

На правах рукописи

БУСЫГИН АЛЕКСАНДР НИКОЛАЕВИЧ

**ЗАПОМИНАЮЩАЯ МАТРИЦА НА ОСНОВЕ НАНОРАЗМЕРНОГО
КОМБИНИРОВАННОГО МЕМРИСТОРНО-ДИОДНОГО КРОССБАРА
ДЛЯ БИОМОРФНОГО НЕЙРОПРОЦЕССОРА**

Специальность 2.2.2 – электронная компонентная база
микро- и нанoeлектроники, квантовых устройств

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени
кандидата физико-математических наук

Тюмень 2021

Работа выполнена на кафедре прикладной и технической физики Федерального государственного автономного образовательного учреждения высшего образования «Тюменский государственный университет» (ФГАОУ ВО «ТюмГУ»)

Научный руководитель: **Удовиченко Сергей Юрьевич,**
доктор физико-математических наук,
профессор кафедры прикладной и технической
физики, руководитель НОЦ «Нанотехнологии»
ТюмГУ

Официальные оппоненты:

Ведущая организация:

Защита диссертации состоится “__” _____ 2022 г. в ____ на заседании диссертационного совета 24.1.106.01 при ФГБУН Институт проблем технологии микроэлектроники и особочистых материалов Российской академии наук по адресу: 142432, Московская обл., Черноголовка, ул. Академика Осипьяна 6.

С диссертацией можно ознакомиться в библиотеке ИПТМ РАН и на сайте http://www.iptm.ru/tes_council/diss.ru.html

Автореферат разослан « » 2022 г.

Ученый секретарь
диссертационного совета

Иржак Артемий Вадимович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы.

Нейронные сети, построенные на простых нейронах и используемые в информационных технологиях, предназначены для аппаратного ускорения расчетов и обеспечивают работу компьютерного зрения, машинного обучения и других систем со слабым искусственным интеллектом. Принятие решения в таких нейросетях происходит в результате выбора наиболее правдоподобного решения на основе ранее заложенных ассоциаций.

В [А3] представлена разработка биоморфного нейропроцессора на основе мемристорно-диодного кроссбара, реализующего аппаратную биоморфную импульсную нейросеть с большим числом нейронов для традиционных задач обработки информации, в том числе распознавания паттернов в видео и аудио информации, а также для воспроизведения работы кортикальной колонки мозга или её фрагмента. В качестве ключевых узлов аппаратной части нейропроцессора используются сверхбольшие запоминающая и логическая матрицы, представляющие собой массив синапсов и задающие вес и маршрут связи между нейронами соответственно. Указанные матрицы должны быть сверхбольшими потому, что каждый нейрон в сети может обладать большим количеством синаптических связей. Аналогов разработанного биоморфного нейропроцессора нет. Уникальность нейропроцессора состоит в том, что он построен на основе оригинальной электрической биоморфной модели нейрона и является биоморфным еще и с точки зрения выполнения функций биоморфной нейросети, созданной на основе оригинальной программной биоморфной модели нейрона.

В отличие от нейропроцессоров на простых нейронах биоморфный нейропроцессор дает возможность принимать решения не только на основе заранее заложенных ассоциаций, но и на основе новых ассоциаций (нового знания), формируемых в процессе обработки сигналов в динамично меняющихся условиях. Осмысливание полученных новых ассоциаций при совершенствовании оригинальной биоморфной нейросети [1] процессора обеспечит переход от слабого к сильному искусственному интеллекту.

В настоящее время существующие аппаратные средства в виде запоминающих матриц с мемристорным кроссбаром [2-6] выполняют узкоспециальные функции, в основном матричное умножение, которое представляет собой взвешенное суммирование входных импульсов напряжения. Аппаратные реализации ассоциативного самообучения и формирование новой ассоциации в импульсной нейросети с дискретными мемристорными синапсами представлены в [7-11]. Однако, предложенные электрические цепи аппаратной

реализации ассоциативной памяти не могут быть использованы для построения сверхбольшой аппаратной нейросети. Причиной является отсутствие интеграции мемристоров в кроссбары и наличие в схемах нейронов и синапсов большого числа активных электронных элементов с высоким энергопотреблением. Хотя аппаратные средства [5, 6] содержат в себе интегрированные мемристорные кроссбары, ассоциативное обучение в них является отложенным из-за программного расчета подстройки весов в периферийной системе после прохождения информации. Процедура обновления весов приводит к потере информации, так как во время этого входная информация игнорируется.

