International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Improving the Variation-Tolerance of Memristor Synapse by Selecting the Optimal Memristance

Son Ngoc Truong

Abstract*: Memristor crossbar array that has a lot of advantageous features such as non-volatile, high-density, and low-power, is potentially used for realizing artificial neural networks. One of the important factors affecting the performance of crossbar circuits is memristance variation. The variation of memristance causes the variation of synaptic weights resulting in the degradation of the performance of the neural networks. In this paper, the variation-tolerance of memristor synapse is improved by minimized the gradient of the synaptic weight function. To reduce the gradient of synaptic weight function, the memristance values of memristors must be close to the high resistance state. This can be achieved by selecting the value of bias resistance because memristance values of memristors in crossbar-based neural network are distributed around the bias resistance value. In the simulation result, we measure the recognition of crossbar circuit for recognizing the MNIST handwritten digits with the variation is in range of 0% to 10% and the value of bias resistance varies from 30KΩ to 80KΩ. For the bias resistance is as low as 30KΩ, the recognition rate of the crossbar circuit is as low as 17%, when the variation is 10%. When the bias resistance is 80KΩ, the recognition is 70%, when the variation is 10%. When the range of memristance is close to the high resistance state, the recognition rate is improved by 53% when the variation is 10%.

***Keywords*: Memristor, memristive synapse, variation-tolerance, neural network.**

I. INTRODUCTION

Artificial neural network is a multidisciplinary field that is inspired by the perception mechanism of human brain for a variety of applications. Artificial neural networks are predominantly realized by software since they are based on several complicated algorithms. However, the artificial neural network requires a large power consumption and high resource for the huge number of computational tasks such as addition and multiplication. Hardware realization of artificial neural networks is an alternative approach to accelerate the performance of the artificial neural networks. VLSI (Very-Large Scale Integration) implementations of artificial neural networks have gained more advantage in the past two decades [1]-[4]. However, the multiplication and addition are costly to implement in hardware [1]. Furthermore, the VLSI technology which is predominantly based on the CMOS (Complementary-Metal-Oxide Semiconductor) is approaching the end of their capabilities because scaling CMOS down faces several fundamental limiting factors stemming from electron thermal energy and quantum-mechanical tunneling [5],[6]. The nano-scale memristor has been considered as a potential device for realizing artificial neural networks. Memristor was mathematically postulated by Leon O. Chua in 1971 and experimentally demonstrated by HP Lab in 2008 as the fourth fundamental circuit element [7], [8]. The flexibly modifiable conductance of memristor is ideal for mimicking the synaptic plasticity of biological synapse in nervous system [9]-[13]. Memristor-based synapse has been extensively studied as a potential solution to overcome the limitation of VLSI-based synapse in artificial neural networks.

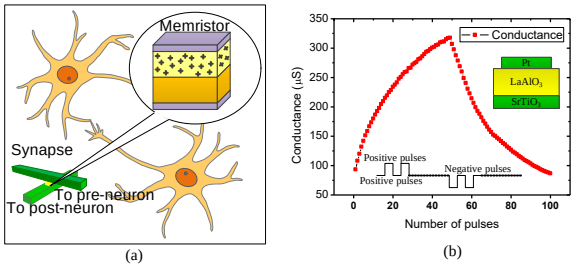

Fig. 1.(a) a conceptual diagram of a memristor-based synapse [9]. (b) the change of a memristor's conductance with positive pulse and negative pulses [17].

