

МОДЕЛИРОВАНИЕ ПРОЦЕССОВ И МАТЕРИАЛОВ

SIMULATION OF PROCESSES AND MATERIALS

Известия высших учебных заведений. Материалы электронной техники. 2020. Т. 23, № 3. С. 186—195.
DOI: 10.17073/1609-3577-2020-3-186-195

УДК 621.315.61:(004.3+004.93)

Математическое моделирование самообучающейся нейроморфной сети, основанной на наноразмерных мемристивных элементах с 1T1R-кроссбар-архитектурой*

© 2020 г. А. Ю. Морозов^{1,2}, К. К. Абгарян^{1,2,§}, Д. Л. Ревизников^{1,2}

¹ *Федеральный исследовательский центр «Информатика и управление»
Российской академии наук,
ул. Вавилова, д. 44, корп. 2, Москва, 119333, Россия*

² *Московский авиационный институт (национальный исследовательский университет),
Волоколамское шоссе, д. 4, Москва, 125993, Россия*

Аннотация. Искусственные нейронные сети играют важную роль в современном мире. Основная их область применения – это задачи распознавания и обработки изображений, речи, а также робототехника и беспилотные системы. Использование нейронных сетей связано с большими вычислительными затратами. Отчасти именно этот факт сдерживал их прогресс, и только с появлением высокопроизводительных вычислительных систем началось активное развитие данной области. Тем не менее, вопрос ускорения работы нейросетевых алгоритмов все еще актуален. Одним из перспективных направлений является создание аналоговых реализаций искусственных нейронных сетей, так как аналоговые вычисления проводятся на порядки быстрее, чем цифровые. В качестве базового элемента, на котором строятся такие системы, выступает мемристор. Мемристор представляет собой резистор, проводимость которого зависит от суммарного пройденного через него заряда. Объединение мемристоров в матрицу (кроссбар) позволяет реализовать на аппаратном уровне один слой искусственных синапсов. Традиционно в качестве аналогового метода обучения применяется метод STDP, основанный на правиле Хебба. Выполнено моделирование двухслойной полносвязной сети с одним слоем синапсов. Мемристивный эффект может проявляться в разных веществах (в основном в разных оксидах), поэтому важно понимать, как характеристики мемристоров будут влиять на параметры нейронной сети. Рассмотрены два оксида: оксид титана (TiO_2) и оксид гафния (HfO_2). Для каждого оксида выполнена параметрическая идентификация соответствующей математической модели для наилучшего согласования с экспериментальными данными. Проведены настройка нейронной сети в зависимости от используемого оксида и моделирование процесса ее обучения распознаванию пяти шаблонов.

Ключевые слова: мемристор, оксид титана, оксид гафния, нейроморфная сеть, импульсная нейронная сеть, STDP, распознавание

Введение

Искусственные нейронные сети используются во многих областях современной жизни и позволяют решать актуальные, важные и практически значимые задачи, которые зачастую не поддаются решению с помощью классических подходов. Для ускорения работы нейросетевых алгоритмов ведутся разработки специальных процессоров, основанных на принципах действия человеческого мозга и представляющих собой аппаратную реализацию

* Статья подготовлена по материалам доклада, представленного на II-й международной конференции «Математическое моделирование в материаловедении электронных компонентов», Москва, 19–21 октября 2020 г.

Александр Юрьевич Морозов^{1,2} — канд. физ.-мат. наук, научный сотрудник, ORCID: 0000-0003-0364-8665, e-mail: morozov@infway.ru; **Карина Карленовна Абгарян**^{1,2,§} — доктор физ.-мат. наук, доцент, заведующая отделом, ORCID: 0000-0002-0059-0712, e-mail: kristal83@mail.ru; **Дмитрий Леонидович Ревизников**^{1,2} — доктор физ.-мат. наук, профессор, ORCID: 0000-0003-0998-7975, e-mail: reviznikov@mai.ru

§ Автор для переписки

импульсных (спайковых) нейронных сетей. Перспективным в данном направлении представляется использование аналоговых вычислений вместо цифровых, так как они производятся на порядки быстрее. В связи с этим актуальной задачей является создание аналоговых нейроморфных систем.

Мемристор — это резистор, проводимость которого меняется в зависимости от суммарного протекшего через него электрического заряда и который является элементарной ячейкой долгосрочной энергонезависимой памяти [1, 2]. Объединение мемристоров в матрицу (кроссбар) позволяет выполнять быстрое аналоговое произведение матрицы на вектор [3, 4]. За счет определенного сходства мемристивных элементов с биологическим синапсом перспективным представляется их использование для аналоговой реализации самообучающихся импульсных нейронных сетей.

Ранее авторами уже была исследована возможность применения мемристоров для аналоговой реализации сверточных нейронных сетей второго поколения [5]. В работе [5] получены оценки характеристик мемристивных элементов, при которых их можно было бы использовать для аппаратной реализации соответствующих нейросетевых алгоритмов.

В настоящей работе выполняется моделирование двухслойной полносвязной импульсной сети с одним слоем мемристивных элементов (синапсов). Используется 1T1R-кроссбар-архитектура, в которой каждому мемристорам соответствует один транзистор. Благодаря такой комбинации имеется возможность выполнять обучение сети на аппаратном уровне с помощью метода STDP (*Spike Timing Dependent Plasticity*) [6—11]. Мемристивный эффект может проявляться в разных веществах (в основном в разных оксидах), поэтому важно понимать, как характеристики мемристоров будут влиять на параметры нейронной сети.

Цель работы — моделирование процесса функционирования нейроморфной сети с мемристивными элементами в качестве синаптических весов, основанных на разных оксидах, а также исследование возможностей адаптации используемой модели нейроморфной сети на случай применения различных типов мемристоров.

Рассматриваются несколько существующих математических моделей мемристоров и выполняется сравнение их характеристик с экспериментальными данными по мемристорам на основе оксида титана (TiO_2) и оксида гафния (HfO_2) соответственно. В третьем разделе формулируется математическая модель схемотехнического решения, реализующего однослойную самообучающуюся импульсную нейронную сеть с мемристивными элементами в качестве синаптических весов. В четвертом разделе выполняется численное моделирование работы

пяти взаимосвязанных нейронов с 320 синапсами для двух разных оксидов. В заключении формулируются основные результаты работы.

Математические модели мемристора

Мемристивный эффект, как правило, возникает за счет перемещения ионов в сверхтонком диэлектрическом слое при приложении электрического поля. Применительно к различным оксидам часто говорят о перемещении вакансий кислорода и формировании/разрушении проводящих филаментов. Большинство известных моделей мемристора сформулированы в виде динамической системы относительно состояния мемристора. Параметр состояния мемристора — это величина, которая соответствует положению границы, разделяющей области с низкой и высокой концентрацией вакансий кислорода, толщине проводящего слоя, или толщине непроводящего барьера, в котором возникает туннельный ток электронов. В зависимости от закона изменения параметра состояния мемристора можно выделить несколько математических моделей — в частности, модели линейного [12] и нелинейного дрейфа [13], модель, основанная на барьере Симмонса [14]. Для ограничения переменной состояния вводятся специальные функции окна [15—18]. Экспериментальные данные показывают, что изменение состояния происходит не при любом значении напряжения, а начиная с некоторого порога, в связи с этим в модели добавляются пороговые условия [19—21].

Рассматриваются несколько моделей мемристоров. Первая модель является переменнорезисторной моделью тонкопленочного мемристора, основанной на экспоненциальной модели дрейфа легирующей примеси [22]:

$$\begin{aligned}
 R &= R_{\text{on}}x + R_{\text{off}}(1 - x), \\
 I_M &= \frac{V_M}{R}, \\
 \frac{dx}{dt} &= \begin{cases} \mu_v \frac{V_p}{D^2} \exp\left(\frac{R_{\text{on}}}{V_p} I_M\right), & V_M \geq V_p, \\ \mu_v \frac{V_n}{D^2} \exp\left(\frac{R_{\text{on}}}{V_n} I_M\right), & V_M \leq V_n, \\ \mu_v \frac{R_{\text{on}}}{D^2} I_M, & V_n < V_M < V_p, \end{cases} \quad (1)
 \end{aligned}$$

где $x \in [0, 1]$ — переменная состояния; $R_{\text{on}}, R_{\text{off}}$ — минимальное и максимальное сопротивление мемристора; I_M, V_M, R — текущее значение тока, напряжения и сопротивления мемристора; V_p, V_n — значения напряжений, при которых происходит переключение состояния; μ_v — коэффициент легирующей подвижности; D — толщина полупроводниковой пленки.

