PAPER Special Section on Novel Device Architectures and System Integration Technologies

A VLSI Spiking Feedback Neural Network with Negative Thresholding and Its Application to Associative Memory

Kan'ya SASAKI^{†a)}, Student Member, Takashi MORIE^{††}, Member, and Atsushi IWATA[†], Fellow

SUMMARY An integrate-and-fire-type spiking feedback network is discussed in this paper. In our spiking neuron model, analog information expressing processing results is given by the relative relation of spike firing. Therefore, for spiking feedback networks, all neurons should fire (pseudo-)periodically. However, an integrate-and-fire-type neuron generates no spike unless its internal potential exceeds the threshold. To solve this problem, we propose negative thresholding operation. In this paper, this operation is achieved by a global excitatory unit. This unit operates immediately after receiving the first spike input. We have designed a CMOS spiking feedback network VLSI circuit with the global excitatory unit for Hopfield-type associative memory. The circuit simulation results show that the network achieves correct association operation.

key words: spiking neuron model, feedback network, global excitatory unit, negative thresholding, associative memory

1. Introduction

In the conventional neural network models and their hardware implementation, the internal potential and the output of a neuron are often represented by analog values based on neuronal firing rate cording or firing population cording [1], [2]. These approaches are used for making VLSI implementation easy rather than for emulating real neuronal operations [3], [4].

However, recently, the spiking neuron models attract a lot of attention with expectation of their higher information processing ability [5], [6]. They represent analog information by the timing of a neuronal spike firing. Additionally, they operate asynchronously, which means that they have no global control signal. Therefore, these models could operate faster than synchronous neural network models. So far, the spiking neuron models have mainly been applied to feedforward networks; for example, a spiking feedforward network for image data processing was reported [7]. However, there have been few reports about spiking feedback networks. A spiking feedback network based on the Hodgkin-Huxley neuron model was proposed [8]. Although this neuron model can exhibit various nonlinear dynamics, it is difficult to realize it by CMOS VLSI circuits because of the complexity of the model. A spiking feedback net-

Manuscript received April 10, 2006.

a) E-mail: kanya@dsl.hiroshima-u.ac.jp

DOI: 10.1093/ietele/e89-c.11.1637

work based on the integrate-and-fire neuron model was also proposed [9]. This simplified neuron model can be easily implemented in CMOS circuits. However, the neuron model proposed in [9] only represents binary information by spikes.

The spiking neuron model treated in this paper expresses analog information by the relative timing of asynchronous spike firing events. Therefore, in this model, neurons cannot express their information unless they fire (pseudo-)periodically. However, a neuron generates no spike unless its internal potential exceeds the threshold. Therefore, naive spiking neuron models cannot be applied to feedback networks.

In this paper, we propose *negative thresholding* operation to overcome this difficulty. In order to achieving this thresholding operation, we introduce a global excitatory unit [10]. We design a spiking feedback network LSI circuit with a global excitatory unit and demonstrate the validity of our model by LSI circuit simulation of associative memory using the network.

This paper is organized as follows. In Sect. 2, a wellknown integrate-and-fire-type spiking neuron model is described. Section 3 proposes our spiking neural network model with a global excitatory unit for feedback networks. In Sect. 4, we propose a CMOS spiking neural network circuit. In Sect. 5, the validity of our model is demonstrated by circuit simulation results of associative memory using our spiking network. Finally, Sect. 6 presents our conclusions.

2. Integrate-and-Fire-Type Spiking Neuron Model

The conventional integrate-and-fire-type (IF) neuron model is shown in Fig. 1. When a neuron receives a spike pulse via a synapse, a post-synaptic potential (PSP) is generated. The internal potential of a neuron is determined by the spatiotemporal summation of PSPs generated by the input spikes. There are two types of PSPs: excitatory and inhibitory.

In the simple IF model, the time courses of PSPs are the same, which we call here a unit PSP, P(t), as a convolutional kernel, and a PSP from neuron *i* to neuron *n*, $PSP_{ni}(t)$, is given by the temporal summation of unit PSPs multiplied by the corresponding synaptic weight w_{ni} :

$$PSP_{ni}(t) = \sum_{t_i^{(f)} \in \mathcal{F}_i} w_{ni} P(t - t_i^{(f)}), \qquad (1)$$

where $\mathcal{F}_i = \{t_i^{(1)}, \dots, t_i^{(n)}\}$ is the set of firing times of neuron

Manuscript revised June 19, 2006.