In nervous systems, synapses are the connections between the presynaptic neurons and the postsynaptic neurons. The strength of the synapses is represented by the synaptic weight. According to the neuron activities including excitatory and inhibitory, synaptic weights can be positive or negative [14]-[16]. Synapses can be modeled by memristors as shown in Fig. 1(a) [9], [10]. The synaptic weight is represented by the conductance of memristors. The conductance of memristors can be modified by applying the positive pulses or negative pulses to the devices. Fig 1(b) shows a measurement of a memristor based on a $LaAlO_3$ film [17]. The conductance of the measured memristor increases when positive pulses are applied to device. By contrast, the conductance decreases as negative pulses are applied to the device, as shown in Fig. 1(b). The memristor circuits realizing the signed synaptic weights have been demonstrated in last decades [18]-[21]. Memristors might be programmed to an undesired memristance values resulting in the process variation [22]-[26]. Memristance variation is one of the factors that degrade the performance of memristor synapses seriously. Furthermore, variation also known as the drift of memristance during read operation [27].

Manuscript published on November 30, 2019.
* Correspondence Author
Son Ngoc Truong*, Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology and Education, Vietnam.

The variation of synaptic is represented by the gradient of synaptic function with respect to the memristance value. To improve the variation-tolerance of memristor synapse, in this paper, we propose an optimal active memristance range, in which memristance value is distributed close to the high resistance state.

II. MEMRISTOR SYNAPSES

Memristor synapse circuits which can perform negative and positive synaptic weightings has been proposed. Fig. 2(a) shows a synapse circuit constituted by 4 memristors for performing zero, positive, and negative synaptic weightings [18], [19]. The synaptic weight is determined by the memristance values of 4 memristors, as indicated in Equation 1 [18], [19].

$$V_{out} = V_{in} w$$

$$w = \frac{M_2}{M_1 + M_2} - \frac{M_4}{M_3 + M_4} \quad (1)$$

By adjusting the memristance values of M_1, M_2, M_3, and M_4, we obtain a zero, positive, or negative synaptic weight. The memristor synapse shown in Fig. 2(a) is interesting, however, it is not suitable for implementing high-density synaptic arrays. For memristor-based synapse array, we can consider the complementary architecture of two memristors as shown in Fig. 2(b) [20]. Here two memristors in two crossbar arrays constitute a synapse for performing negative, zero, or positive synaptic weightings. The synaptic weight in Fig. 2(b) is decided by the difference between two memristance values in two crossbar arrays [20].

$$V_o = \sum v_{in,j} w_j$$

$$Where, w_j = \frac{1}{M_3} - \frac{1}{M_4} \quad (2)$$

To reduce the power consumption and area, S. N. Truong proposed a new memristor synapse which employs only one memristor and a constant-term [21]. The schematic of the proposed memristor synapse is shown in Fig. 2(c).

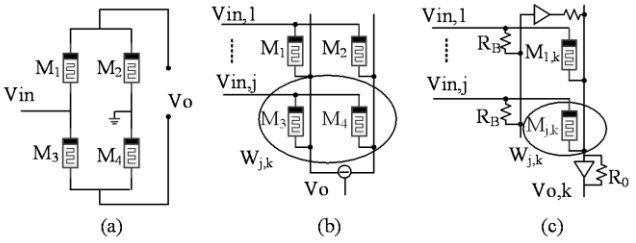

Fig. 2. The memristor-based synapses for implementing the signed synaptic weights. (a). A memristor bridge circuit constituted by four memristors for implementing the signed synaptic weights [18], [19]. (b). A complementary architecture of two memristors for implementing the signed synaptic weights [20]. (c) A new synaptic circuit that utilizes only one memristor [21].

Here a negative voltage is generated by a constant-term circuit sharing for all columns in crossbar array [21]. The synaptic weight in Fig. 2(c) is calculated using the Equation 3 [21].

$$w_{j,k} = R_0 \left(\frac{1}{R_B} - \frac{1}{M_{j,k}} \right) \quad (3)$$

The memristor circuit in Fig. 2(c) can perform a negative and or positive synaptic weight by adjusting the value of memristance [21]. R_B is constant.The memristor synapse circuit in Fig. 2(c) is used to implements a two-layer neural network for the application of handwritten digit recognition is shown in Fig. 3. Fig. 3(a) shows the samples of MNIST dataset of handwritten digits. The character images are resized to 14x14 pixels. Figure 3(b) show a conceptual diagram of a two-layer neural network. The input layer has 196 inputs for 196 pixels, the hidden layer is composed of 256 neurons, and the output layer has 10 neurons for recognizing 10 digits.