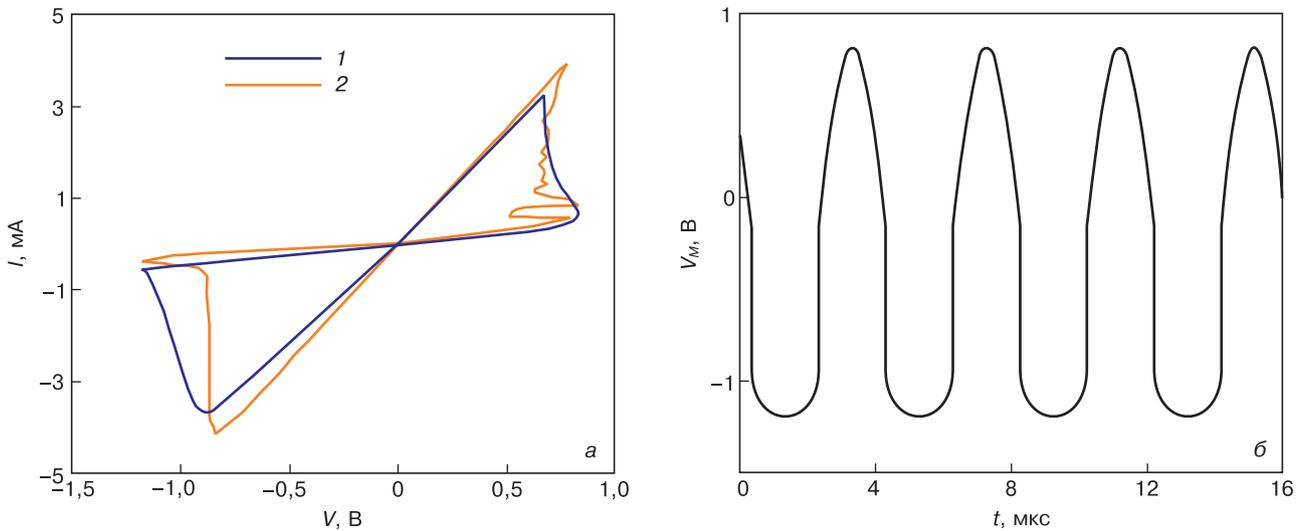


Рис. 1. Сравнение вольт–амперной характеристики модели (1) с экспериментальными данными (2) по оксиду титана (а) и форма входного напряжения (б) для этой модели

Fig. 1. Comparison of the C–V curve of model (1) with experimental data (2) on titanium oxide (a) and the shape of the input voltage (b) for this model

Выполняется моделирование работы мемристора при следующих значениях параметров: $R_{on} = 205 \text{ Ом}$, $R_{off} = 2,13 \text{ кОм}$, $\mu_v = 6 \cdot 10^{-10}$, $V_p = 0,65 \text{ В}$, $V_n = -0,87 \text{ В}$, $D = 620 \text{ нм}$, $x(0) = 0,1$, $t \in [0, 16] \text{ мс}$, $V_M(t)$ представлена на рис. 1, б. Такой выбор параметров и формы напряжения $V_M(t)$ обусловлен получением характеристик мемристора, схожих с экспериментальными характеристиками по оксиду титана, приведенными в работе [12]. На рис. 1, а показана экспериментальная вольтамперная характеристика и модельная.

Здесь наблюдается хорошее согласование в правой части графиков и удовлетворительное — в левой части.

Далее рассматривается модель с нелинейной зависимостью от напряжения. В общем виде уравнение, описывающее состояние мемристора, может быть представлено следующим образом:

$$\frac{dx}{dt} = af(x)V^s,$$

где $x \in [0, 1]$ — переменная состояния; a — постоянная, определяемая свойствами материала; V — текущее значение напряжения; s — нечетное целое число, $f(x)$ — функция окна, используемая для приблизительного представления нелинейных эффектов ионного дрейфа и ограничения границ.

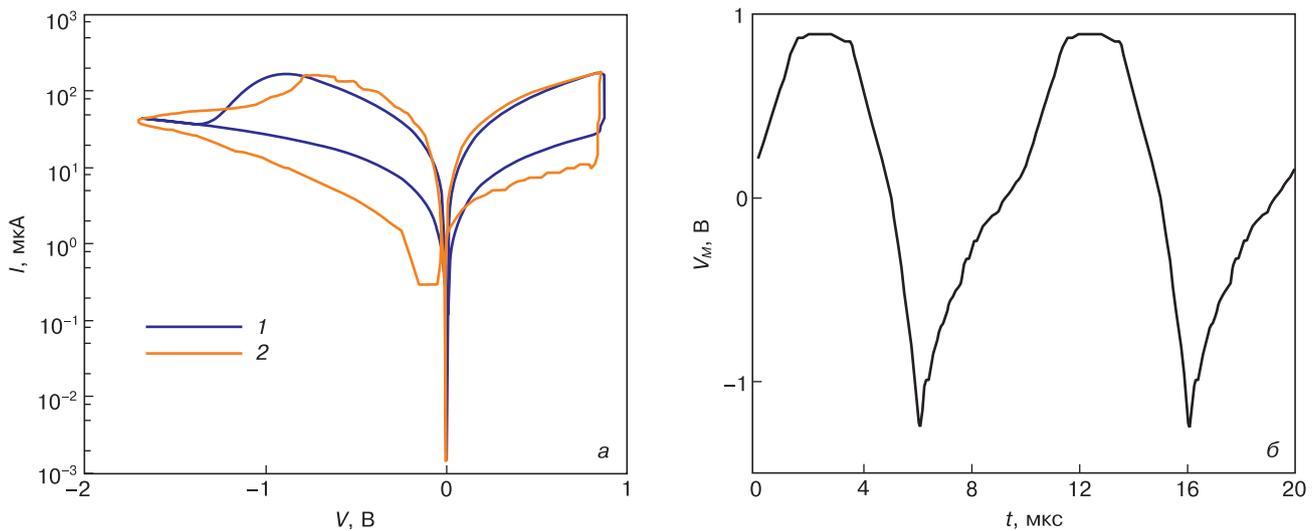


Рис. 2. Сравнение вольт–амперной характеристики модели (1) с экспериментальными данными (2) по оксиду гафния (а) и форма входного напряжения (б) для этой модели

Fig. 2. Comparison of the C–V curve of model (2) with experimental data (2) on hafnium oxide (a) and the shape of the input voltage (b) for this model

В настоящей работе используется модель мемристора данного класса, предложенная в [23]:

$$\frac{dx}{dt} = aV^s \begin{cases} 1 - (1-x)^{2\text{round}\left(\frac{b}{|V|+c}\right)}, & V \leq -v_{\text{thr}}, \\ 1 - x^{2\text{round}\left(\frac{b}{|V|+c}\right)}, & V > v_{\text{thr}}, \\ 0, & -v_{\text{thr}} < V \leq v_{\text{thr}}, \end{cases} \quad (2)$$

$$I = x^n \beta \sinh(\alpha_M V) + \chi [\exp(\gamma V) - 1],$$

где I, V — текущие значения тока и напряжения; v_{thr} — пороговое значение напряжения активации; $n, \beta, \alpha_M, \chi, \gamma$ — подгоночные параметры в выражении для тока; round — функция получения целочисленного результата; b, c — подгоночные коэффициенты основного уравнения.