[†]The authors are with the Graduate School of Advanced Sciences of Matter, Hiroshima University, Higashihiroshima-shi, 739–8530 Japan.

^{††}The author is with the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Kitakyushu-shi, 808–0196 Japan.



Fig. 1 Integrate-and-fire-type neuron model.

i. The type and amplitude of a PSP are determined by the sign (positive or negative) and the absolute value of synaptic weight w_{ni} , respectively. The internal potential of neuron *n*, $I_n(t)$, is given by the spatial summation of *PSP*_{ni}(*t*):

$$I_n(t) = \sum_{i \in \Gamma_n} PS P_{ni}(t),$$
⁽²⁾

where Γ_n is the set of inputs to neuron *n*. The effect of a PSP is temporary, and $I_n(t)$ returns to the resting potential level after the PSP ceases. When $I_n(t)$ exceeds the threshold *th*, neuron *n* fires and generates a spike. After firing, $I_n(t)$ is reset, and a refractory period T_r follows. Spikes generated at neuron *n* are transmitted to other neurons or the neuron itself via output line $i_n(t)$ with transmission delay time T_d .

3. Global Excitatory Unit for Spiking Feedback Networks

In the IF neuron model, if $I_n(t)$ does not exceed the threshold as shown in Fig. 2(a), the neuron obviously generates no spikes. In order to apply the spiking neuron model to feedback networks with continuous states, spikes expressing the zero value (standard timing) should be generated.

To achieve this, we propose *negative thresholding* operation that starts immediately after the first spike input. Here, we introduce a global excitatory unit *GEU* for this thresholding operation. Operation of the spiking neuron model with *GEU* is shown in Fig. 2(b). In this model, the potential generated by *GEU* at neuron n, $G_n(t)$, is given by

$$G_n(t) = w_{nG} \{ 1 - e^{-(t - t_i^{(1)})/\tau_G} \} \mathcal{H}(t - t_i^{(1)})$$
(3)

for all *i*, where w_{nG} is a synaptic weight from *GEU* to neuron n, τ_G is the time constant for rise time, $\mathcal{H}(\cdot)$ is the Heaviside step function. Here, for simplicity, w_{nG} is set at a constant positive value w_G . *GEU* receives spikes from all neurons via excitatory synapses, and it is activated by the earliest input spike, which we call a *GEU-trigger spike* and is generated at $t_i^{(1)}$. The activated *GEU* gives a continuous-level stimulus to all neurons. Therefore, *GEU* gives the same effect as a decrease in the threshold levels of all neurons, which means



Fig.2 Examples of input/output spike trains and neuronal internal potentials without (a) and with (b) *GEU*.



Fig. 3 Spiking feedback network model with GEU.

negative thresholding. If once *GEU* is activated, even a neuron at the resting state can fire, and a neuron with stronger inhibition generates a spike with later timing.

Figure 3 shows a spiking feedback network model with *GEU*, where all neurons are connected each other. If synaptic weights are symmetric $(w_{ij} = w_{ji})$, this network realizes Hopfield-type associative memory. The input key pattern is given from *in_n* as initial input spikes. In the associative memory operation, *GEU* must start to operate after all neurons receive initial input spikes. If not, all neurons fire independently of initial inputs. This situation is same as that the threshold levels for firing is set lower than the resting potential of neurons at the beginning of the network operation. In our network, neurons output spikes after all neurons receive initial input spikes. Therefore, we use the earliest output spike from neurons as the *GEU-trigger spike*.

On the other hand, *GEU* must give a continuous-level stimulus until the network recalls a stored pattern. After recalling a stored pattern, if the network starts another recalling operation by receiving another input pattern, *GEU* must be reset before the operation.

In our spiking feedback network model, analog information expressing processing results is given by the relative timing of neuronal spike firing. Because there is no global clock signal, it is required that some neurons fire synchronously and the timing of synchronous firing is used as the relative standard timing for information representation by spike timing.

