Inhibitive Sensory Neuron for Energy Efficient Spiking Neural Network with Time-to-first-spike Coding

Se-Hyun Hwang, Wonbin Jo, Kang-Uk Byeon, and Joon-Kyu Han*

Department of Electronic Engineering, Sogang University, Seoul 04107, Korea

E-mail: joonkyuhan@sogang.ac.kr

Neuromorphic computing is being investigated as a way to improve the energy efficiency of traditional von Neumann computing through the use of spiking neural networks (SNNs). In typical SNNs, rate coding is commonly employed as an encoding scheme, where information is transmitted based on the spiking frequency of input neurons. However, rate coding increases energy consumption since it requires neurons to generate repetitive spikes. This study introduces an inhibitive sensory neuron and a time-to-first-spike (TTFS) coding using it to enhance the energy efficiency of SNNs. When light is applied to the neuron transistor, which operates based on a single transistor latch (STL) phenomenon, electron-hole pairs are generated by photons that reduce the potential barrier between the source and body, allowing for faster spike generation [1]. Moreover, when a high gate voltage (V_G) is applied, spiking is inhibited because the potential barrier is eliminated [2]. By using these two mechanisms, an inhibitory sensory neuron is developed that generates only the first spike and prevents subsequent spikes. This is achieved through a feedback circuit that detects the initial spike and then applies a high V_G. Using TTFS coding significantly reduces the number of spikes generated in the SNN, thereby lowering overall energy consumption. After evaluating TTFS coding through electrical measurements, an MNIST pattern classification is performed using software simulation to demonstrate the effectiveness of TTFS coding.

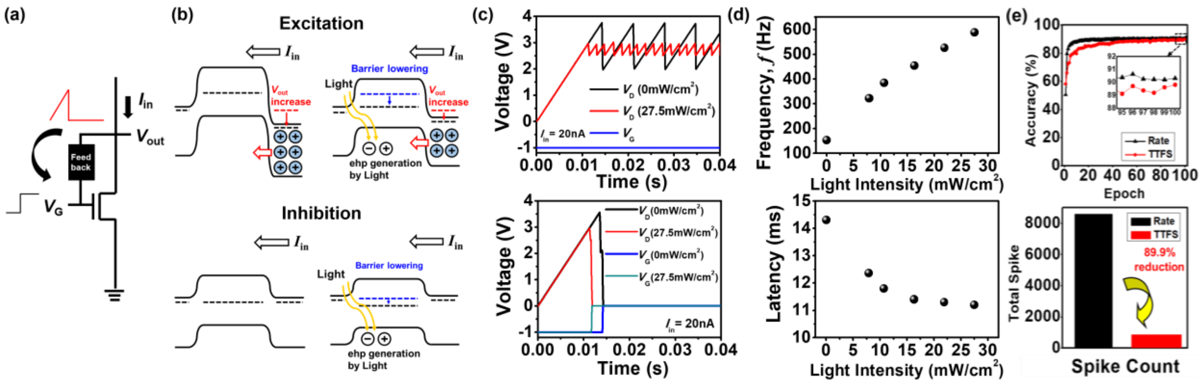

Fig. 1. (a) Circuit diagram of an inhibitory sensory neuron. (b) Energy band diagrams illustrating the principle of inhibitive sensory neuron. (c) Spiking characteristics for rate coding and TTFS coding. (d) Spiking frequency (rate coding) and first spike latency (TTFS coding) as a function of the light intensity. (e) Accuracy and total spike number for classifying MNIST patterns using rate coding and TTFS coding.

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) grant [RS-2024-00334953, 2009-0082580] and Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (Ministry of Education-Ministry of Trade, Industry and Energy) [P0025169, Semiconductor Major Track].

References [1] Han et al., *Nano Letters*, 20, 12 (2020). [2] Han et al., *IEEE Electron Device Letters*, 41, 2 (2019).