Выполняется моделирование работы мемристора при следующих значениях параметров: $n = 5, \beta = 7,069 \cdot 10^{-5} \text{ В}, \alpha_M = 1,8 \text{ В}^{-1}, \chi = 1,946 \cdot 10^{-4} \text{ В}, \gamma = 0,15 \text{ В}^{-1}, a = 1 \text{ В}^{-5}, s = 5, b = 15 \text{ В}, c = 2 \text{ В}, v_{\text{thr}} = 1 \text{ В}, x(0) = 0,4, V(t)$ — рис. 2, б. Часть значений соответствует значениям в исходной работе [23], а часть подобрана для наибольшего соответствия экспериментальным данным по оксиду гафния (HfO_2) приведенным в той же статье [23].

На рис. 2, а приведено сравнение полученной вольтамперной характеристики с экспериментальной кривой для HfO_2 .

Здесь наблюдается удовлетворительное согласование результатов моделирования с экспериментальными данными.

Важной особенностью мемристорных элементов является возможное их несовершенство [24], ведущее к неконтролируемому изменению уровня проводимости в ходе функционирования системы или инициации. Отметим, что в этом случае модель мемристора можно описать как динамическую систему с неопределенностями и использовать для ее исследования соответствующие методы [25, 26].

Математическая модель нейроморфной сети

Рассмотрим работу схемотехнического решения однослойной самообучающейся аналоговой импульсной нейронной сети с мемристорными элементами в качестве синаптических весов (рис. 3). Пришедшие на вход импульсы V_g открывают соответствующие транзисторы, что приводит к протеканию токов через мемристоры с последующим их суммированием в нейронах. Схемотехническая модель нейрона представляет собой параллельную RC цепь и абстрактный генератор импульсов G (рис. 4). Как только значение потенциала на конденсаторе превышает некоторый порог, его потенциал сбрасывается, и генератор импульсов выдает выходной сигнал V_{out} и сигнал обратной связи V_{te} . Помимо этого, в обратной связи постоянно поддерживается некоторый небольшой потенциал, необходимый для функционирования сети в обычном режиме.

Процесс обучения сети происходит согласно правилу STDP (те синаптические связи, которые привели к активации нейрона, усиливаются, другие — ослабевают). Данный механизм обучения реализуется путем обратной связи у нейронов (V_{te}).

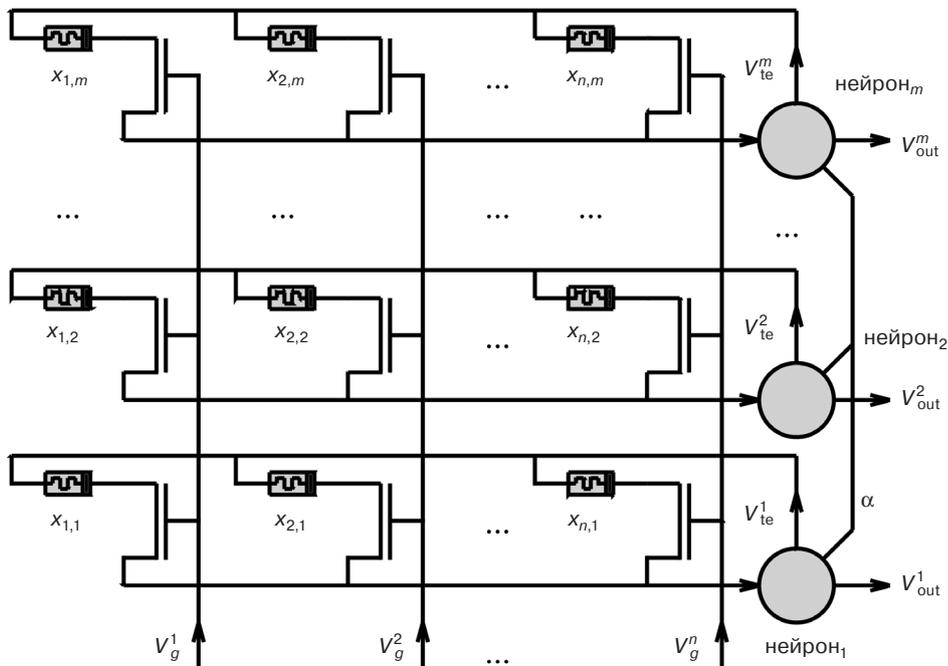


Рис. 3. Схемотехническая реализация импульсной нейронной сети
 Fig. 3. Schematic implementation of a pulsed neural network

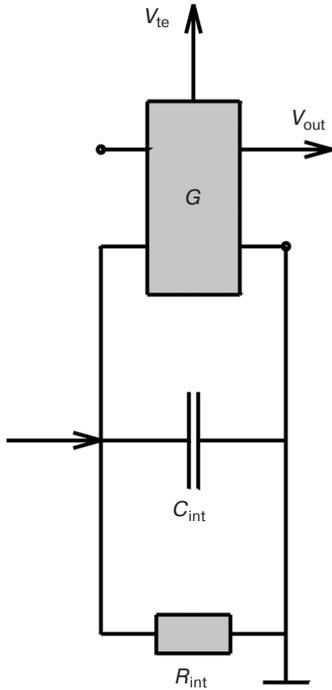


Рис. 4. Схематическая реализация нейрона
Fig. 4. Schematic implementation of a neuron

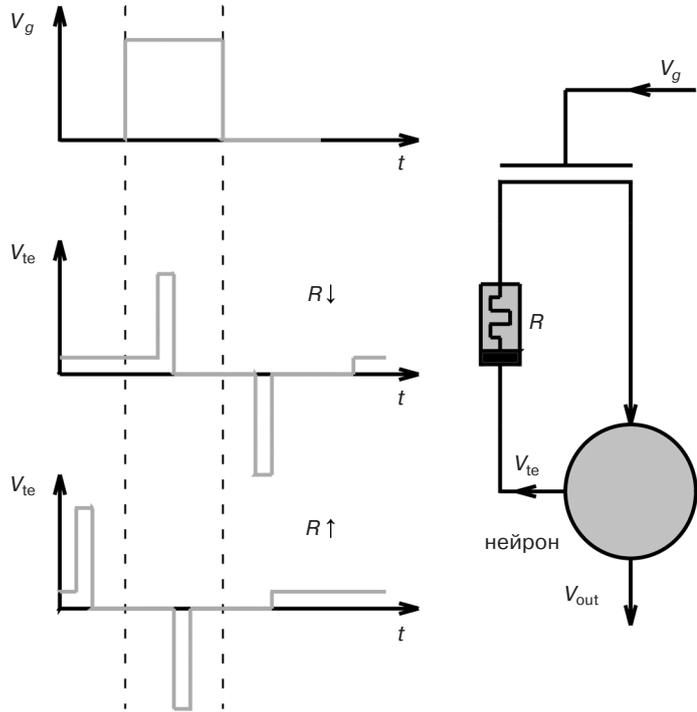


Рис. 5. Схематическая реализация правила обучения STDP
Fig. 5. Schematic implementation of the STDP learning rule

В момент активации нейрона по каналу обратной связи с задержками поступают два противоположных по знаку импульса. Если на синапсе есть активность и пришел положительный импульс обратной связи, то значение проводимости соответствующего мемристора увеличивается, а если пришел отрицательный импульс обратной связи, то проводимость мемристора уменьшается. На рис. 5 проиллюстрирован процесс изменения синаптических весов на примере одного синапса.

Обучение сети происходит следующим образом: на вход в сеть с равной вероятностью подается или произвольный шум, или заранее определенный шаблон. Спустя некоторое время сеть адаптируется к распознаванию шаблона. В случае нескольких шаблонов в выходном слое будет несколько нейронов с дополнительной связью (на рис. 3 помечена буквой α). Когда происходит активация одного нейрона, то он подавляет остальные нейроны (уменьшает значение их потенциала). Распределение шаблонов по нейронам происходит в процессе обучения.