Цель работы: создание сверхбольшой запоминающей матрицы биоморфного нейропроцессора, являющейся массивом синапсов искусственных нейронов на основе комбинированного мемристорно-диодного кроссбара – нового компонента нанoeлектроники.

Задачи исследования:

- Разработать физические основы создания ячейки сверхбольшой запоминающей матрицы.
- разработать электрические схемы, топологии ячейки и сверхбольшой запоминающей матрицы (1000x1000 ячеек) в планарной и 3D геометрии с высокими интеграцией элементов и энергоэффективностью;
- построить физико-математические модели работы ячейки и процессов обработки сигналов и ассоциативного самообучения в запоминающей матрице;
- провести численное моделирование работы отдельной сверхбольшой запоминающей матрицы и сравнить с экспериментальными данными;
- провести численное моделирование матрицы малого размера в составе аппаратной импульсной нейросети и сравнить с экспериментальными данными.

Научная новизна.

1. Разработаны физические основы создания ячейки сверхбольшой запоминающей матрицы: необходимость высокого входного сопротивления, которое достигается при комплементарном включении двух мемристоров и обеспечивает малое энергопотребление; выбор в качестве селективного элемента диода Зенера, обеспечивающего существенно большую интеграцию элементов и меньшее энергопотребление по сравнению с активным селектором – транзистором.
2. Созданы электрическая схема и физико-математическая модель работы ячейки сверхбольшой запоминающей матрицы на основе комбинированного мемристорно-диодного кроссбара, ячейка которого содержит комплементарно включенных мемристоров и диода Зенера.

3. Показано, что применение диода Зенера в качестве селективного элемента, подключенного электрически последовательно к мемристору, позволяет увеличить размер кроссбара за счет уменьшения паразитных токов между ячейками, а также уменьшить деградацию выходного сигнала при суммировании входных импульсов напряжения в сверхбольшой матрице.
4. Предложена топология на кристалле как в планарном исполнении, так и с применением 3D интеграции на основе нового компонента наноэлектроники – комбинированного мемристорно-диодного кроссбара.
5. Разработаны электрическая схема и физико-математические модели процессов обработки сигналов в запоминающей матрице и её ассоциативного самообучения по правилам долговременной пластичности и пластичности, зависимой от времени импульса. Модели самообучения обеспечивают специфичную реализацию этих правил, связанную с оригинальностью мемристорно-диодной ячейки.
6. С помощью численного моделирования при использовании измеренных характеристик изготовленных мемристоров продемонстрированы способности отдельной запоминающей матрицы к обработке сигналов по взвешиванию импульсов напряжения и сложению токов ячеек, а также к ассоциативному самообучению запоминающей матрицы в составе аппаратной импульсной нейросети. Результаты численного моделирования хорошо согласуются с экспериментальными данными, что подтверждает корректность разработанных физико-математических моделей процессов, протекающих в запоминающей матрице, и дает основание для использования этих моделей при проектировании матриц сверхбольшого размера.

Научная и практическая значимость.

Разработанная запоминающая матрица на основе мемристорно-диодного кроссбара является одной из важнейших частей нейропроцессора, который аппаратно воспроизводит работу импульсной (биоморфной) нейросети, способной к ассоциативному самообучению. Запоминающая матрица в составе аппаратной биоморфной нейросети процессора генерирует новые ассоциации по биологически подобному механизму. Это позволяет говорить о формировании процессора нового поколения, который качественно отличается от существующих нейропроцессоров для компьютерного зрения, машинного обучения и других систем со слабым искусственным интеллектом.