In the associative memory using neural networks, the network must converge to a binary (0/1) stored pattern, which means all neurons separate into two groups and these two groups each generate synchronous spikes at different times corresponding to logical 0 and 1. The neurons belong to each of the two groups are often connected each other with excitatory synapses. Therefore, in order to improve the convergence property of our model, it should be satisfied that the condition that neurons connected with excitatory synapses generate synchronous spikes.

For this purpose, we make the transmission delay time T_d equal to the refractory period T_r . If both are different, the following problems may occur. When $T_r > T_d$, even if some neurons generate spikes synchronously at a certain time, the neurons cannot generate synchronous firing at next timing, because these neurons are in a refractory period when the neurons receive spikes. Therefore, stable synchronous firing cannot occurs. In contrast, if $T_r < T_d$, before receiving spikes from neurons that fired synchronously, neurons may fire by spikes from neurons that are firing asynchronously. Therefore, also in this case, stable synchronous firing cannot occurs.

Our *GEU* is purely proposed as a part of the artificial neural network model. However, functions controlling the whole activity of a neuron group are found in various places in the real brain. For example, acetylcholine increases the activity level of neurons in the learning process, and some biogenic amines are used to modulate the synaptic activity simultaneously in widespread areas of the brain [11]. *GEU* may be considered as a model for realizing such functions.

4. Spiking Neural Network Circuit

We have designed a spiking feedback network LSI circuit with *GEU* using a $0.35 \,\mu$ m CMOS technology. The synapse and neuron circuits are shown in Fig. 4. Figure 5 shows the timing diagrams of both circuits.

Figure 6 shows *delay* & *inv* 1/2 circuits and these timing diagrams. When *delay* & *inv* 1 receives an input pulse, V_a decreases immediately and increases slowly by PMOS transistor P_1 . As a result, this circuit outputs an inverted pulse whose fall timing coincides with the rise timing of the input pulse and whose pulse width is determined by V_{bias1} . In contrast, when *delay* & *inv* 2 receives an input pulse, V_c decreases slowly by NMOS transistor N_1 and increases immediately. As a result, this circuit outputs an inverted pulse whose rise timing coincides with the fall timing of the input pulse and whose pulse width is determined by V_{bias2} .

As shown in Fig. 4(a), the synapse circuit operate as an excitatory or inhibitory synapse according to the binary *sign* signal. The *sign* signal determines the sign of the synaptic weight. The synaptic weight is determined by the current value of the current source, which is made of a MOS FET. Bias voltage V_w , which is fed into the gate terminal of the MOS FET, determines this current value. The synapse circuit receives a spike input i_n generated by the corresponding



Fig. 4 Spiking neuron circuits: (a) synapse and (b) neuron.



Fig. 5 Timing diagrams of spiking neuron circuits: (a) synapse and (b) neuron.



Fig.6 Delay and inversion circuits and their timing diagrams: (a),(b) *delay* & *inv* 1, (c),(d) *delay* & *inv* 2.

neuron. In order to generate switching signal *psp* by the spike input, we use a *delay* & *inv* 1 and *NOR* circuits. Bias signal V_{bias1} of *delay* & *inv* 1 determines the temporal width

of the *psp* signal. The *psp* signal turns on a current switch, and the current source charges or discharges capacitor C_n in the neuron circuit, whose terminal voltage represents the neuronal internal potential. The spatiotemporal summation of PSPs by input spikes is performed by this capacitor.

In the neuron circuit, the neuronal internal potential I_n is compared with the threshold voltage th by a differential pair. The internal potential returns to the resting potential by leak resister R_n connected to the capacitor. In order to realize the transmission delay, spike pulses are delayed by using delay & inv 1/2 circuits. When the internal potential exceeds the threshold voltage, a spike pulse is generated. After neuronal spike firing, a refractory period is generated by increasing the threshold voltage. A spike pulse to other neurons is generated after the refractory period. The detail of the operation of the neuron circuit is as follows. When the terminal voltage of the capacitor I_n exceeds the threshold voltage th, the signal V_1 rise to 'High.' The signal V_2 is generated by V_1 with *delay* & *inv* 2. The signal V_3 is generated by logical AND operation between V_1 and V_2 . The signal V_4 is generated by V_3 . When V_4 is 'High,' th increases. The temporal width of V_1 represents the absolute refractory period, and the temporal width of V_4 represents the relative refractory period. Therefore, the temporal width between the rising edge of V_1 and the falling edge of V_4 represents the refractory period of the neuron. After the refractory period, a spike pulse to other neurons i_n is generated. Therefore, the transmission delay time T_d equals to the refractory period. The transmission delay time is represented by the summation of the pulse widths of signals V_3 and V_4 . These pulse widths are determined by the bias voltages V_{bias1} and V_{bias2} in delay & inv 1/2.