Сформулируем комплексную математическую модель однослойной самообучающейся импульсной нейронной сети (см. рис. 3). Математическая модель задается следующими соотношениями [27]:

$$\frac{dx_{i,j}}{dt} = \begin{cases} F_X \left(\frac{V_{te}^j - V_{int}^j}{R_{i,j}}, V_{te}^j - V_{int}^j, x_{i,j} \right), & V_g^i(t) > 0, \\ 0, & V_g^i(t) = 0, \end{cases} \quad (3)$$

$$R_{i,j} = F_R(x_{i,j}, V_{te}^j - V_{int}^j),$$

$$\frac{dV_{int}^j}{dt} = \frac{1}{C_{int}} \left[\sum_{i=1}^n \frac{\hat{V}_{te}^j - V_{int}^j}{R_{i,j}} - \frac{V_{int}^j}{R_{int}} \right] - \max_{i=1,m} \left[\theta(V_{int}^i - V_{th}) \hat{\alpha}_{i,j} \right] \delta \left(\prod_{i=1}^m (V_{int}^i - V_{th}) \right) V_{int}^j, \quad (4)$$

$$\frac{d\tau_j}{dt} = 1 - \delta(V_{int}^j - V_{th}) \tau_j, \quad (5)$$

$$V_{te}^j = \begin{cases} V_{te}^+, & \tau_j \leq \tau_s, \\ V_{te}^-, & \frac{\tau_r}{2} < \tau_j \leq \frac{\tau_r}{2} + \tau_s, \\ V_{te}^0, & \tau_r < \tau_j, \\ 0, & \tau_s < \tau_j \leq \frac{\tau_r}{2} \vee \frac{\tau_r}{2} + \tau_s < \tau_j \leq \tau_r, \end{cases} \quad (6)$$

$$V_{out}^j = \begin{cases} V_{out}^+, & \tau_j \leq \tau_{out}, \\ 0, & \tau_{out} < \tau_j, \end{cases} \quad (7)$$

$$\hat{\alpha}_{i,j} = 1 - \alpha(1 - \delta_{ij}), \quad i = \overline{1, n}, j = \overline{1, m},$$

— и дополняется начальными условиями

$$x_{i,j}(0) = \text{random}[0, 1], V_{int}^j(0) = 0, \tau_j(0) > \max(\tau_r, \tau_{out}),$$

где n — количество входов; m — количество нейронов; V_g^i — текущее значение напряжения на i -м входе нейронной сети; V_{te}^j — текущее значение напряжения в обратной связи j -го нейрона; V_{out}^j — теку-

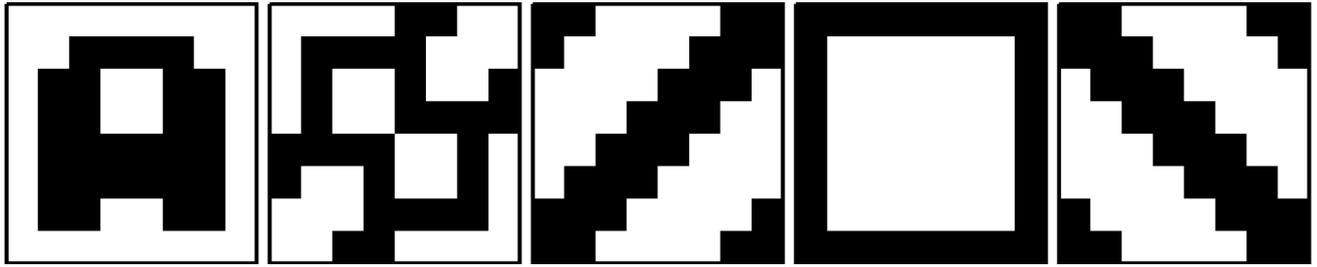


Рис. 6. Распознаваемые шаблоны
Fig. 6. Recognizable patterns

щее значение напряжения на выходе j -го нейрона; τ_j — время, прошедшее после последней активации j -го нейрона; V_{int}^j — напряжение на конденсаторе j -го нейрона; $R_{\text{int}}, C_{\text{int}}$ — значение сопротивления и емкости у нейронов; $V_{\text{te}}^+, V_{\text{te}}^-, V_{\text{te}}^0$ — значения амплитуды импульсов обратной связи и значение напряжения по умолчанию; V_{out}^+ — амплитуда выходного импульса; V_{th} — уровень напряжения активации нейрона; $R_{i,j}$ — значение сопротивления мемристора i -го синапса j -го нейрона; $x_{i,j}$ — состояние мемристора i -го синапса j -го нейрона, $x_{i,j} \in [0, 1]$; τ_r — длительность сигнала в обратной связи после активации нейрона; τ_s — длительность одного импульса в сигнале обратной связи, $2\tau_s < \tau_r$; τ_{out} — длительность одного импульса на выходе сети; α — коэффициент подавления; $\hat{V}_{\text{te}}^j = \max(0, \min(V_{\text{te}}^j, V_{\text{te}}^0))$ и отвечает за то, чтобы импульсы обратной связи не вносили вклад в накопление потенциала внутри нейрона; Δ_{ij} — символ Кронекера; $\Delta(x)$ — дельта-функция; $\theta(x)$ — функция Хэвисайда. Соотношения (3) задают модель мемристора. Функция $F_X(I, v, x)$ определяет скорость изменения переменной состояния в зависимости от тока (I), напряжения (v) и текущего состояния (x). Функция $F_R(x, v)$ определяет зависимость сопротивления мемристора от состояния и приложенного напряжения. Соотношение (4) задает модель нейрона, который представляет собой параллельную RC цепочку (рис. 4), включенную последовательно с резистором (все мемристоры синапсов на уровне нейрона можно рассматривать как один резистор). Уравнение (5) реализует механизм счетчика времени после последней активации нейрона. Как только напряжение V_{int}^j на конденсаторе достигает порогового значения V_{th} , переменная τ_j обнуляется. Аналогично происходит в уравнении (4): после активации нейрона накопленный им потенциал сбрасывается, а у других нейронов — уменьшается прямо пропорционально коэффициенту α . Уравнения (6) и (7) определяют форму импульсов в обратной связи и на выходе нейронной сети. В начальный момент времени переменная τ_j выбирается таким образом, чтобы избежать преждевременного появления импульсов в обратной связи и на выходе.

Отметим, что приведенная математическая модель описывает только один слой нейронной сети. Для моделирования многослойной сети достаточно

соединить выход k -го слоя со входом $(k + 1)$ -го слоя: $V_g^{i(k+1)}(t) = V_{\text{out}}^{j(k)}(t)$, $i = j$. Таким образом, область применения описанной математической модели не ограничивается однослойной сетью.

Результаты моделирования

Рассматривается задача распознавания пяти шаблонов (рис. 6) $m = 5$ и $n = 8 \times 8 = 64$. В процессе моделирования работы нейронной сети, каждую эпоху обучения (равную $\tau_r/2$ с), компоненты вектора $V_g(t)$ могут с равной вероятностью или быть случайным шумом (V_g^j имеет дискретное распределение), или принимать с равной вероятностью одно из пяти значений, которые задаются в соответствии с распознаваемыми шаблонами. Запишем вектор V_g (для наглядности в виде матрицы) для первых двух шаблонов:

$$V_g^{(1)} = \begin{pmatrix} 0, 0, 0, 0, 0, 0, 0, 0, \\ 0, 0, 2, 2, 2, 2, 0, 0, \\ 0, 2, 2, 0, 0, 2, 2, 0, \\ 0, 2, 2, 0, 0, 2, 2, 0, \\ 0, 2, 2, 2, 2, 2, 2, 0, \\ 0, 2, 2, 2, 2, 2, 2, 0, \\ 0, 2, 2, 0, 0, 2, 2, 0, \\ 0, 0, 0, 0, 0, 0, 0, 0 \end{pmatrix}^T \text{ В,}$$

$$V_g^{(2)} = \begin{pmatrix} 0, 0, 0, 0, 2, 2, 0, 0, \\ 0, 2, 2, 2, 2, 0, 0, 0, \\ 0, 2, 0, 0, 2, 0, 0, 2, \\ 0, 2, 0, 0, 2, 2, 2, 2, \\ 2, 2, 2, 2, 0, 0, 2, 0, \\ 2, 0, 0, 2, 0, 0, 2, 0, \\ 0, 0, 0, 2, 2, 2, 2, 0, \\ 0, 0, 2, 2, 0, 0, 0, 0 \end{pmatrix}^T \text{ В.}$$