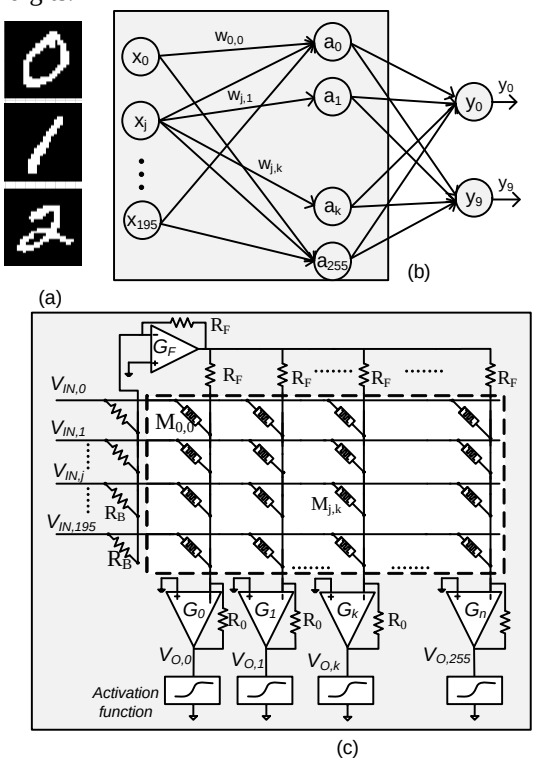

Fig. 3. (a) The handwritten digits of MNIST dataset, (b). The conceptual diagram of a multiple neural network for handwritten digit recognition, and (c) the schematic of a memristor crossbar circuit for implementing a hidden layer in a two-layer neural network. Fig. 3(c) shows a memristor crossbar circuit for implementing the first layer of neural network. The memristor crossbar has 196 rows for the input neurons, 256 columns corresponding to 256 neurons in hidden layer. The output of the kth neuron is calculated as follows [21]

$$V_{O,k} = \sum_{j=1}^{m} w_{j,k} V_{IN,j}$$
$$where\ w_{j,k} = R_0\left(\frac{1}{R_B} - \frac{1}{M_{j,k}}\right) \quad (4)$$

In Equation 4, the synaptic weight is determined by the value of R_B, R_0, and the memristance value. R_0 is a coefficient for amplifying the synaptic weight. R_B is a bias resistance. The signed synaptic weight is obtained by adjusting the valued of memristance, $M_{j,k}$, around the bias resistance R_B. It emerges that the synaptic weight is a nonlinear function of R_B and $M_{j,k}$. The change of memristance makes the synaptic weight change nonlinearly. Here R_B can be selected between low resistance state (LRS) and high resistance state (HRS). The bias resistance, R_B, strongly affects the distribution of memrsitance values. In this work, we will discover the range of R_B so as to reduce the effect of the variation of $M_{j,k}$. To do this, we calculated the gradient of the synaptic weight function in Equation 4 with respect to the memristance value. We obtain the Equation 5.

$$\frac{\partial w}{\partial M_{j,k}} = \frac{1}{M_{j,k}^2} \quad (5)$$

Equation 5 presents a nonlinear relationship between the variation of the memristance and the variation of the corresponding synaptic weight. The gradient of synaptic weight function with respect to the memristance is a function of memristance as indicated in Equation 5. The gradient of synaptic weight represents the variation of synaptic weight caused by the variation of memristance value. The relationship of the gradient of the synaptic weight and the memristance value in Equation 5 is shown in Fig. 4. As we can see in Fig. 4, the gradient decreases as memristance increases. To minimize the gradient of synaptic weight function, memristance should be close to the high resistance state, as shown in Fig. 4. The active range of memristance is decided by the value of bias resistance R_B since the memristance values are distributed around the value of R_B.