The neuron and synapse circuits shown in Fig. 4 are also used as the GEU circuit and the synapse circuits between GEU and neuron, respectively. In order to achieve a constant output from GEU, we use a *psp* with a longer pulse width than the whole operation time in the synapse circuits from GEU to neurons.

5. Simulation of Associative Memory Using Spiking Neural Network

5.1 Simulation Condition

Figure 7 shows a block diagram of the spiking feedback neural network circuit, which consists of 36 neurons with symmetric connections and a *GEU*. The synaptic weights w_{ij} , which are expressed by the sum of autocorrelation matrixes of the stored pattern vectors, are given by the following equation:

$$w_{ij} = \sum_{k=1}^{N} (2I_i^k - 1)(2I_j^k - 1), \text{ for } i \neq j,$$
(4)

where $w_{ii} = 0$, and I_i^k is the *i*-th element of the *k*-th stored pattern vector. In the simulation, the number of stored patterns *N* was five, and the stored patterns are shown in Fig. 8.



Fig.7 Block diagram of a spiking feedback neural network circuit with *GEU*.



The elements I_i^k were randomly chosen under the condition that the numbers of '0' (black) and '1' (white) are equal. As a result, $w_{ij} \in \{\pm 5, \pm 3, \pm 1\}$. We used gray-level (5-levels) input patterns that converge to stored pattern #1 shown in Fig. 8.

We evaluated the network performance by circuit simulation of the designed CMOS circuit. The simulations were performed by a high-speed circuit simulator, HSIM. The reason why we used circuit simulation instead of usual software simulation is that it directly leads to the VLSI implementation of our model.

In the simulation, the spike width was set at 20 ns. The time step corresponding to one level in gray-level patterns was set at 25 ns. Therefore, the firing timing of input spikes was limited in {0, 25, 50, 75, 100} ns, and the time span for receiving input spikes was 100 ns. The transmission delay time T_d was set at 200 ns. Input spikes generated at 0 ns and 100 ns express '1' (white) and '0' (black) pixels, respectively.



Fig. 9 Typical time course of PSP for input i(t).

Since a PSP is generated by an RC circuit, the unit PSP is given by

$$P(s) = P_0\{(1 - e^{-s/\tau})\mathcal{H}(s)\mathcal{H}(t_p - s) + (1 - e^{-t_p/\tau})e^{-(s-t_p)/\tau}\mathcal{H}(s - t_p)\},$$
(5)

where P_0 is a constant, τ is the time constant for decay, t_p is the time span between the time when the unit PSP starts to increase and the time when it has the peak value. A typical time course of P(t) is shown in Fig. 9. In our simulation, $P_0 = 0.05$, $t_p = 100$ ns, and $\tau = 50$ ns.

5.2 Simulation Results and Discussion

A simulation result about the network without GEU is shown in Fig. 10, which shows the outputs of twelve neurons corresponding to pixels 1 to 12 in the first and second rows of the pattern shown in Fig. 8. In the case without GEU, most neurons generate approximately periodical spikes because of feedback with transmission delay. However, some neurons generate no spikes at certain intervals because their internal potentials do not exceed the threshold. Thus, the networks without GEU cannot express analog information by spike timing properly, and therefore cannot operate as an associative memory.