Полученные результаты важны при переходе от существующих самообучающихся аппаратных импульсных нейросетей на дискретных мемристорах к сверхбольшим нейросетям с синапсами, интегрированными в мемристорно-диодный кроссбар.

С помощью запоминающей матрицы, встроенной в аппаратную импульсную нейросеть, поступление новой неизвестной информации отождествляется с наблюдаемой генерацией новых ассоциаций в биоморфном нейропроцессоре. Переход от слабого к сильному искусственному интеллекту возможен при совершенствовании биоморфной нейросети [1], которая позволит осмысливать эти ассоциации.

Предложенные электрические схемы и топологии сверхбольшой запоминающей матрицы совместимы с имеющимся оборудованием, применяемым для создания КМОП микросхем и защищены патентами РФ на изобретение №2649657 и №2697623.

Разработанная запоминающая матрица войдет в состав нейропроцессора при проектировании его экспериментального образца совместно с АО НИИ Молекулярной Электроники (НИИМЭ) – головного предприятия приоритетного технологического направления «Электронные технологии» РФ и его изготовлению на ПАО «Микрон», г. Москва, Зеленоград. Ранее было достигнуто соглашение о сотрудничестве между ТюмГУ и компанией АО «Ангстрем-Т» по созданию экспериментального образца аппаратного устройства нейропроцессора.

На защиту выносятся:

1. Разработанные физические основы, физико-математическая модель работы и электрическая схема ячейки сверхбольшой матрицы на основе комплементарно включенных мемристоров и диода Зенера с малым энергопотреблением и существенно большей интеграцией элементов по сравнению с активным селектором – транзистором.
2. Разработанные физико-математические модели процессов обработки сигналов и ассоциативного самообучения по правилам долговременной пластичности и пластичности, зависимой от времени импульса, электрическая схема и топология сверхбольшой планарной и 3D запоминающей матрицы на основе нового компонента нанoeлектроники – комбинированного мемристорно-диодного кроссбара.
3. Результаты численного моделирования процессов обработки сигналов при взвешивании импульсов напряжения и сложении токов ячеек в отдельной запоминающей матрице и ассоциативного самообучения запоминающей матрицы в составе аппаратной импульсной нейросети биоморфного нейропроцессора с использованием измеренных характеристик изготовленных мемристоров. Корректность разработанных физико-математических моделей, следующая из согласия результатов численного моделирования с экспериментальными данными и дающая основание для использования этих моделей при проектировании матриц сверхбольшого размера.

4. Генерация новых ассоциаций при численном моделировании нейросети с массивом синапсов, интегрированных в мемристорно-диодный кроссбар запоминающей матрицы, в отличие от ассоциативного самообучения аппаратных нейросетей с дискретными мемристорными синапсами.

Степень достоверности и апробация результатов.

Достоверность разработанных физико-математических моделей и полученных в работе результатов по численному моделированию больших электрических схем, содержащих мемристорно-диодные кроссбары, обеспечена применением известного симулятора SPICE (Simulation Program with Integrated Circuit Emphasis), использованием измеренных характеристик изготовленных мемристоров и согласием расчетных результатов с экспериментальными данными.

Основные результаты работы докладывались и обсуждались на Международных и Всероссийских конференциях в том числе:

- VIII, IX и XII конференциях Нанотехнологического общества России (Москва, 2016, 2017, 2021);
- First International Workshop «Nanoelectronic Memristive Devices for Quantum and Neuromorphic Computing (MEM-Q)» (Kurchatov Institute, Moscow, 2018);
- 6-я Международная научная конференция «Электронная компонентная база и микроэлектронные модули» (Республика Крым, г. Ялта, 2020)
- Международная научная конференция «Применение технологий виртуальной реальности и смежных информационных систем в междисциплинарных задачах FIT-M 2020» (г. Москва, 2020)
- Международная научно-техническая конференция «Smart energy systems 2021» (г. Казань, 2021)

Исследования, выполненные в ходе диссертационной работы, поддержаны грантами РФФИ №19-07-00272 «Электрофизические свойства комбинированного мемристорного-диодного кроссбара - нового компонента наноэлектроники, предназначенного для изготовления запоминающей и логической матриц нейропроцессора» и №19-37-90030 «Генерация нового знания в нейросети на основе массива мемристорных синапсов в запоминающей матрице биоморфного нейропроцессора и принципы увеличения быстродействия и энергоэффективности обработки информации на специализированном устройстве по сравнению с существующими вычислительными средствами».