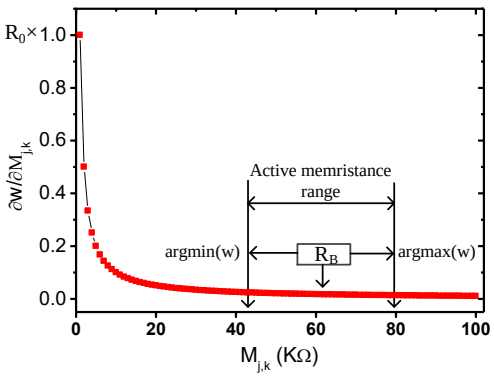

Fig. 4. The gradient of synaptic weight function with respect to the memristance valueThe synaptic weights obtained by the training process are then converted to the memristance values using Equation 4. The valued of R_0 and R_B are selected so as to the achieved memristance values must be in a range of from LRS to HRS. In Equation 4, R_B and R_0 are constant, the memristance increases makes the synaptic weight increase. The synaptic weights obtained from the training process can be negative, zero, and positive. These synaptic weights are converted to the values of memristance using Equation 4. Therefore, the memristance values are distributed around the value of R_B for perform negative, zero, and positive synaptic weightings. As we can see in Fig. 4, if R_B is close to HRS, memristance values are distributed close to HRS and the gradient of synaptic weights are very small. The value of R_B should be selected so as to satisfy that the maximum synaptic weight can be converted to the memristance value which must be less than or equal to HRS. Thereby, we can choose the value of R_0 and R_B according to the following assumption.

$$\arg\max(w) = HRS \quad (7)$$

The optimal bias resistance is calculated using the Equation 7. Where the maximum of synaptic weight is converted to the value of memristance which is equivalent to the highest resistance of memristor device. The active memristance is defined as a range of memristance that can represent the minimum and the maximum of synaptic weights, as shown in Fig. 4.

(

III. RESULT AND DISCUSSION

The two-layer neural network with memristor synapses in Fig. 3 is tested for handwritten digit recognition. The crossbar circuit is simulated by using Cadence Spectre [28]. Memirstor is modelled using Verilog-A model [17]. The neural network is implemented using two crossbar circuits. The hidden layer is realized by a crossbar circuit presented in Fig. 3(c). The 256 neurons in hidden layer are realized by 256 columns. The activation circuits are realized using Op-amp circuits [29]. The output layer can be designed using the crossbar array as the hidden layer. The crossbar circuit is trained by 60,000 training samples. The number of testing samples is 10,000. The recognition rate of the crossbar circuit without variation is as high as 97%. The variation of memristance increases from 0% to 10%. The distribution of memristance values is controlled by the value of the bias resistance R_B. In this simulation, the recognition rate is measured with varying the value of R_B from 30KΩ to 80KΩ. When the R_B is set to be 30KΩ, the recognition rate degrades sharply when the memristance variation increases. It is due to the fact that the distribution of memristances close to LRS resulting in the high gradient of synaptic weight, as presented in previous section. However, when the active memristance range is close to high resistance value determined by R_B=80KΩ, the recognition rate is as high as 70% when the variation is 10%. To improve the variation-tolerance of memristor synapses, the active memristor range should be close to the HRS. This can be done by choosing the value of R_B close to HRS and satisfy the Equation 6.

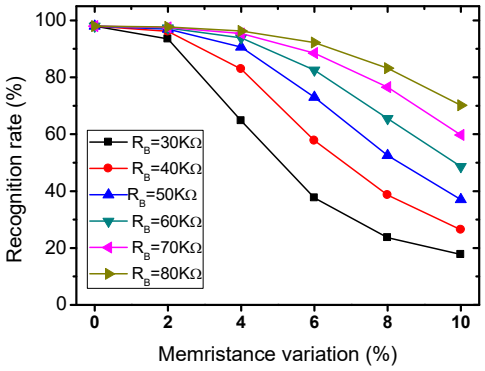

Fig. 5. The comparison of the recognition rate when the variation increases from 0% to 10% among different active memristance ranges selected by different values of R_B

The distribution of memristance values is decided by the value of bias resistance. The gadient of synaptic weigt function is minimized if memristance values move close to HRS. The optiaml range of memristance is obtanied by selecting the value of R_B using Equation 6. By using the optimal active memristance range, the varition-tolerance is improved significantly.