В зависимости от используемой модели мемристора выполняется настройка параметров математической модели (3)—(7). Для модели (1), которая соответствует мемристорам на основе оксида титана (TiO_2), имеем следующие значения параметров: $R_{\text{int}} = 200 \text{ Ом}$, $C_{\text{int}} = 45 \text{ мкФ}$, $V_{\text{te}}^+ = 0,7 \text{ В}$, $V_{\text{te}}^- = -0,9 \text{ В}$, $V_{\text{te}}^0 = 10 \text{ мВ}$, $V_{\text{out}}^+ = 2 \text{ В}$, $V_{\text{th}} = 9 \text{ мВ}$, $\tau_r = 3 \text{ мс}$,

$\tau_s = 50$ мкс, $\tau_{out} = 1,5$ мс. Значения V_{te}^+ и V_{te}^- подбираются из расчета, чтобы инициировать переключение мемристора, и чем они больше по абсолютному значению, тем переключение мемристора происходит быстрее. Длительность импульсов τ_s в обратной связи также влияет на скорость переключения мемристора: чем больше длительность, тем быстрее произойдет переключение. Значение порога активации нейрона V_{th} не должно превышать напряжение по умолчанию в обратной связи V_{te}^0 , иначе нейрон не активируется. В то же время напряжение по умолчанию V_{te}^0 не должно быть слишком большим, чтобы не вызывать переходные процессы в мемристорах. Значения R_{int} и C_{int} подбираются в соответствии с диапазоном изменения сопротивления мемристивных элементов и отвечают за быстроту накопления заряда в нейроне, поэтому дополнительно должны быть согласованы с длительностью одной эпохи. Параметры τ_{out} и V_{out}^+ в случае однослойной сети не оказывают никакого влияния.

Напряжение V_g^i имеет дискретное распределение: $V_g^i = 0$ В с вероятностью 0,73 и $V_g^i = 2$ В с вероятностью 0,27. На рис. 7 показан процесс адаптации синаптических весов к распознаваемым образцам. Цвет соответствует значению переменной состоя-

ния соответствующего мемристора: чем темнее, тем проводимость больше; чем светлее, тем меньше. В начальный момент времени все веса инициализируются случайными значениями, и в процессе работы сети постепенно изменяются. Примерно с 700-й эпохи начинают просматриваться шаблоны, распознаванию которых обучается сеть: происходит запоминание информации нейронной сетью.

Далее рассматривается модель мемристора (2), которая соответствует оксиду гафния (HfO_2). Параметры математической модели нейронной сети следующие: $R_{int} = 1$ кОм, $C_{int} = 45$ мкФ, $V_{te}^+ = 1,55$ В, $V_{te}^- = -1,6$ В, $V_{te}^0 = 10$ мВ, $V_{out}^+ = 2$ В, $V_{th} = 2,5$ мВ, $\tau_r = 15$ мс, $\tau_s = 1$ мс, $\tau_{out} = 7,5$ мс. Настройка параметров осуществлялась по аналогии с моделью мемристора на основе оксида титана: $V_g^i = 0$ В с вероятностью 0,85 и $V_g^i = 2$ В с вероятностью 0,15. На рис. 8 показано изменение синаптических весов сети в процессе обучения.

В отличие от предыдущего примера здесь сеть обучалась дольше. Отметим, что приведенные наборы параметров нейронной сети являются одним из возможных наборов и не являются оптимальными с позиции скорости обучения сети. В процессе обучения шаблоны распределяются по нейронам про-

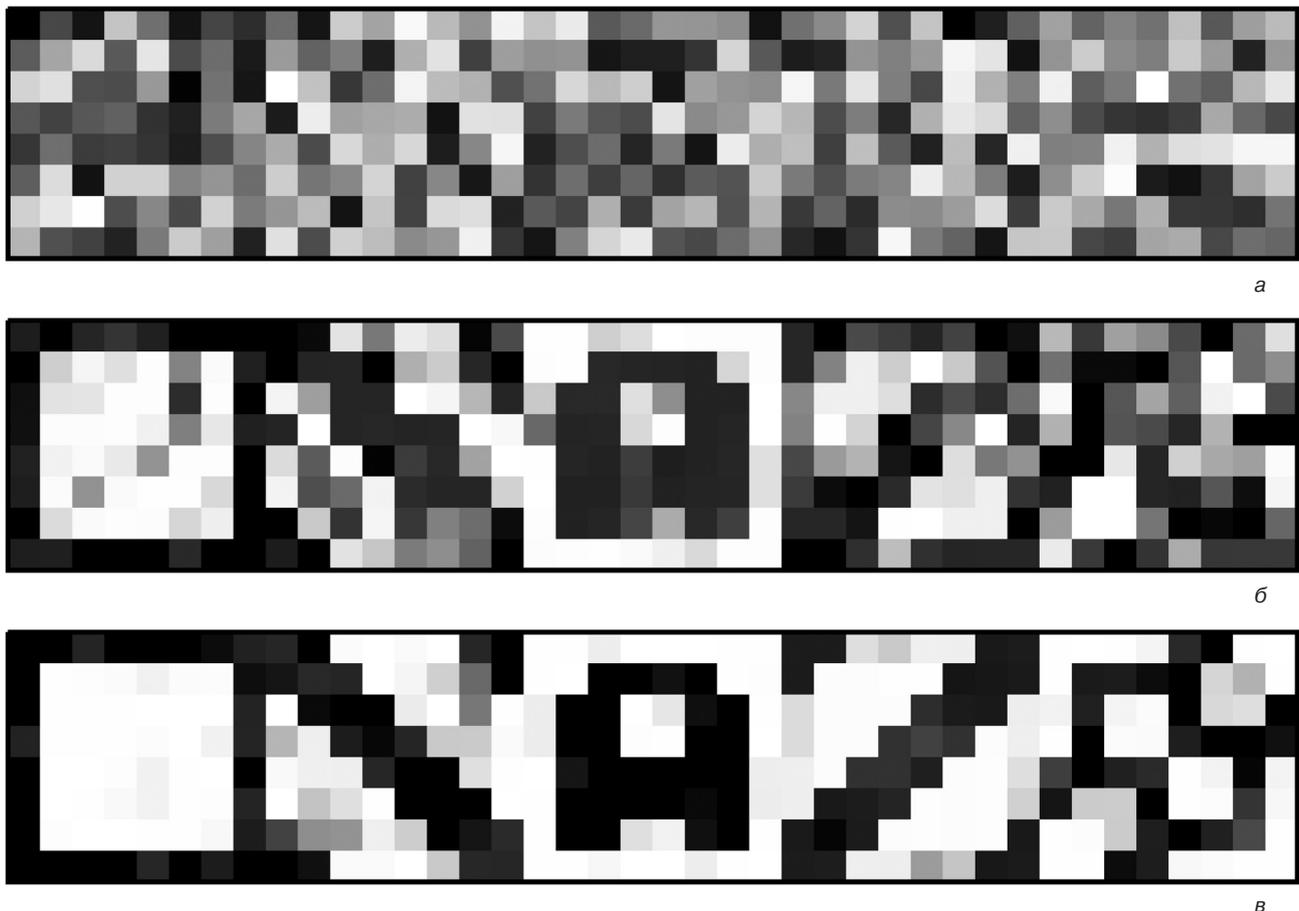


Рис. 7. Изменение синаптических весов в процессе обучения нейроморфной сети на основе оксида титана: а — 0; б — 700 эпох; в — 1500 эпох

Fig. 7. Change in synaptic weights during training of a neuromorphic network based on titanium oxide: (a) 0; (б) 700 epochs; (в) 1500 eras