Публикации. По результатам исследований опубликовано 16 научных работ, в том числе восемь в рецензируемых изданиях: три статьи в журналах, входящих в международные базы; две статьи, входящих в международную базу и в перечень ВАК РФ; одна статья в журнале из международной базы перечня ВАК РФ, а также два патента на изобретение.

Личный вклад соискателя. Разработка физико-математических моделей, весь объем численного моделирования и обработка полученных результатов выполнялись автором лично. Автор выражает благодарность научному руководителю, а также коллегам из НОЦ «Нанотехнологии» за помощь в проведении исследований и обсуждении полученных результатов.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения и списка литературы. Работа изложена на 94 страницах, включает 35 рисунков и 1 таблицу. Список литературы содержит 92 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, сформулированы цель работы и основные задачи, показана научная новизна и научная и практическая значимость работы, сформулированы положения, выносимые на защиту, и личный вклад автора в получении результатов работы, приводятся сведения об апробации работы, публикациях, структуре и объеме диссертации.

В первой главе проведен анализ опубликованных работ по запоминающим матрицам с точки зрения интеграции элементов и энергопотребления. Выполнена оценка пригодности существующих запоминающих матриц к вычислениям, необходимым для реализации сверхбольших нейросетей со способностью к самообучению.

Вторая глава посвящена разработке физических основ создания ячейки, электрических схем и топологий ячейки и сверхбольшой запоминающей матрицы нейропроцессора на базе мемристорно-диодного кроссбара, обладающей высокими интеграцией элементов и энергоэффективностью. Представлены физические основы создания ячейки сверхбольшой запоминающей матрицы: необходимость высокого входного сопротивления, которое достигается при комплементарном включении двух мемристоров и обеспечивает малое энергопотребление; выбор в качестве селективного элемента диода Зенера, обеспечивающего существенно большую интеграцию элементов и меньшее энергопотребление по сравнению с активным селектором – транзистором.

Ячейка кроссбара (рис.1) содержит два мемристора M1 и M2 с общим электродом, совмещенным с анодом диода Зенера D1 [A1, A2]. Применение диода Зенера в качестве селективного элемента, подключенного электрически последовательно к мемристорам, позволяет увеличить размер кроссбара за счет уменьшения паразитных токов между ячейками, а также уменьшить деградацию выходного сигнала при суммировании входных импульсов напряжения в матрице. Выбор в пользу пассивного элемента – диода по сравнению с традиционно применяемым в качестве селективного активного элемента –

транзистора сделан из-за его меньших размеров и энергопотребления. Кроме этого транзистор требует наличия дополнительного управляющего проводника.

На нижнем электроде ячейки 3 расположены диоды Зенера, состоящие из последовательно наносимых слоев сильнолегированного n-полупроводника, слаболегированного n-полупроводника и сильнолегированного p-полупроводника. На аноде диода располагается общий металлический электрод комплементарных мемристоров, непосредственно соприкасающийся с вышележащим сплошным мемристорным слоем, на котором расположены верхние электроды ячейки 1 и 2. Комплементарное включение двух мемристоров позволяет постоянно поддерживать высокое входное сопротивление ячейки в рабочем режиме, что обеспечивает ее низкое энергопотребление.