IV. CONCLUSION

In this work, we proposed an optimal active memristance range for improving the variation-tolerance of memristor synapses in a memristor-based neural network. The effect of memristance variation is mitigated by minimizing the gradient of synaptic weight function. To minimize the gradients, memristance values should be distributed close to the high resistance state. By selecting the appropriate memristance range which is close to the high resistance state, the recognition rate of crossbar circuit is improved by 53% when variation increases to 10%.

REFERENCES

1. A. Ardakani et al, "VLSI Implementation of Deep Learning Neural Network Using Integral Stochastic Computing," IEEE Trans. Very Large Scale Integration System, vol. 25, iss. 10, pp. 2688-2699, Feb. 2017
2. M. Walker, P. Hasler, and L. Akers " CMOS neural network for pattern association," IEEE Micro, vol. 9, no. 5, pp. 68-71, Oct. 1989
3. B. V. Benjamin et al, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," Proc. IEEE, vol. 102, no.5, pp.-699-716, Apr. 2014
4. P. A. Merolla et al, "A million spiking- neuron integrated circuit with a scalable communication network and interface," Science, vol. 345, no. 6197, pp 668-673, Aug. 2014
5. P. M. Solomon, "Device Innovation and Material Challenges at the Limit of CMOS Technology," Annu. Rev. Mater. Sci., vol. 30, pp. 681-697, Aug. 2000
6. T. P. Brđanin and B. Dokić, "Strained Silicon Layer in CMOS Technology," Electronics, vol. 18, no. 2, pp. 63-69, Dec. 2014
7. L. O. Chua, "Memristor – the missing circuit element," IEEE Trans. Circuit Theory, vol. CT-18, no. 5, pp. 507-519, Sep. 1971.
8. D. B. Strukov, G. S. Sinder, D. R. Stewart, and R. S. Williams, "The missing memristor found," Nature, vol. 453, pp. 80-83, May 2008.
9. S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," Nano Letters, vol. 10, no. 4, pp. 1297-1301, Mar. 2010.
10. D. Kuzum, R.G.D. Jeyasingh, B. Lee, and H.S.P. Wong, "Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing," Nano Lett., vol. 12, no. 5 pp. 2179–2186, May 2012
11. K. D. Cantley, A. Subramaniam, H. J. Stiegler , R. A., Chapman, and E. M. Vogel, "Hebbian learning in spiking neural networks with nanocrystalline silicon TFTs and memristive synapse," IEEE Trans. Nanotechnology, vol. 10, no. 5, pp. 1066-1073, 2011.
12. G. S. Snider, "Self-organized computation with unreliable, memristive nanodevices," Nanotechnology, vol. 18, no. 36, pp. 1-13, 2007.
13. H. Kim, M. P. Sah, C. Yang, T. Roska, and L. O. Chua, "Neural synaptic weighting with a pulse-based memristor circuit," IEEE Trans. on Circuit and Systems-I , vol. 59, no.1, pp. 148-158, 2012.
14. L. F. Abbott and W. G. Regehr, "Synaptic computation," Nature, vol. 431, 796-803, 2004.
15. R. S. Zucker, W. G. and Regehr, "Short-term synaptic plasticity," Annu. Rev. Physiol.vol. 64, pp. 355–405, 2002.
16. R. Lamprecht and J. LeDoux, "Structural plasticity and memory," Nat. Rev. Neurosci, vol. 5, no. 1, 45-54, 2004.
17. S. N. Truong K. V. Pham , W. S. Yang, S. H. Shin, K. Pedrotti, and K. S. Min, "New pulse amplitude modulation for fine tuning of memristor synapses," Microelectronics Journal, vol. 55, pp. 162-168, 2016
