

УДК 621.396:621.372

**ПОСТРОЕНИЕ СЕТИ ХЕММИНГА НА ОСНОВЕ КРОССБАРА
С БИНАРНЫМИ МЕМРИСТОРАМИ**

М. С. Тарков

Институт физики полупроводников им. А. В. Ржанова СО РАН, г. Новосибирск, Россия

Описаны свойства аналоговых и бинарных мемристоров (резисторов с памятью), которые могут быть использованы для аппаратной реализации синапсов нейронов, а также мемристорные матрицы, называемые кроссбарами. Бинарные мемристоровы, сопротивление которых принимает только два значения (максимальное и минимальное), основаны на механизме переключения филамента и распространены более широко, чем аналоговые мемристоровы. Они гораздо более устойчивы к статистическим флуктуациям по сравнению с аналоговыми. Предложена аппаратная реализация ассоциативной памяти Хемминга на основе использования кроссбара на бинарных мемристорах и КМОП-схемотехники. Максимальное сопротивление бинарного мемристора соответствует значению -1 компоненты хранимого эталонного вектора, а минимальное — значению $+1$. Показано, что кроссбар на бинарных мемристорах реализует свойства первого слоя сети Хемминга, согласно которым выходной сигнал нейрона первого слоя неотрицателен. При этом он максимален для нейрона, эталонный вектор которого наиболее близок к вектору входных данных. Для заданной размерности эталонного вектора получено соотношение между максимальным и минимальным сопротивлениями бинарных мемристоров, которое гарантирует корректную работу первого слоя сети Хемминга. Моделирование в системе LTSPICE предложенной схемы памяти Хемминга подтвердило её работоспособность.

Ключевые слова: ассоциативная память Хемминга, мемристор, кроссбар, КМОП-технология, LTSPICE.

DOI 10.17223/20710410/40/9

**CONSTRUCTION OF A HAMMING NETWORK BASED
ON A CROSSBAR WITH BINARY MEMRISTORS**

M. S. Tarkov

*Rzhanov Institute of Semiconductor Physics SB RAS, Novosibirsk, Russia***E-mail:** tarkov@isp.nsc.ru

The properties of analog and binary memristors (resistors with memory) are described. The memristors can be used for the hardware implementation of neurons synapses. The memristor matrices are called crossbars. The binary memristors, whose resistance takes only two values (maximum and minimum), are based on the switching filament mechanism and are distributed more widely than analog memristors. They are much more stable to statistical fluctuations compared to analog memristors. The Hamming associative memory's hardware realization based on the use of a binary memristors crossbar and CMOS circuitry is proposed. The maximum binary memristor resistance corresponds to the stored reference vector component value -1 , and the minimum resistance corresponds to the value $+1$. It is shown that the binary memristors crossbar

realizes the Hamming network first layer properties according to which the output first layer neuron signal is non-negative. This signal is maximal for a neuron with the reference vector closest to the input vector. For a given reference vector dimension, the relationship between the maximum and minimum binary memristors resistances is obtained. It guarantees the Hamming network first layer correct operation. Simulation in the LTSPICE system of the proposed Hamming memory scheme confirmed its operability.

Keywords: *associative Hamming memory, memristor, crossbar, CMOS-technology, LTSPICE.*

Введение

Искусственная нейронная сеть обычно использует матрицу весовых коэффициентов для представления множества синапсов слоя нейронов. Соответственно вычисление активации слоя нейронов можно рассматривать как умножение этой матрицы весов на вектор входных сигналов слоя. Аппаратная реализация нейронной сети требует много памяти для хранения матрицы весов слоя нейронов и является дорогостоящей. Решение этой проблемы упрощается при использовании в качестве ячейки памяти устройства, называемого мемристором. Мемристор был предсказан теоретически в 1971 г. Леоном Чуа [1]. Первую физическую реализацию мемристора продемонстрировала в 2008 г. лаборатория фирмы «Hewlett Packard» в виде тонкоплёночной структуры TiO_2 [2]. В России первый мемристор на основе TiO_2 получен в 2012 г. в Тюменском государственном университете [3]. Мемристор имеет много достоинств, таких, как энергонезависимость хранения информации, малое потребление энергии, высокая плотность интеграции и замечательная масштабируемость. Уникальная способность сохранять следы возбуждения устройства делает его идеальным кандидатом для реализации синапсов в электронных нейронных сетях [4].