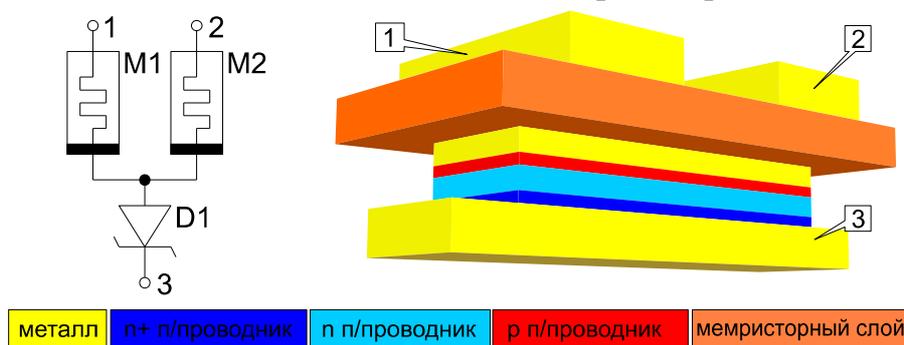


Рис.1. Электрическая схема и топология отдельной ячейки кроссбара.

Разработанные ячейки объединены в планарные двухслойные запоминающие матрицы [A7], которые вертикально объединены в 3D структуру из трех горизонтально расположенных и зеркально ориентированных по отношению друг к другу кроссбаров (рис.2), позволяющую увеличить интеграцию элементов и сократить число проводников [A2].

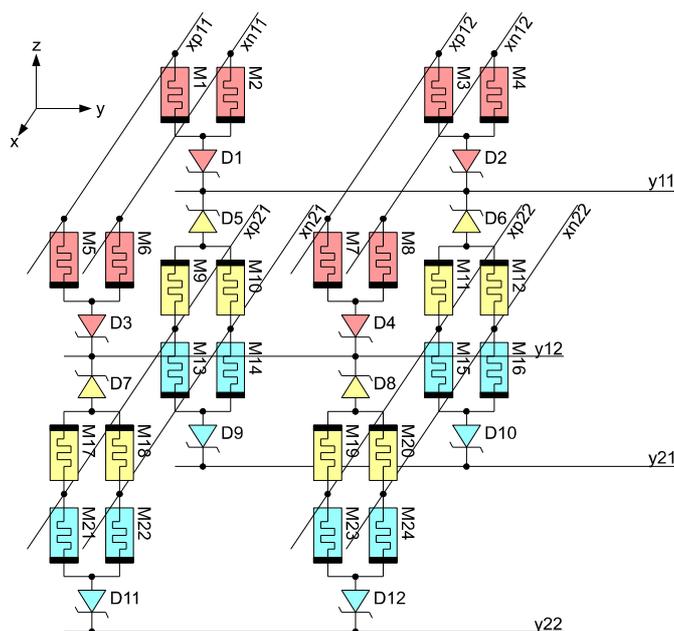


Рис.2. Электрическая схема фрагмента трехмерной запоминающей матрицы.

Три цвета соответствуют трём комбинированным кроссбарам.

Соответствующая топология 3D матрицы приведена на рис.3. Каждый кроссбар на рисунке содержит 18 ячеек [A2, A8].

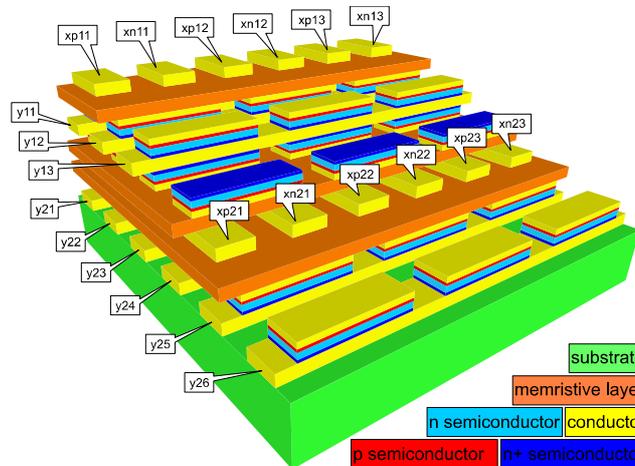


Рис.3. Топология фрагмента 3D запоминающей матрицы.

