn})^2 + \mathrm{c_{hn}} & (v < r_h) \\ a_{hp}(v - b_{hp})^2 + c_{hp} & (v \geq r_h), \end{cases}$$

where *v* denotes the membrane potential, and *n* and *q* are the fast and slow variables, respectively, that abstractly

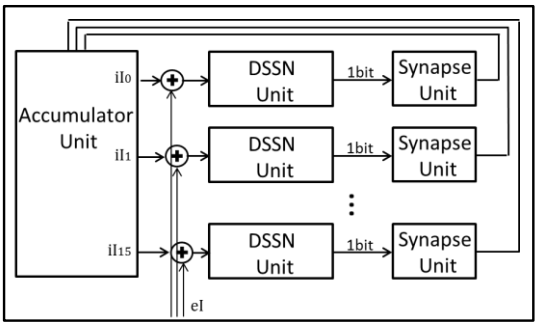

Fig. 2 The architecture of 16 fully-connected network. The network is composed of 16 DSSN, 16 silicon synapse, and an accumlator unit

describe the activity of the ion channels. The slow variable q controls the slow dynamics of the neuronal activity and has a key role to realize spike-frequency adaptation and burst firing. Variable u is the slowest that modifies the structure of the fast subsystem comprising *v* and *n*. If sustained stimulus to the fast subsystem is taken as a bifurcation parameter, a saddle-node bifurcation is observed in the RS, FS, and LTS modes, and a homoclinic - loop bifurcation in the IB mode. The parameter I_0 is a bias constant and I_{stim} represents the input stimulus. The other parameters determine the dynamical properties of the model. All of the variables and constants in this qualitative model are purely abstracted and have no dimension.

The model is solved by the Euler's method ($\Delta t = 0.0001$). We developed differential equations for efficient implementation. The equations of the circuit for calculation of *v* are

$$v_{\mathrm{next}} = v\_vv + v\_v + v\_n + v\_q + v\_I + v\_c,$$

$$v\_vv_{ij} = v^2 \cdot (\Delta t \cdot a_{fj} \cdot \emptyset_i/\tau),$$

$$v\_v_{ij} = v \cdot (\Delta t \cdot b_{fj} \cdot a_{fj} \cdot \emptyset_i/\tau + 1),$$

$$v\_n_{ij} = n \cdot (-\Delta t \cdot \emptyset_i/\tau),$$

$$v\_q_{ij} = q \cdot (-\Delta t \cdot \emptyset_i/\tau),$$

$$v\_I_{ij} = I_{\mathrm{stim}} \cdot (\Delta t \cdot \emptyset_i/\tau),$$

$$v\_c_{ij} = \Delta t \cdot (b_{gj} \cdot b_{gj} \cdot a_{gj} + c_{gj})/\tau,$$

$$v\_\mathrm{x} = \begin{cases} v\_x_{0n} & \text{when} \quad u < r_{u0} \text{ and } v > 0, \\ v\_x_{0p} & \text{when} \quad r_{u0} < u < r_{u1} \text{ and } v > 0, \\ v\_x_{1n} & \text{when} \quad r_{u1} < u \text{ and } v > 0, \\ v\_x_{1p} & \text{when} \quad u < r_{u0} \text{ and } v > 1, \\ v\_x_{2n} & \text{when} \quad r_{u0} < u < r_{u1} \text{ and } v > 1, \\ v\_x_{2p} & \text{when} \quad r_{u1} < u \text{ and } v > 1, \end{cases}$$

$$\text{for} \quad x = vv, v, n, q, I, \text{and } c.$$

where, *i* denotes *0*, *1*, or *2* and *j* denotes *n* or *p*. In these,

equations, we can calculate a product of parameters in advance and store it as a constant value. The

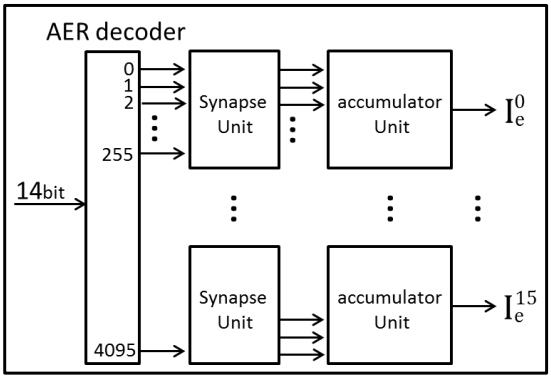

Fig. 3 The architecture of external connection. This module has 16 silicon synapse units and 16 accumlator units and an AER decoder.

multiplication of a variable and a parameter is realized by shifters and adders (Fig.1(A)). To reduce the number of adders, coefficient was approximated by the sum of 2^{-x}; x denotes arbitrary integer number. Multiplier circuit is used once in an Euler's method's step for calculation of v^2. v_{ij} are screened by the multiplexer depending on the value of *v* and *u* (Fig. 1(C)). We constructed *n*-circuit and *q*-circuit in the same way. The DSSN unit consumes 1 multiplier and 4 multiplexers. The numbers of adders are 35, 33, 60, and 88 for RS, FS, LTS, and IB classes.