Мемристор ведет себя подобно синапсу: он «запоминает» полный электрический заряд, прошедший через него [5]. Память, основанная на мемристорах, может достигать очень высокой степени интеграции 100 Гбит/см^2 , в несколько раз более высокой, чем на основе технологии флэш-памяти [6]. Эти уникальные свойства делают мемристор многообещающим устройством для создания массово-параллельных нейроморфных систем [7–9].

Понятие мемристора покрывает широкий диапазон явлений [10]:

- а) изменение фазы, вызванное нагреванием в устройствах памяти, основанных на халькогенидах;
- б) образование проводящего филамента, вызванное электрохимическими процессами окисления/восстановления, наблюдаемыми в оксидах металлов, халькогенидах и полимерах;
- в) дрейф/диффузия ионов в аморфных плёнках;
- г) образование проводящего филамента, вызванное нагреванием и наблюдаемое в некоторых оксидах.

В «аналоговых» мемристорах (например, на основе окиси титана TiO_2) непрерывное изменение напряжения вызывает непрерывное и постепенное (инкрементное) изменение проводимости. Тот же эффект достигается при применении последовательности положительных или отрицательных программирующих импульсов. Сопротивление аналогового мемристора (рис. 1) может быть представлено [2] как

$$M(p) = p \cdot R_{\text{он}} + (1 - p)R_{\text{офф}},$$

где p ($0 \leq p \leq 1$) — положение фронта примеси относительно полной толщины h плёнки TiO_2 ; $R_{\text{он}}$ и $R_{\text{оф}}$ — минимальное и максимальное сопротивления мемристора. Когда на мемристор подаётся напряжение V выше некоторого порога V_{th} , его сопротивление уменьшается за счёт расширения примесной зоны D , имеющей низкое сопротивление, и сокращения зоны U чистого окисла, имеющей высокое сопротивление. Соответственно сопротивление мемристора возрастает при подаче напряжения V ниже $-V_{\text{th}}$ за счёт сокращения зоны D и расширения зоны U . После отключения напряжения текущее сопротивление мемристора сохраняется.

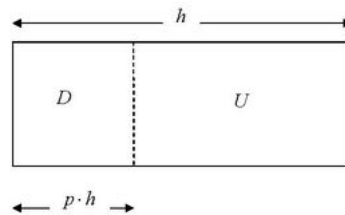


Рис. 1. Мемристор

В «цифровых» (бинарных) мемристорах сопротивление изменяется резко (с низкого на высокое и обратно), когда приложено напряжение, «пересекающее» порог. Скачкообразное изменение сопротивления обусловлено образованием проводящего филамента. Например, при образовании Ag-филамента в аморфном кремнии (a-Si, который служит материалом электролита) наблюдается устойчивый эффект резкого изменения сопротивления. Такой мемристор имеет чёткий порог переключения и резкое изменение сопротивления при напряжении считывания 1 В. Время переключения составляет менее 10 нс. Особенно интересны реализации синаптических устройств на окиси гафния HfO_2 , поскольку такая память имеет блестящие характеристики: большая масштабируемость (до размеров менее 10 нм), малое время переключения (порядка 1 нс).

Бинарные мемристоры [11, 12], основанные на механизме переключения филамента, распространены более широко, чем аналоговые, материалы для которых встречаются значительно реже и требуют более сложного процесса обработки. При переключении филамента мемристоры могут иметь либо высокое сопротивление (ВС), либо низкое (НС). Таким образом, мы можем хранить только 1 или 0 (–1) при переключении филамента бинарного мемристора. Бинарные мемристоры гораздо более устойчивы к статистическим флуктуациям по сравнению с аналоговыми. Это связано с тем, что ВС намного выше, чем НС, несмотря на большое количество флуктуаций в НС и ВС.

На рис. 2 представлен мемристорный кроссбар, который является типичной структурой мемристорной памяти. Он содержит мемристор на каждом пересечении горизонтальных и вертикальных проводников. Мемристорный кроссбар интересен для реализации матриц соединений в нейронных сетях, поскольку он может обеспечить большое число сигнальных связей и вычислить взвешенную комбинацию входных сигналов. На вертикальные шины кроссбара подаются компоненты вектора входных сигналов. Каждая горизонтальная шина позволяет вычислить скалярное произведение вектора входных сигналов на вектор весов, заданных проводимостями мемристоров соответствующей строки кроссбара.