В третьей главе представлены разработка физико-математических моделей ячейки и процессов обработки сигналов, результаты численного моделирования работы матрицы и сравнение с экспериментальными данными.

С целью разработки сверхбольшой запоминающей матрицы создана физико-математическая модель работы ячейки на основе упрощенных моделей мемристора и диода Зенера. В модели мемристора изменение параметра состояния было жестко ограничено по сравнению с [12], поскольку неабсолютная точность рациональных чисел в компьютерной системе приводит к выходу параметра состояния за границы допустимого интервала. Для диода Зенера построена идеализированная модель, нелинейная вольт-амперная характеристика которой представляет собой кусочную функцию из трех прямых линий [A2].

В ячейке матрицы производится умножение напряжения входных импульсов на весовой коэффициент с помощью закона Ома. Напряжение на выходе ячейки, подключенной к схеме (например, сумматора) с высоким входным сопротивлением при приложении к контактам мемристоров ячейки напряжения разной полярности амплитудой U_{input} имеет вид:

$$U_i = \left(\frac{2U_{input}}{R_{1i} + R_{2i}} R_{2i} - U_{input} \right) \frac{1}{R_d + R_p} R_p, \quad (1)$$

где R_{1i} и R_{2i} – сопротивления мемристоров комплементарной пары, R_d – сопротивление прямосмещенного диода Зенера, R_p – входное сопротивление сумматора. Сопротивление мемристоров представляет собой континуум значений в пределах от минимального R_{on} до максимального R_{off} .

Так как мемристоры образуют комплементарную пару, сумма их сопротивление остается постоянной: $R_{1i} + R_{2i} = K = const$. Учитывая, что

сопротивление открытого диода мало, $R_D \ll R_P$, то справедливо выражение для напряжений на выходе ячейки:

$$U_i = \frac{U_{input}}{K} (2R_{2i} - K). \quad (2)$$

Выходной ток ячейки при подключении к схеме с низким входным сопротивлением:

$$I_i = \left(\frac{2}{K} R_{2i} - 1 \right) \frac{U_{input}}{R_D} \quad (3)$$

Процесс изменения состояния ячейки производился путем последовательного изменения сопротивления двух мемристоров. Численный расчет с использованием разработанной модели ячейки запоминающей матрицы показывает (рис.4), что запись одной ячейки из комплементарных мемристоров в матрице размером 100×100 требует в 8 раз меньше энергии при добавлении в каждую ячейки диода Зенера, значительно уменьшающего паразитные токи через соседние ячейки. Для моделирования взят худший возможный при записи случай, когда все ячейки матрицы изначально находились в одинаковом состоянии. Отношение сопротивлений мемристора в низкопроводящем и высокопроводящем состояниях $R = (R_{off} - R_{on})/R_{on} = 10$.

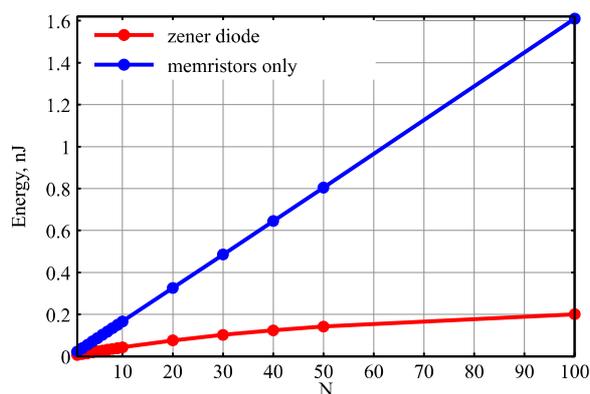


Рис.4. Затраты энергии на запись одной ячейки комплементарных мемристоров без диода (синий цвет) и с диодом Зенера (красный цвет) в зависимости от числа ячеек в квадратной матрице $N \times N$.