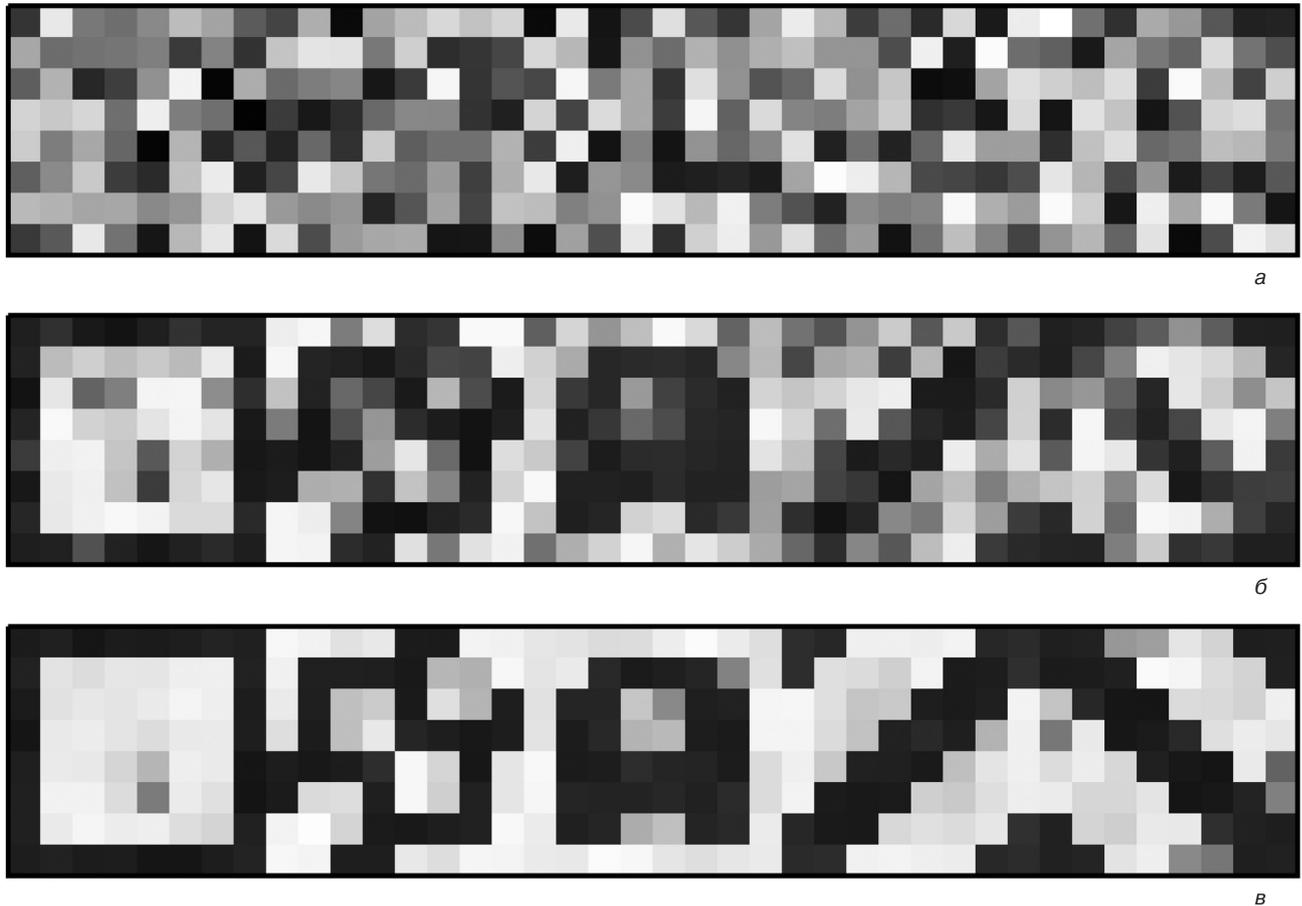


Рис. 8. Изменение синаптических весов в процессе обучения нейроморфной сети на основе оксида гафния: а — 0; б — 3000 эпох; в — 6000 эпох

Fig. 8. Change of synaptic weights in the process of training a neuromorphic network based on hafnium oxide: (a) 0; (б) 3000 epochs; (в) 6000 epochs

извольным образом, поэтому в двух приведенных примерах нейроны, отвечающие за один и тот же шаблон, разные.

Заключение

Работа посвящена математическому моделированию самообучающейся нейроморфной сети, основанной на наноразмерных мемристивных элементах с 1T1R-кроссбар-архитектурой. Рассмотрены несколько математических моделей, описывающих мемристоры на основе оксида титана и оксида гафния. Выполнено сравнение характеристик моделей с экспериментальными данными. Сформулирована комплексная математическая модель импульсной нейроморфной сети с механизмом обучения согласно правилу STDP. Выполнено моделирование работы двух нейронных сетей с мемристивными элементами в качестве синаптических весов на основе разных оксидов, состоящих из пяти нейронов с 320 синапсами. Произведена настройка параметров нейроморфной сети в зависимости от используемого оксида, что говорит об универсальности и гибкости используемой математической модели. В процессе

работы нейронные сети успешно обучились распознавать определенные шаблоны.

Библиографический список

1. Wong H.-S. P., Lee H. Y., Yu S., Chen Y. S., Wu Y., Chen P. S., Lee B., Frederic T. Metal-oxide RRAM // Proceedings of the IEEE. 2012. V. 100, N 6. P. 1951—1970. DOI: 10.1109/JPROC.2012.2190369
2. Yang J. J., Strukov D. B., Stewart D. R. Memristive devices for computing // Nature Nanotechnology. 2013. V. 8, N 1. P. 13—24. DOI: 10.1038/nnano.2012.240
3. Li C., Hu M., Li Y., Jiang H., Ge N., Montgomery E., Zhang J., Song W., Dávila N., Graves C. E., Li Z., Strachan J. P., Lin P., Wang Z., Barnell M., Wu Q., Williams R. S., Yang J. J., Xia Q. Analogue signal and image processing with large memristor crossbars // Nature Electronics. 2018. V. 1, N 1. P. 52—59. DOI: 10.1038/s41928-017-0002-z
4. Hu M., Graves C. E., Li C., Li Y., Ge N., Montgomery E., Davila N., Jiang H., Williams R. S., Yang J. J., Xia O., Strachan J. P. Memristor-based analog computation and neural network classification with a dot product engine // Advanced Materials. 2018. V. 30, N 9, P. 1705914. DOI: 10.1002/adma.201705914
5. Морозов А. Ю., Ревизников Д. Л., Абгарян К. К. Вопросы реализации нейросетевых алгоритмов на мемристивных кроссбарах // Известия вузов. Материалы электронной техники. 2019. Т. 22, № 4. С. 272—278. DOI: 10.17073/1609-3577-2019-4-272-278
6. Diehl P., Cook M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity // Frontiers in Computational Neuroscience. 2015. V. 9. P. 99. DOI: 10.3389/fncom.2015.00099
7. Ambrogio S., Milo V., Wang Z.-Q., Ramaswamy N., Balatry S., Carboni R., Calderoni A., Lelmbi D. Neuromorphic learning and recognition with one-transistor-one-resistor synapses and bist-

able metal oxide RRAM // IEEE Transactions on Electron Devices. 2016. V. 63, N 4. P. 1508—1515. DOI: 10.1109/TED.2016.2526647

8. Guo Y., Wu H., Gao B., Qian H. Unsupervised learning on resistive memory array based spiking neural networks // Frontiers in Neuroscience. 2019. V. 13. Art. N 812. DOI: 10.3389/fnins.2019.00812

9. Milo V., Laudato M., Ambrosi E., Chicca E., Pedretti G., Briccalli A., Bianchi S., Ielmini D. Resistive switching synapses for unsupervised learning in feed-forward and recurrent neural networks // International Symposium on Circuits and Systems. Florence (Italy): IEEE, 2018. P. 1—5. DOI: 10.1109/ISCAS.2018.8351824