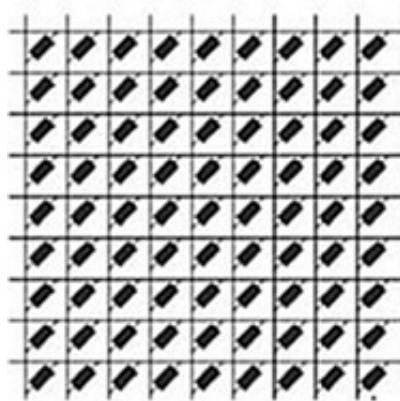


Рис. 2. Мемристормый кроссбар

1. Программирование кроссбара на бинарных мемристорах

Установка сопротивлений бинарных мемристоров кроссбара осуществляется на основе методики, предложенной в [12]. Рассмотрим установку мемристора в состояние +1 (НС) на примере фрагмента кроссбара, представленного на рис. 3. Для определённости рассмотрим мемристор M11. Чтобы сделать низким сопротивление этого мемристора, сделаем установки напряжений, указанные в таблице (в вольтах):

H1	H2	H3	V1	V2	V3
+7	0	0	-7	+7	+7

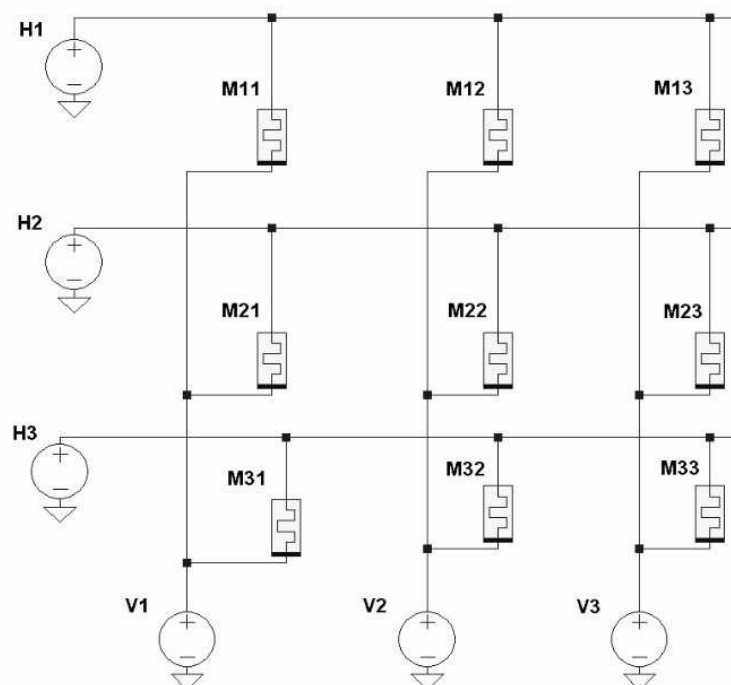


Рис. 3. Фрагмент кроссбара

Эксперимент показал, что при таких установках напряжений сопротивление мемристора M11 станет минимальным (100 Ом) в течение 2 нс. При этом сопротив-

ления мемристоров M12 и M13 останутся неизменными, а сопротивления остальных мемристоров (M21, M22, M23, M31, M32, M33) изменятся незначительно (не более 0,5 %).

Чтобы установить мемристор M11 в состояние -1 (BC), нужно поменять полярность напряжений в таблице. При этом время установки увеличивается до 5 нс.

2. Сеть Хемминга и её реализация с использованием мемристорного кроссбара

В сети Хемминга [13] нейроны первого слоя рассчитывают расстояния Хемминга между поданным на вход сети вектором x и векторами весов (эталонными векторами) x^i нейронов этого слоя. Расстоянием Хемминга $d(x, y)$ между векторами x и y называется количество несовпадающих компонент этих векторов. Значения выходных сигналов нейронов первого слоя определяются по формуле

$$y_i = 1 - d(x^i, x)/N, \quad (1)$$

где N — число компонент вектора x . Из (1) следует, что $y_i \in [0, 1]$, причём для $x = x^i$ выполняется $y_i = 1$ (максимально возможное значение).