Физико-математическая модель процессов, происходящих в электрической схеме запоминающей матрицы с комплементарными мемристорами [A2] отличается от соответствующей модели матрицы с одним мемристором в ячейке тем, что мемристоры образуют резистивный делитель напряжения. Основная процедура в запоминающей матрице – взвешенное суммирование, эквивалентное умножению матрицы на вектор, в рамках этой модели производится с использованием закона Ома для умножения входных напряжений на проводимости ячеек и первого закона Кирхгофа, для суммирования получившихся токов.

Численный расчет процесса прохождения одного импульса амплитудой 1 В с использованием разработанной модели ячейки запоминающей матрицы показывает (рис. 5) деградацию выходного напряжения от размера квадратной матрицы $N \times N$ для R равным 10, 100 и 1000.

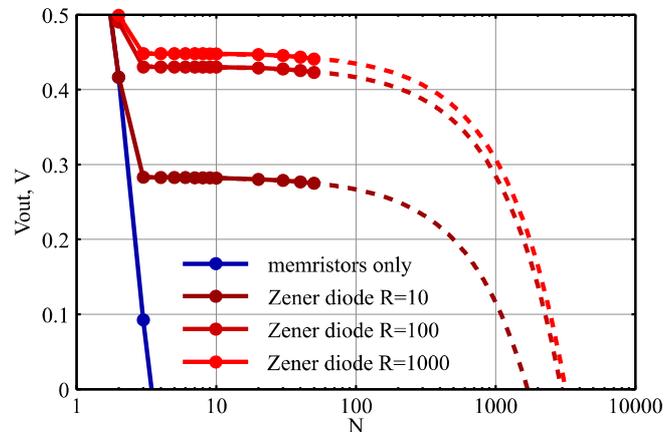


Рис.5. Зависимость выходного напряжения от размера матрицы при последовательном чтении при разных R .

Из рис.5 видно, что при отсутствии в ячейках диода Зенера выходное напряжение снижается практически до нуля уже в 3×3 матрице. При добавлении диода Зенера происходит снижение выходного напряжения в пределах от 50% до 70% в матрице того же размера, а дальнейшее увеличение размера матрицы слабо влияет на величину выходного сигнала. Медленно изменяющийся уровень выходного напряжения $\sim 0,3$ В достаточен для выполнения дальнейшей процедуры суммирования. Размер матрицы при этом оказывается 1000×1000 , что считается сверхбольшим. Следовательно, матрица с комплементарными мемристорами и без нелинейного селективного элемента не может быть использована в качестве сверхбольшой матрицы нейропроцессора.

Суммирование напряжений выполняется последовательной подачей напряжения на выбранные ячейки. На выходе сумматора:

$$U_S = \sum_i U_i = \frac{U_{input}}{K} \sum_i (2R_{2i} - K). \quad (4)$$

Таким образом, соотношение электрических величин в аппаратной части, описываемое выражением (4), соответствует аргументу активационной функции в формуле (1) следующим образом:

$$x_i \rightarrow U_{input}, \quad w_i \rightarrow \frac{2R_{1i} - K}{K}.$$

Сигнал U_{input} на входные шины подается импульсами, причем на один вход ячейки импульс приходит неизменным, а на другой – в инвертированном виде. Импульсы на общей шине создают колебания потенциала U_i , которые

поступают на вход сумматора, на выходе которого уровень напряжения U_s определяется частотой поступающих импульсов на вход матрицы. Таким образом, импульсы напряжения в комплементарной паре мемристоров, работающих как делитель напряжения, взвешиваются аналогично синаптическому взвешиванию в нейронной сети, и в зависимости от записанных в мемристорах проводимостей поступают в сумматор. Взаимовлияния ячеек не будет поскольку входные инверторы неиспользуемых ячеек находятся в высокоимпедансном состоянии.

С помощью численного моделирования проведена проверка работоспособности одного столбца матрицы при выполнении операций взвешивания и суммирования импульсов напряжения, которая заключалась в организации трех состояний ячеек: оба мемристора закрыты, один из мемристоров пары открыт [A2]. На рис.6а показано изменение во времени выходного напряжения сумматора V_{out} .