We adopted the silicon synapse circuit in [4] based on the kinetic synapse model [11].The equations are given by

$$\frac{dI_s}{dt} = \begin{cases} \alpha(1 - I_s) & ([T] = 1) \\ -\beta I_s & ([T] = 0) \end{cases}$$

where, I_s represent the post-synaptic current and the value of [T] is 1 when the membrane potential of the presynaptic neuron is over the threshold (0). This silicon syanpse unit costs 1 multiplexer and 2 adders, and it requires 2 clocks by a step.

2.2 Circuit architecute

Our circuit is composed of 16 fully-connected neurons and external input from an address event representation (AER)–based stimulus receiver module. The fully-connected neuronal network consumes 16 DSSN and 16 silicon synapse units and an accumlator unit comprising 1 multiplier and 1 adder and an 18kb block random access memory that stores synaptic weights (Fig. 2). The synapse units calculate postsynaptic input depending on the 1-bit signal that denotes whether the membrane potential *v* of the pre-synaptic DSSN is over the threshold or not. The accumlator unit calculates weighted sum of the post-synaptic input. The I_{stim} is calculated as follows.

$$I_{\text{stim}}^i = C_0 \sum_{j=1}^{16} w_{ij} I_s^j + I_e^i$$

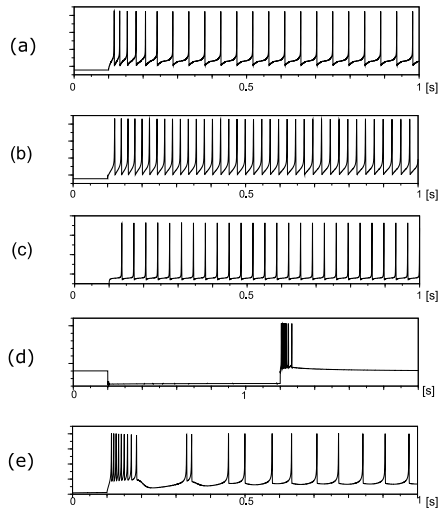

Fig .4 Waveforms generated by the DSSN unit of RS (a), FS (b), LTS (c, d), IB (e). Excitable step inputs rise at t = 0.1 (a-c,e). inhibitory step inputs are provided from t = 0.2 to t = 1.0 (d).

where i and j are the indices of the post- and pre-synaptic neurons respectively, and I_e^i denotes an external input to the *i* th neuron, respectively. The parameter c_0 is a coefficient to scale I_{stim} into an appropriate range.

The units for the external connection is composed of 16 silicon synapse units and 16 accumlator units and an AER decoder (Fig.3). AER decoder receives 14-bit input signal by a clock. First 12-bit of input signal specifies an address of 4096 synapses by AER decoder. Next 1-bit represents rising or trailing of the pulse input, and the value of the bit is stored in the register corresponding to the specified address. Therfore, we can assigne input pulse having an arbitrary length. Last 1-bit is an enable signal.

Synaptic weights are loaded from the 16 18kb block random acces memory. A synapse unit updates 256 post-synaptic current depending on the decoded input signal at each Δt step, and the accumulator unit calculates an weighted sum of the post-synaptic inputs.

$$I_e^i = c_1 \sum_{k=1}^{256} w_{ik} I_s^k$$

where i and k are the indices of the neurons and synapses, respectively. The parameter c_1 is a coefficient to scale I_e^i into an appropriate range.

In our system, a Δt step corresponds to 1024 clocks, and 1024 input signals are acceptable by 1 step. The DSSN unit require 10000 steps to simulate a second, and it runs 10 times faster than real-time under the assumptuion that a clock correspond to 10 ns. We represent all variables and synaptic weights using 18-bit signed fixed point with 14-bit fractions.

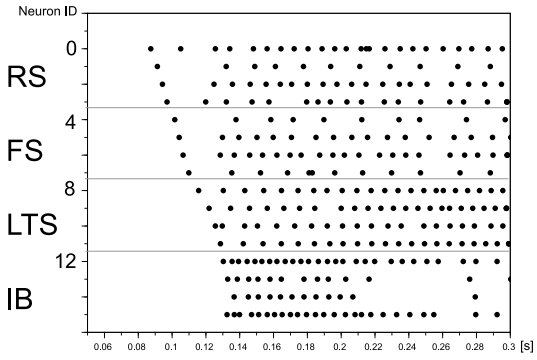

Fig.5 Raster plots of the spikes in the network. Neurons are RS (Neuron ID = 0-3), FS (4-7), LTS (8-11), IB (12-15).