Сигналы y_i становятся начальными состояниями нейронов второго слоя, которые функционируют по принципу WTA («Winner Takes All»). Этот слой определяет победителя среди нейронов первого слоя, т. е. нейрон, выходной сигнал которого наиболее близок к 1. Такой нейрон указывает на эталонный вектор x^i с минимальным расстоянием Хемминга до входного вектора x .

Поставим в однозначное соответствие эталонному вектору сети Хемминга x^i , $i = 1, \dots, p$, вектор весов w^i , задаваемых проводимостями мемристоров кроссбара, по правилу

$$w_j^i = \begin{cases} g_L, & \text{если } x_j^i = -1, \\ g_H, & \text{если } x_j^i = 1, \end{cases}$$

где $g_H > g_L$, $j = 1, \dots, N$. Пусть w — произвольный вектор размерности N с компонентами g_L и g_H .

Утверждение 1. $(x^i, w^i) - (x^i, w) \geq g_H - g_L > 0$, если $w \neq w^i$.

Доказательство. Пусть m компонент вектора w^i равны g_H , а n компонент — g_L , $m + n = N$. Согласно определению вектора x^i , имеем

$$(x^i, w^i) = mg_H - ng_L. \quad (2)$$

Чтобы получить из w^i вектор w , нужно заменить k ($k \in \{1, \dots, m\}$) компонент g_H на g_L и/или l ($l \in \{1, \dots, n\}$) компонент g_L на g_H . Тогда

$$(x^i, w) = (x^i, w^i) + k(g_L - g_H) - l(g_H - g_L). \quad (3)$$

Из (2) и (3) следует $(x^i, w^i) - (x^i, w) = (k + l)(g_H - g_L) \geq g_H - g_L > 0$, так как $g_H > g_L$. ■

Из утверждения 1 следует, что вычисления по формуле (1) можно заменить на вычисление скалярного произведения (x, w^i) , которое может быть реализовано на мемристорном кроссбаре. Мемристорный кроссбар с указанными выше свойствами не реализует формулу (1) вычислений первого слоя сети Хемминга, но обеспечивает требуемое свойство этого слоя: нейрон с весовым вектором w^i , представленный i -й строкой кроссбара, для эталонного вектора w^i дает максимальный выходной сигнал $y_i = (x^i, w^i)$.

Будем называть вектор w нетривиальным, если хотя бы одна его компонента равна g_H . Тогда $(x^i, w) \geq g_H - (N - 1)g_L$, откуда следует

Утверждение 2. $(x^i, w) > 0$ для нетривиального w , если

$$g_L \leq \frac{g_H}{N-1}. \quad (4)$$

Неравенство (4) задаёт ограничение на отношение между максимальной g_H и минимальной g_L проводимостями мемристоров кроссбара, которое обеспечивает неотрицательность выходного сигнала нейрона первого слоя сети Хемминга. Предложенный здесь подход к построению ассоциативной памяти Хемминга позволил сократить вдвое количество используемых мемристоров по сравнению с реализацией в [14].

3. Пример реализация сети Хемминга

На рис. 4 представлены изображения символов «L», «T», «X» размером 3×3 пикселей. Эти изображения могут быть описаны эталонными векторами (построчная развертка изображений)

$$\begin{aligned} L &= (1, -1, -1, 1, -1, -1, 1, 1, 1), \\ T &= (1, 1, 1, -1, 1, -1, -1, 1, -1), \\ X &= (1, -1, 1, -1, 1, -1, 1, -1, 1), \end{aligned}$$

где 1 соответствует белым пикселям, а -1 — чёрным.

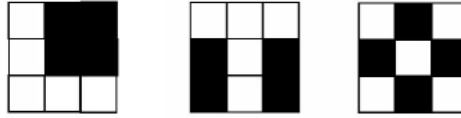


Рис. 4. Изображения символов

Запоминание этих изображений в памяти Хемминга реализовано мемристорным кроссбаром, представленным на рис. 5, где каждая строка соответствует одному из изображений. Компоненте 1 эталонного вектора в кроссбаре соответствует мемристор с минимальным сопротивлением $R_L = 100$ Ом (с максимальной проводимостью $g_H = 1/R_L$), а компоненте -1 — мемристор с максимальным сопротивлением $R_H = 100$ кОм (с минимальной проводимостью $g_L = 1/R_H$). Обозначив векторы весов через w^L , w^T и w^X , можем записать выходные сигналы строк кроссбара (рис. 5):

$$\text{SumL} = (x, w^L), \quad \text{SumT} = (x, w^T), \quad \text{SumX} = (x, w^X),$$

где x — вектор входных сигналов, компоненты которого — напряжения 0,3 В (соответствует 1) и $-0,3$ В (соответствует -1). Модуль напряжений выбран ниже порога бинарного мемристора, чтобы исключить возможность изменения проводимости мемристоров кроссбара при его эксплуатации.