18. H. Kim, M. P. Sah, C. Yang, T. Roska, and L. O. Chua, "Neural synaptic weighting with a pulse-based memristor circuit," IEEE Trans. Circuit Syst. I, vol. 59, no. 1, pp. 148–158, Jan. 2012
19. S. P. Adhikari, C. Yang, and H. S. Kim, "Memristor Bridge Synapse-Based Neural Network and Its Learning," IEEE Trans. Neural Network and Learning Systems, vol. 23, no. 9, pp. 1426-1435, Sep. 2012
20. M. Hu, H, Li, Q. Wu, G. S. Rose, and Y. Chen, "Memristor crossbar based hardware realization of BSB recall function," International Joint Conference on Neural Networks, pp. 1-7, Jun. 2012.
21. S. N. Truong and K. S. Min "New memristor-based crossbar array architecture with 50-% area reduction and 48-% power saving for matrix-vector multiplication of analog neuromorphic computing," Journal of Semiconductor Technology and Science, vol. 14, no. 3, pp. 356-363, Jun. 2014.
22. D. Niu, Y. Chen, C. Xu, and Y. Xie, "Impact of Process Variations on Emerging Memristor," IEEE/ACM Design Automation Conference, Jun. 2010, pp. 877-882.
23. M. Hu, H. Li, Y. Chen, and R. E. Pino, "Geometry variations analysis of TiO2 thin-film and spintronic memristors," IEEE Asia and South Pacific Design Automation Conference, 2011, pp. 25-30.
24. J. Rajendran, H. Manem, R. Karri, and G. S. Rose, "An Approach to Tolerate Process Related Variations in Memristor-Based Applications," IEEE International Conference on VLSI Design, 2011, pp. 18-23
25. J. Rajendran, R. Karri, and G. S. Rose "Improving Tolerance to Variations in Memristorbased Applications Using Parallel Memristors," IEEE Trans. IEEE Transactions on Computers, vol. 64 , no. 3, pp. 733-746. Mar. 2015
26. S. N. Truong, S. H. Shin, S. D. Byeon, J. S. Song, H. S. Mo, and K. S. Min, "Comparative Study on Statistical-Variation Tolerance between Complementary Crossbar and Twin Crossbar of Binary Nano-scale Memristors for Pattern Recognition," Nanoscale Research Letters, vol.10, no. 1, pp. 405-413, Dec. 2015.
27. S. N. Truong, K. V. Pham, W. S. Yang, and K. S. Min, "Reduction of Memristance Drift in Twin Memristor Crossbar," Proceeding of the 15th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA), pp. 63-64, Dresden, Germany, Aug. 2016
28. Virtuoso Spectre Circuit Simulator User Guide, Cadence, San Jose, CA, USA, 2004.
29. K. V. Pham, T. V. Nguyen, S. B. Tran, H. K. Nam, M. J. Lee, B. J. Choi, S. N. Truong, and K. S. Min, "Memristor binarized neural networks," Jounal of Seminconductor Technology and Science, vol. 18 no. 5, pp. 568- 577, Oct. 2018.

AUTHORS PROFILE

Son Ngoc Truong received the B.S. and M.S. degrees in Electronic Engineering from Ho Chi Minh City University of Technology and Education, Vietnam, in 2006 and 2011, respectively, and the Ph.D degree in Electronic Engineering from Kookmin University, Seoul, Korea, in 2016. He was a postdoctoral researcher fellow at Kookmin University, Seoul, Korea in 2016 and 2017. He is currently lecturer at Ho Chi Minh City University of Technology and Education, Vietnam. His research interests include circuits and systems with memristor device and memristor-based crossbars for neuromorphic computing systems.