In contrast, in a simulation result of the network with *GEU* shown in Fig. 11, all neurons generate approximately periodical spikes. At the beginning of the simulation, all neurons generate spikes whose timing corresponds to the gray-level input values. However, after several spike generations, the neurons that express *black* and *white* pixels (hereafter we call them *black*-pixel neurons and *white*-pixel neurons, respectively) in stored pattern #1 form themselves in two groups each of which generates synchronous spikes with different timing. Thus, by using *GEU*, all neurons can generate spikes with the timing according to their analog internal potentials, and the network can operate as an associative memory. In this result, the time difference between spikes belonging to the respective two groups is 100 ns, which is equal to $T_d/2$.

Figure 12 shows the relationship between the synaptic weights from *GEU* to all neurons, w_G , and the normalized calculation steps required for recalling, N_r . Here, N_r is defined as recalling time divided by the inter-spike interval after convergence; i.e., 200 ns in this simulation. From Fig. 12, N_r has the minimum value when $w_G = 2.0$.



Fig. 10 Simulation result without GEU.



Fig. 11 Simulation result with GEU.

For $w_G < 0.5$ and $w_G > 3.0$, the network could not converge to any stored patterns. When w_G is too small $(w_G < 0.5)$, the network cannot express analog information by spike timing as in the case without *GEU* shown in Fig. 10. On the other hand, when w_G is too large $(w_G > 3.0)$, the effect of *GEU* is so strong that the network cannot attain correct recalling. Therefore, it is necessary to set w_G at an appropriate value.

Figure 13 shows the convergence characteristic of spiking networks by using Manhattan distance d_M from the in-



Fig. 12 Relationship between the synaptic weight from *GEU* and the normalized calculation steps required for recalling.



Fig. 13 Recalling process: changes in Manhattan distance d_M from the input pattern to stored pattern #1 as a function of the normalized calculation steps.





Fig. 15 Simulation result with GEU to stored pattern #6.



put pattern to stored pattern #1, where $w_G = 2.0$. We used two different input patterns for evaluating each Manhattan distance. In this figure, $d_M = 0$ means stored pattern #1, $d_M = 36$ means the inverse pattern of #1, and $d_M = 18$ means the other stored patterns. It is found from the simulation results that these networks can recall the correct pattern #1 when $d_M < 10$.

We evaluated the spiking network operation about other five stored patterns shown in Fig. 14. Figure 15 shows a simulation result of the successful recall of stored pattern #6.

We also carried out simulations when the input spike-

Fig. 16 Simulation result when the input spike-timing range is smaller than half of the transmission delay time.

timing range T_{in} was changed. Figure 16 shows a simulation result for $T_{in} = 67$ ns, which is shorter than $T_d/2$ (= 100 ns). On the other hand, Fig. 17 shows a simulation result for $T_{in} = 150$ ns, which is longer than $T_d/2$. For both results shown in Figs. 16 and 17, the output spike-timing difference between the two groups after convergence is the same as $T_d/2$. Therefore, in our associative memory using a spiking feedback network, the output spike-timing difference is



Fig. 17 Simulation result when the input spike-timing range is larger than half of the transmission delay time.

independent of the input spike-timing range, and the output spike-timing difference is equal $T_d/2$.

The reason why our network has such characteristics is as follows. In our network operating as an associative memory, the *black*-pixel neurons and the *white*-pixel ones inhibit each other. Therefore, the *black*-pixel neuron group and the *white*-pixel neuron group generate spikes at the different timing when the inhibition is weakest. This corresponds to the timing point when the distance between spikes expressing *black* and *white* is longest. This time span is equal to $T_d/2$.

On the other hand, if the input spike-timing range is larger than T_d , some neurons output spikes before all neurons receive initial input spikes. These output spikes affect other neurons' operations, and as a result, the network may not perform the correct memory operation. Therefore, the input spike-timing range must be smaller than T_d .

6. Conclusions

We proposed a spiking feedback neural network using integrate-and-fire neurons with negative thresholding. A global excitatory unit was introduced for the negative thresholding. We designed a CMOS spiking feedback network VLSI circuit with the global excitatory unit for Hopfield-type associative memory. The circuit simulation results verified that the network operates as an associative memory. The network has a characteristic that the output spike-timing difference is independent of the input spiketiming range. The spiking feedback network proposed in this paper will construct a different type of associative memory from conventional rate-coding-based neural networks, and its VLSI hardware lead to new applications.