Второй слой сети Хемминга (рис. 6) реализован на основе КМОП-технологии [14, 15] и преобразует вектор входных сигналов по принципу WTA, то есть максимальной входной компоненте он ставит в соответствие 1 на выходе (высокое напряжение, для схемы рис. 6 оно составляет около 4,5 В), а остальным входным компонентам ставится в соответствие 0 (напряжение, близкое к нулю, для схемы рис. 6 оно составляет около 0,1 мВ). Для примера на рис. 5 выполняется $\text{SumT} > \text{SumL}$ и $\text{SumT} > \text{SumX}$. В этом случае на выходе второго слоя (рис. 6) получаем $\text{OutT} \approx 4,5$ В, $\text{OutL} = \text{OutX} \approx 0,1$ мВ. Схема памяти Хемминга (рис. 5 и 6) реализована в системе моделирования LTSPICE [16].

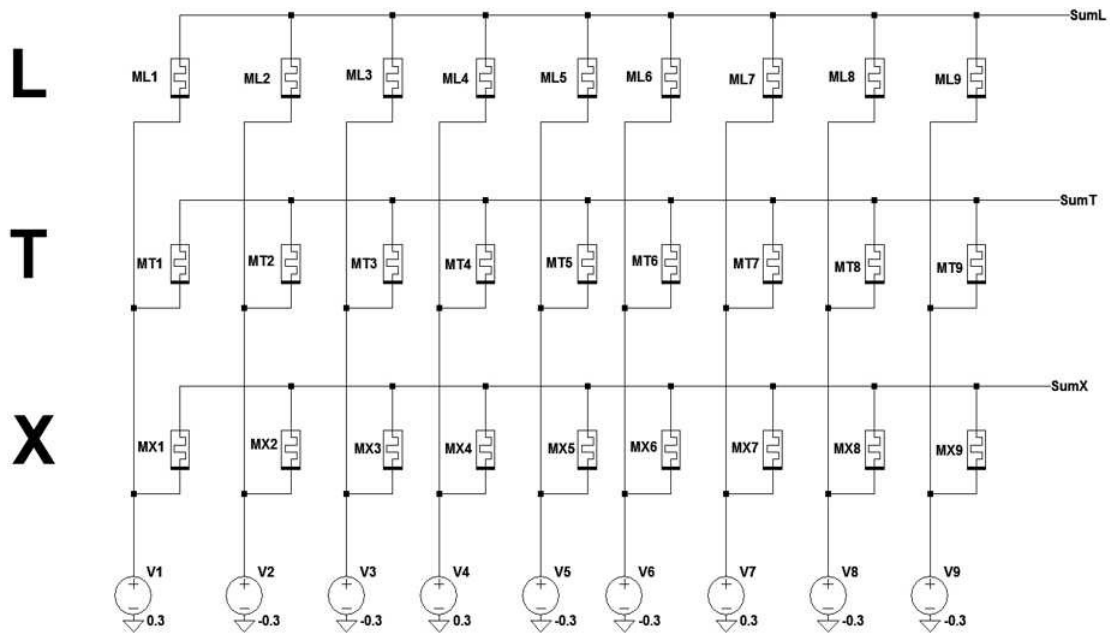
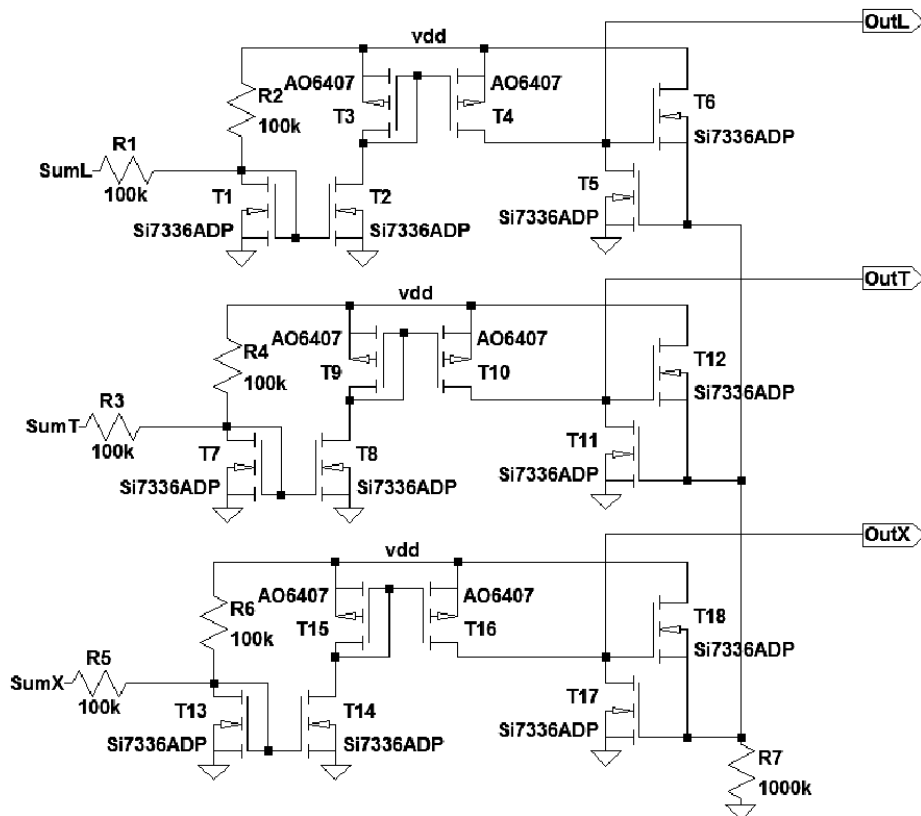