10. Pedretti G., Bianchi S., Milo V., Calderoni A., Ramaswamy N., Ielmini D. Modeling-based design of brain-inspired spiking neural networks with RRAM learning synapses // International Electron Devices Meeting. San Francisco (CA, USA): IEEE, 2017. P. 28.1.1—28.1.4. DOI: 10.1109/IEDM.2017.8268467

11. Milo V., Ielmini D., Chicca E. Attractor networks and associative memories with STDP learning in RRAM synapses // IEEE International Electron Devices Meeting. San Francisco (CA, USA): IEEE, 2017. P. 11.2.1—11.2.4. DOI: 10.1109/IEDM.2017.8268369

12. Strukov D. B., Snider G. S., Stewart D. R., Williams R. S. The missing memristor found // Nature. 2008. V. 453, N 7191. P. 80—83. DOI: 10.1038/nature06932

13. Yang J. J., Pickett M. D., Xuema L., Ohlberg D. A. A., Stewart D. R., Williams R. S. Memristive switching mechanism for metal/oxide/metal nanodevices // Nature nanotechnology. 2008. V. 3, N 7. P. 429—433. DOI: 10.1038/nnano.2008.160

14. Pickett M. D., Strukov D. B., Borghetti J. L., Yang J. J., Snider G. S., Stewart D. R., Williams R. S. Switching dynamics in titanium dioxide memristive devices // J. Appl. Phys. 2009. V. 106, N 7. Art. N 074508. DOI: 10.1063/1.3236506

15. Joglekar Y. N., Wolf S. J. The elusive memristor: properties of basic electrical circuits // European Journal of Physics. 2009. V. 30, N 4. P. 661. DOI: 10.1088/0143-0807/30/4/001

16. Bielek Z., Bielek D., Biolkova V. SPICE model of memristor with nonlinear dopant drift // Radioengineering. 2009. V. 18, N 2. P. 210—214. URL: https://www.radioeng.cz/fulltexts/2009/09_02_210_214.pdf

17. Prodromakis T., Peh B. P., Papavassiliou C., Toumazou C. A versatile memristor model with nonlinear dopant kinetics // IEEE Transactions on Electron Devices. 2011. V. 58, N 9. P. 3099—3105. DOI: 10.1109/TED.2011.2158004

18. Zha J., Huang H., Liu Y. A novel window function for memristor model with application in programming analog circuits // IEEE Transactions on Circuits and Systems II: Express Briefs. 2015. V. 63, N 5. P. 423—427. DOI: 10.1109/TCSII.2015.2505959

19. Kvatinsky S., Ramadan M., Friedman E. G., Kolodny A., Weiser U. C. TEAM: ThrEshold adaptive memristor model // IEEE Transactions on Circuits and Systems I: Regular Papers. 2013. V. 60, N 1. P. 211—221. DOI: 10.1109/TCSI.2012.2215714

20. Kvatinsky S., Ramadan M., Friedman E. G., Kolodny A. VTEAM: A general model for voltage-controlled memristors // IEEE Transactions on Circuits and Systems II: Express Briefs. 2015. V. 62, N 8. P. 786—790. DOI: 10.1109/TCSII.2015.2433536

21. Yakopcic C., Taha T. M., Subramanyam G., Pino R. E., Rogers S. A memristor device model // IEEE Electron Device Letters. 2011. V. 32, N 10. P. 1436—1438. DOI: 10.1109/LED.2011.2163292

22. Zheng G., Mohanty S. P., Kougianos E., Okobiah O. Polynomial metamodel integrated Verilog-AMS for memristor-based mixed-signal system design // 56th International Midwest Symposium on Circuits and Systems (MWSCAS). Columbus (OH, USA): IEEE, 2013. P. 916—919. DOI: 10.1109/MWSCAS.2013.6674799

23. Mladenov V. Analysis of memory matrices with HfO₂ memristors in a PSpice environment // Electronics. 2019. V. 8, N 4. P. 383 (16pp). DOI: 10.3390/electronics8040383

24. Teplov G. S., Gornev E. S. Multilevel bipolar memristor model considering deviations of switching parameters in the Verilog-A language // Russ. Microelectron. 2019. V. 48, N 3. P. 131—142. DOI: 10.1134/S1063739719030107

25. Morozov A. Y., Reviznikov D. L. Adaptive interpolation algorithm based on a kd-tree for numerical integration of systems of ordinary differential equations with interval initial conditions // Differential Equations. 2018. V. 54, N 7. P. 945—956. DOI: 10.1134/S0012266118070121

26. Morozov A. Yu., Reviznikov D. L., Gidasov V. Yu. Adaptive interpolation algorithm based on a kd-tree for the problems of chemical kinetics with interval parameters // Mathematical Models and Computer Simulations. 2019. V. 11, N 4. P. 622—633. DOI: 10.1134/S2070048219040100

27. Morozov A. Y., Abgaryan K. K., Reviznikov D. L. Mathematical model of a neuromorphic network based on memristive elements // Chaos, Solitons & Fractals. 2021. V. 143, Art. N 110548. DOI: 10.1016/j.chaos.2020.110548

Работа выполнена при поддержке гранта РФФИ № 19-29-03051 мк.

Статья поступила в редакцию 15 августа 2020 г.

Izvestiya vuzov. Materialy elektronnoi tekhniki = Materials of Electronics Engineering. 2020, vol. 23, no. 3, pp. 186—195. DOI: 10.17073/1609-3577-2020-3-186-195

Mathematical modeling of a self-learning neuromorphic network based on nanosized memristive elements with 1T1R crossbar architecture

A. Yu. Morozov^{1,2}, K. K. Abgaryan^{1,2,§}, D. L. Reviznikov^{1,2}

¹ *Federal Research Centre “Information and Control” of the Russian Academy of Sciences, 44 Vavilov Str., Moscow 119333, Russia*

² *Moscow Aviation Institute (National Research University), 4 Volokolamskoe shosse, 4, Moscow 125993, Russia*

Abstract. Artificial neural networks play an important role in the modern world. Their main field of application is the tasks of recognition and processing of images, speech, as well as robotics and unmanned systems. The use of neural networks is associated with high computational costs. In part, it was this fact that held back their progress, and only with the advent of high-performance computing systems did the active development of this area begin. Nevertheless, the issue of speeding up the work of neural network algorithms is still relevant. One of the promising directions is the creation of analog implemen-

Information about authors:

Alexander Yu. Morozov^{1,2}: Cand. Sci. (Phys.–Math.), Researcher, ORCID: 0000-0003-0364-8665 (morozov@infway.ru); **Karine K. Abgaryan**^{1,2,§}: Dr. Sci. (Phys.–Math.), Associate Professor, Head of Department, ORCID: 0000-0002-0059-0712 (kristal83@mail.ru); **Dmitry L. Reviznikov**^{1,2}: Dr. Sci. (Phys.–Math.), Professor, ORCID: 0000-0003-0998-7975 (reviznikov@mai.ru)

§ Corresponding author

tations of artificial neural networks, since analog calculations are performed orders of magnitude faster than digital ones. The memristor acts as the basic element on which such systems are built. A memristor is a resistance, the conductivity of which depends on the total charge passed through it. Combining them into a matrix (crossbar) allows one layer of artificial synapses to be implemented at the hardware level. Traditionally, the STDP method based on Hebb's rule has been used as an analog learning method. In this work, we are modeling a two-layer fully connected network with one layer of synapses. The memristive effect can manifest itself in different substances (mainly in different oxides), so it is important to understand how the characteristics of memristors will affect the parameters of the neural network. Two oxides are considered: titanium oxide (TiO₂) and hafnium oxide (HfO₂). For each oxide, a parametric identification of the corresponding mathematical model is performed to best fit the experimental data. The neural network is tuned depending on the oxide used and the process of training it to recognize five patterns is simulated.