References

- D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, Parallel Distributed Processing, The MIT Press, Cambridge, MA, 1986.
- [2] T. Morie and Y. Amemiya, "An all-analog expandable neural network LSI with on-chip backpropagation learning," IEEE J. Solid-State Circuits, vol.29, no.9, pp.1086–1093, 1994.
- [3] A.F. Murray and L. Tarassenko, Analogue Neural VLSI—A Pulse Stream Approach, Chapman & Hall, London, UK, 1994.
- [4] Y. Hirai and M. Yasunaga, "A PDM digital neural network system with 1,000 neurons fully interconnected via 1,000,000 6bit synapses," Proc. Int. Conf. on Neural Information Processing (ICONIP), pp.1251–1256, 1996.
- [5] W. Maass, "Networks of spiking neurons: The third generation of neural network models," Neural Netw., vol.10, no.9, pp.1659–1671, 1997.
- [6] W. Maass and C.M. Bishop, ed., Pulsed Neural Networks, The MIT Press, Cambridge, MA, 1999.
- [7] S.J. Thorpe and J. Gautrais, "Rapid visual processing using spike asynchrony," in Advances in Neural Information Processing Systems, ed. M.C. Mozer, M.I. Jordan, and T. Petsche, vol.9, pp.901– 907, The MIT Press, 1997.
- [8] H. Hasegawa, "An associative memory of hodgkin-huxley neuron networks with willshaw-type synaptic couplings," J. Physical Society of Japan, vol.70, no.7, pp.2210–2219, 2001.
- [9] Y. Li and J.G. Harris, "A spiking recurrent neural network," IEEE Computer Society Annual Symposium on VLSI Emerging Trends in VLSI Systems Design, pp.321–322, Feb. 2004.
- [10] K. Sasaki, T. Morie, and A. Iwata, "A spiking neural network with negative thresholding and its application to associative memory," 2004 IEEE Int. Midwest Symposium on Circuits and Systems (MWSCAS2004), pp.III-89–III-92, Hiroshima, July 2004.
- [11] B.G. Wallace, J.G. Nicholls, A.R. Martin, and P.A. Fuchs, From Neuron to Brain, Sinauer Associates, Inc., Massachusetts, USA, 2001.



Kan'ya Sasaki received the M.E. degree in advanced sciences of matter from Hiroshima University in 2004. He is currently pursuing a Ph.D. at the Department of Semiconductor Electronics and Integration Science, Graduate School of Hiroshima University.



Takashi Moriereceived the B.S. andM.S. degrees in physics from Osaka University, Osaka, Japan, and the Dr.Eng. degree fromHokkaido University, Sapporo, Japan, in 1979,1981 and 1996, respectively.From 1981 to1997, he was a member of the Research Staffat Nippon Telegraph and Telephone Corpora-tion (NTT).From 1997 to 2002, he was an as-sociate professor of the department of electri-cal engineering, Hiroshima University, Higashi-Hiroshima, Japan.Since 2002 he has been a

professor of Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Kitakyushu, Japan. His main interest is in the area of VLSI implementation of neural networks, mixed/merged analog-digital circuits, and new functional devices. Dr. Morie is a member of IEEE, the Japan Society of Applied Physics and the Japanese Neural Network Society.



Atsushi Iwata received the B.E., M.S. and Ph.D. degrees in electronics engineering from Nagoya University, Nagoya, Japan, in 1968, 1970, and 1994, respectively. From 1970 to 1993, he was with the Electrical Communications Laboratories, Nippon Telegraph and Telephone Corporation. Since 1994 he has been a professor of electrical engineering at Hiroshima University, Hiroshima, Japan. His research is in the field of integrated circuit design where his interests have included circuit architecture and

design techniques for analog-digital mixed LSIs, wireless interconnections, substrate noise reduction, 3D integrated image processors and bio-inspired signal processing LSIs. He was the Program Chairman for the 1995 Symposium on VLSI Circuits and the Chairman for the 1999 Symposium on VLSI Circuits. He was an Associate Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS. Dr. Iwata received an Outstanding Panelist Award for the 1990 International Solid-State Circuits Conference. He is a member of the Japanese Neural Network Society.