Рис. 5. Первый слой сети Хемминга


 Рис. 6. Второй слой сети Хемминга ($v_{dd} = 7\text{ V}$)

Заключение

Предложена аппаратная реализация ассоциативной памяти Хемминга на основе использования кроссбара на бинарных мемристорах и КМОП-схемотехники. Показано,

что кроссбар на бинарных мемристорах, где максимальное сопротивление мемристора соответствует значению -1 компоненты хранимого эталонного вектора, а минимальное — значению $+1$, реализует свойства первого слоя сети Хемминга, согласно которым выходной сигнал нейрона первого слоя неотрицателен, причём он максимален для нейрона, эталонный вектор которого наиболее близок к входному. Для заданной размерности эталонного вектора получено соотношение между максимальным и минимальным сопротивлениями бинарных мемристоров, которое гарантирует корректную работу первого слоя сети Хемминга. Моделирование в системе LTSPICE предложенной схемы памяти Хемминга подтвердило её работоспособность.

ЛИТЕРАТУРА

1. Chua L. Memristor — the missing circuit element // IEEE Trans. Circuit Theory. 1971. V. 18. P. 507–519.
2. Strukov D. B., Snider G. S., Stewart D. R., and Williams R. S. The missing memristor found // Nature. 2008. V. 453. P. 80–83.
3. <http://www.utmn.ru/presse/teleradiokanal-evrazion/videonovosti-tyumgu/89986/>
4. Pershin Y. and Di Ventra M. Experimental demonstration of associative memory with memristive neural networks // Neural Networks. 2010. V. 23. No. 7. P. 881–886.
5. Chua L. Resistance switching memories are memristors // Appl. Phys. A: Mater. Sci. & Process. 2011. V. 102. No. 4. P. 765–783.
6. Ho Y., Huang G. M., and Li P. Nonvolatile memristor memory: device characteristics and design applications // Proc. Int. Conf. ICCAD. 2009. P. 485–490.
7. Jo S. H., Chang T., Ebong I., et al. Nanoscale memristor device as synapse in neuromorphic systems // Nanoletters. 2010. V. 10. No. 4. P. 1297–1301.
8. Kavehei O. Memristive Devices and Circuits for Computing, Memory, and Neuromorphic Applications. PhD Thesis. The University of Adelaida, Australia, 2011.
9. Lehtonen E. Memristive Computing. Doctoral Thesis. University of Turku, Finland, 2012.
10. Lu W., Kim K.-H., Chang T., and Gaba S. Two-terminal resistive switches (memristors) for memory and logic applications // Proc. 16th Asia and South Pacific Design Automation Conf., Yokohama, Japan, 2011.
11. Truong S. N., Ham S.-J., and Min K.-S. Neuromorphic crossbar circuit with nanoscale filamentary-switching binary memristors for speech recognition // Nanoscale Res. Lett. 2014. No. 9:629. <http://www.nanoscalereslett.com/content/9/1/629>
12. Yakopcic C., Taha T. M., Subramanyam G., and Pino R. E. Memristor SPICE model and crossbar simulation based on devices with nanosecond switching time // Proc. Int. Joint Conf. Neural Networks. Dallas, Texas, USA. August 4–9, 2013.
13. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002. 344 с.
14. Zhu X., Yang X., Wu C., et al. Hamming network circuits based on CMOS/memristor hybrid design // IEICE Electronics Express. 2013. V. 10. No. 12. P. 1–9.
15. Lazzaro J., Ruckebusch S., Mahowald M. A., and Mead C. A. Winner-take-all networks of $O(n)$ complexity // Advances in Neural Information Processing Systems. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1989. P. 703–711.
16. Володин В. Я. Компьютерное моделирование электронных схем. СПб.: БХВ-Петербург, 2010. 400 с.