3. Result

Here, we show the waveforms generated by the DSSN unit in response to external stimulus inputs. Figure 4(a) shows the waveform of the RS class that is a most typical neuron class in the cortex. A characteristic behavior of this class is spike-frequency adaptation; that is, the spike frequency decreases over time in response to a constant stimulation input. On the other hand, FS neurons fire with almost constant frequency (Fig.4(b)). LTS neurons exhibit periodic firing in reaction to excitable input stimulus (Fig.4(c)) and generate a burst firing just after the termination of a sufficient hyperpolarizing stimulus (Fig.4(d)). IB neurons generate the bursting at the onset of a stimulus, then continue to spiking (Fig. 4(e)).

Figure 5 shows raster plots of the spikes in the network. We applied 12-bit signal correspond to each neuron by rotation and evoked spike, and additional spikes were evoked by the inner connection. Synaptic weights were assigned randomly.

4. Conclusion

In this paper, we constructed digital circuit of the DSSN model that supports four cortical and thalamic neuron classes. The implemented circuit could generate activities qualitatively comparable to the ionic-conductance model for each neuron class. It consumed only 1 multiplier for calculation of an Eular's Method's step, which is an expensive module in an FPGA. We confirmed the behavior of a network of 16 fully-connected DSSNs that can receive stimuli via 4096 synapses. Each synapse is controlled by 12-bit input signal through an AER decoder.

Our previous study implemented the DSSN model that supports the Class I and Class II in the Hodgkin's classification. We expect the circuit constructed in this work will be a basis for digital silicon neuronal networks that can support a wide variety of neuronal activities more elaborately than the I&F-based circuits.

In future works, we will implement the DSSN model that supports the square-wave bursting and elliptic bursting which were already realized by software simulation. Larger-scale network will also be pursued.

5. Acknowledgement

This work was supported by JSPS Grant-in-Aid for scientific Exploratory Research Number 25240045.

6. Reference

1. R. K. Weinstein and R. H. Lee. Architectures for high-performance fpga implementations of neural models. J. Neural Eng, 3(1):21–34, 2006.
2. M. J. Pearson, A. G. Pipe, B. Mitchinson, K. Gurney, C. Melhuish, I. Gilhespy, and M. Nibouche. Implementing spiking neural networks for real-time signal-processing and control applications: A model-validated fpga approach. IEEE Transactions on Neural Networks, 18(5):1472–1487, 2007.
3. D. B. Thomas and W. Luk. Fpga accelerated simulation of biologically plausible spiking neural networks. In K. L. Pocek and D. A. Buell, editors, FCCM, pages 45–52. IEEE Computer Society, 2009
4. J. Li, Y. Katori, and T. Kohno. An fpga-based silicon neuronal network with selectable excitability silicon neurons. Frontiers in neuroscience, 6(183), 2012.
5. T. Kohno and K. Aihara. Digital spiking silicon neuron: Concept and behaviors in gj-coupled network. Proceedings of International Symposium on Artificial Life and Robotics 2007, 2007.
6. A. L. Hodgkin. The local electric changes associated with repetitive action in a non-medullated axon. The Journal of physiology, 107(2):165–181, Mar. 1948
7. X.-J. Wang and J. Rinzel. Oscillatory and bursting properties of neurons. In M. A. Arbib, editor, The Handbook of Brain Theory and Neural Networks, pages 835–840. MIT Press, MA, 2nd edition, 2003.
8. R. Brette and W. Gerstner. Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity. J. Neurophysiol., 94:3637–3642, 2005.
9. T. Nanami and T. Kohno. Simple cortical and thalamic neuron models for digital arithmetic circuit implementation. Inpreparation.
10. M. Pospischil, M. Toledo-Rodriguez, C. Monier, Z. Piwkowska, T. Bal, Y. Fregnac, H. Markram, and A. Destexhe. Minimal hodgkin-huxley type models for different classes of cortical and thalamic neurons. Biological Cybernetics, 99(4-5):427–441, 2008.
11. Alain Destexhe, Zachary F. Mainen, and Terrence J. Sejnowski. Kinetic Models of Synaptic Transmission. in: Methods in Neuronal Modeling, Edited by Koch, C. and Segev, I. (2nd Edition) MIT Press, Cambridge, 1998