Keywords: memristor, titanium oxide, hafnium oxide, neuromorphic network, impulse neural network, STDP, recognition

References

1. Wong H.-S. P., Lee H. Y., Yu S., Chen Y. S., Wu Y., Chen P. S., Lee B., Frederic T. Metal-oxide RRAM. *Proceedings of the IEEE*, 2012. vol. 100, no. 6, pp. 1951—1970. DOI: 10.1109/JPROC.2012.2190369
2. Yang J. J., Strukov D. B., Stewart D. R. Memristive devices for computing. *Nature Nanotechnology*, 2013, vol. 8, no. 1, pp. 13—24. DOI: 10.1038/nnano.2012.240
3. Li C., Hu M., Li Y., Jiang H., Ge N., Montgomery E., Zhang J., Song W., Dávila N., Graves C. E., Li Z., Strachan J. P., Lin P., Wang Z., Barnell M., Wu Q., Williams R. S., Yang J. J., Xia Q. Analogue signal and image processing with large memristor crossbars. *Nature Electronics*, 2018, vol. 1, no. 1, pp. 52—59. DOI:10.1038/s41928-017-0002-z
4. Hu M., Graves C. E., Li C., Li Y., Ge N., Montgomery E., Davila N., Jiang H., Williams R. S., Yang J. J., Xia O., Strachan J. P. Memristor-based analog computation and neural network classification with a dot product engine. *Advanced Materials*, 2018, vol. 30, no. 9, pp. 1705914. DOI: 10.1002/adma.201705914
5. Morozov A. Yu., Reviznikov D. L., Abgaryan K. K. Issues of implementing neural network algorithms on memristor crossbars. *Izvestiya Vysshikh Uchebnykh Zavedenii. Materialy Elektronnoy Tekhniki = Materials of Electronics Engineering*, 2019, vol. 22, no. 4, pp. 272—278. DOI: 10.17073/1609-3577-2019-4-272-278
6. Diehl P., Cook M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in Computational Neuroscience*, 2015. vol. 9, pp. 99. DOI: 10.3389/fncom.2015.00099
7. Ambrogio S., Milo V., Wang Z.-Q., Ramaswamy N., Balatny S., Carboni R., Calderoni A., Ielmini D. Neuromorphic learning and recognition with one-transistor-one-resistor synapses and bistable metal oxide RRAM. *IEEE Transactions on Electron Devices*, 2016, vol. 63, no. 4, pp. 1508—1515. DOI: 10.1109/TED.2016.2526647
8. Guo Y., Wu H., Gao B., Qian H. Unsupervised Learning on Resistive Memory Array Based Spiking Neural Networks. *Frontiers in Neuroscience*, 2019, vol. 13, p. 812. DOI: 10.3389/fnins.2019.00812
9. Milo V., Laudato M., Ambrosi E., Chicca E., Pedretti G., Briccalli A., Bianchi S., Ielmini D. Resistive switching synapses for unsupervised learning in feed-forward and recurrent neural networks. *International Symposium on Circuits and Systems*. Florence (Italy): IEEE, 2018, pp. 1—5. DOI: 10.1109/ISCAS.2018.8351824
10. Pedretti G., Bianchi S., Milo V., Calderoni A., Ramaswamy N., Ielmini D. Modeling-based design of brain-inspired spiking neural networks with RRAM learning synapses. *International Electron Devices Meeting*, San Francisco (CA, USA): IEEE, 2017, pp. 28.1.1—28.1.4. DOI: 10.1109/IEDM.2017.8268467
11. Milo V., Ielmini D., Chicca E. Attractor networks and associative memories with STDP learning in RRAM synapses. *International Electron Devices Meeting*, San Francisco (CA, USA): IEEE, 2017, pp. 11.2.1—11.2.4. DOI: 10.1109/IEDM.2017.8268369
12. Strukov D. B., Snider G. S., Stewart D. R., Williams R. S. The missing memristor found. *Nature*. 2008, vol. 453, no. 7191, pp. 80. DOI: 10.1038/nature06932
13. Yang J. J., Pickett M. D., Xuema L., Ohlberg D. A. A., Stewart D. R., Williams R. S. Memristive switching mechanism for metal/oxide/metal nanodevices. *Nature nanotechnology*, 2008, vol. 3, no. 7, pp. 429—433. DOI: 10.1038/nnano.2008.160
14. Pickett M. D., Strukov D. B., Borghetti J. L., Yang J. J., Snider G. S., Stewart D. R., Williams R. S. Switching dynamics in titanium dioxide memristive devices. *J. Appl. Phys.*, 2009, vol. 106, no. 7, pp. 074508. DOI: 10.1063/1.3236506
15. Joglekar Y. N., Wolf S. J. The elusive memristor: properties of basic electrical circuits. *European J. Physics*, 2009, vol. 30, no. 4, pp. 661. DOI: 10.1088/0143-0807/30/4/001
16. Birolek Z., Birolek D., Biolkova V. SPICE model of memristor with nonlinear dopant drift. *Radioengineering*, 2009, vol. 18, no. 2, pp. 210—214. URL: https://www.radioeng.cz/full-texts/2009/09_02_210_214.pdf
17. Prodromakis T., Peh B. P., Papavassiliou C., Toumazou C. A versatile memristor model with nonlinear dopant kinetics. *IEEE Transactions on Electron Devices*, 2011. vol. 58, no. 9, pp. 3099—3105. DOI: 10.1109/TED.2011.2158004
18. Zha J., Huang H., Liu Y. A novel window function for memristor model with application in programming analog circuits. *IEEE Transactions on Circuits and Systems II: Express Briefs*. 2015. vol. 63, no. 5, pp. 423—427. DOI: 10.1109/TCSII.2015.2505959
19. Kvatinisky S., Friedman E. G., Kolodny A., Weiser U. C. TEAM: ThrEshold adaptive memristor model. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2013, vol. 60, no. 1, pp. 211—221. DOI: 10.1109/TCSI.2012.2215714
20. Kvatinisky S., Ramadan M., Friedman E. G., Kolodny A. VTEAM: A general model for voltage-controlled memristors. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2015, vol. 62, no. 8, pp. 786—790. DOI: 10.1109/TCSII.2015.2433536
21. Yakopcic C., Taha T. M., Subramanyam G., Pino R. E., Rogers S. A memristor device model. *IEEE Electron Device Letters*, 2011, vol. 32, no. 10, pp. 1436—1438. DOI: 10.1109/LED.2011.2163292
22. Zheng G., Mohanty S. P., Kougiyanos E., Okobiah O. Polynomial metamodel integrated Verilog-AMS for memristor-based mixed-signal system design. *International Midwest Symposium on Circuits and Systems (MWSCAS)*, Columbus (OH, USA): IEEE, 2013, pp. 916—919. DOI: 10.1109/MWSCAS.2013.6674799
23. Mladenov V. Analysis of memory matrices with HfO₂ memristors in a PSpice environment. *Electronics*, 2019, vol. 8, no. 4, p. 383. DOI: 10.3390/electronics8040383
24. Teplov G. S., Gornev E. S. Multilevel bipolar memristor model considering deviations of switching parameters in the Verilog-A language. *Russian Microelectronics*, 2019. vol. 48, no. 3, pp. 131—142. DOI: 10.1134/S1063739719030107
25. Morozov A. Y., Reviznikov D. L. Adaptive interpolation algorithm based on a kd-tree for numerical integration of systems of ordinary differential equations with interval initial conditions. *Differential Equations*, 2018, vol. 54, no. 7, pp. 945—956. DOI: 10.1134/S0012266118070121
26. Morozov A. Yu., Reviznikov D. L., Gidasov V. Yu. Adaptive interpolation algorithm based on a kd-tree for the problems of chemical kinetics with interval parameters. *Mathematical Models and Computer Simulations*, 2019, vol. 11, no. 4, pp. 622—633. DOI: 10.1134/S2070048219040100
27. Morozov A. Y., Abgaryan K. K., Reviznikov D. L. Mathematical model of a neuromorphic network based on memristive elements. *Chaos, Solitons & Fractals*, 2021, vol. 143, p. 110548. DOI: 10.1016/j.chaos.2020.110548

Acknowledgments

This work was supported by the RFBR grant No. 19-29-03051 MK.

Received August 15, 2020