REFERENCES

1. Chua L. Memristor — the missing circuit element. IEEE Trans. Circuit Theory, 1971, vol. 18, pp. 507–519.
2. Strukov D. B., Snider G. S., Stewart D. R., and Williams R. S. The missing memristor found. Nature, 2008, vol. 453, pp. 80–83.
3. <http://www.utmn.ru/presse/teleradiokanal-evrazion/videonovosti-tyumgu/89986/>
4. Pershin Y. and Di Ventra M. Experimental demonstration of associative memory with memristive neural networks. Neural Networks, 2010, vol. 23, no. 7, pp. 881–886.
5. Chua L. Resistance switching memories are memristors. Appl. Phys. A: Mater. Sci. & Process., 2011, vol. 102, no. 4, pp. 765–783.
6. Ho Y., Huang G. M., and Li P. Nonvolatile memristor memory: device characteristics and design applications. Proc. Int. Conf. ICCAD, 2009, pp. 485–490.
7. Jo S. H., Chang T., Ebong I., et al. Nanoscale memristor device as synapse in neuromorphic systems. Nanoletters, 2010, vol. 10, no. 4, pp. 1297–1301.
8. Kavehei O. Memristive Devices and Circuits for Computing, Memory, and Neuromorphic Applications. PhD Thesis. The University of Adelaida, Australia, 2011.
9. Lehtonen E. Memristive Computing. Doctoral Thesis. University of Turku, Finland, 2012.
10. Lu W., Kim K.-H., Chang T., and Gaba S. Two-terminal resistive switches (memristors) for memory and logic applications. Proc. 16th Asia and South Pacific Design Automation Conf., Yokohama, Japan, 2011.
11. Truong S. N., Ham S.-J., and Min K.-S. Neuromorphic crossbar circuit with nanoscale filamentary-switching binary memristors for speech recognition. Nanoscale Res. Lett., 2014, no. 9:629. <http://www.nanoscalereslett.com/content/9/1/629>
12. Yakopcic C., Taha T. M., Subramanyam G., and Pino R. E. Memristor SPICE model and crossbar simulation based on devices with nanosecond switching time. Proc. Int. Joint Conf. Neural Networks, Dallas, Texas, USA, August 4–9, 2013.
13. Osovskiy S. Neironnye seti dlya obrabotki informatsii [Neural Networks for Data Processing]. Moscow, Finansy i Statistika, 2002. 344 p. (in Russian)
14. Zhu X., Yang X., Wu C., et al. Hamming network circuits based on CMOS/memristor hybrid design. IEICE Electronics Express, 2013, vol. 10, no. 12, pp. 1–9.
15. Lazzaro J., Ryckebusch S., Mahowald M. A., and Mead C. A. Winner-take-all networks of $O(n)$ complexity. Advances in Neural Information Processing Systems. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc., 1989, pp. 703–711.
16. Volodin V. Ya. Kompyuternoe modelirovanie elektronnykh skhem [Computer Simulation of Electronic Circuitry] SPb., BHV-Peterburg Publ., 2010. 400 p. (in Russian)