Сложение токов, протекающих через закрытые диоды Зенера, осуществляется при низком входном сопротивлении усилителя. Сумма выходных токов ячеек I_s :

$$I_s = \sum_i \left(\frac{2}{K} R_{2i} - 1 \right) \frac{U_{input}}{R_D} \quad (5)$$

Взвешивание напряжений входных импульсов с последующим суммированием токов позволяет уменьшить паразитные токи между ячейками и произвести корректное сложение одновременно пришедших импульсов. На рис.6б. показан выходной ток одной шины матрицы как результат сложения токов из двух ячеек, находящихся в разных синаптических состояниях [A3].

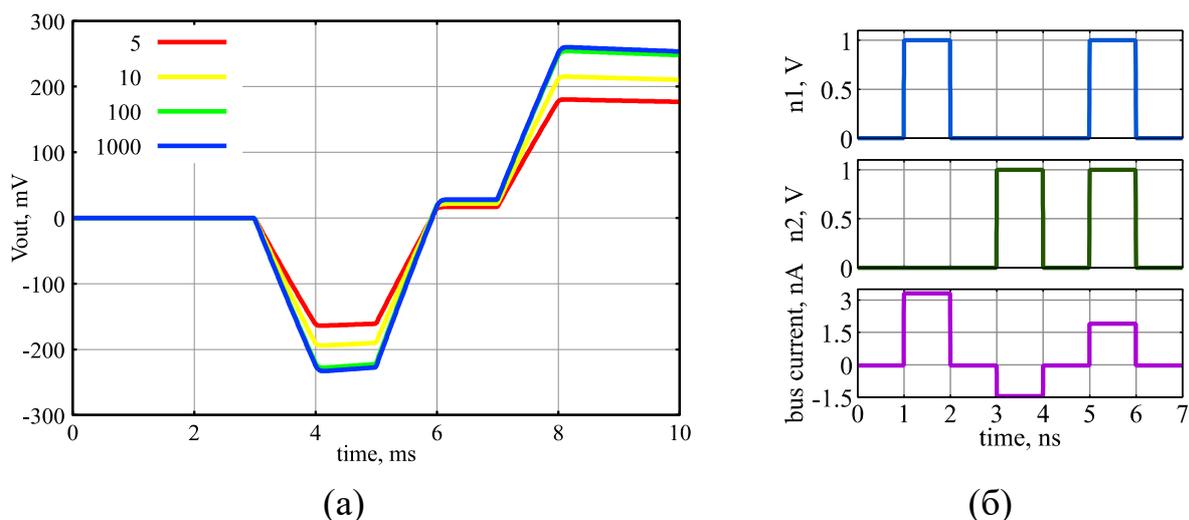


Рис.6. Взвешенное суммирования входных импульсов напряжения: (а) – при высоком входном сопротивлении сумматора. Цвет соответствует различным R ; (б) – при низком входном сопротивлении сумматора.

SPICE-модели позволяют оптимизировать параметры элементов ячейки и матрицы в целом. Из рисунка видно, что чем выше отношение R сопротивлений мемристоров в низкопроводящем и высокопроводящем состояниях, тем сильнее состояние ячейки влияет на выходное напряжение сумматора.

При моделировании матричного умножения входного вектора на матрицу чисел [A4] на входные проводники кроссбара запоминающей матрицы поступают импульсы напряжения. Выходные токи кроссбара являются суммой токов от отдельных ячеек, которые в свою очередь однозначно определяются значениями сопротивления мемристоров. На рис.7 показаны выходные токи кроссбара размером 2x2 при трех разных состояниях весовых коэффициентов. Выходной ток с первого выходного проводника обозначен как Out[0], а второй – Out[1]. Для сравнения приводятся экспериментальные данные по матричному умножению из [13, 14].

Результаты моделирования показывают, что полученный выходной вектор токов соответствует результату матричного умножения матрицы на вектор, полученному по математическим правилам.

Из-за высокого коэффициента усиления в преобразователе ток-напряжение и превышения максимального входного напряжения АЦП в микроконтроллере ожидаемый высокий ток на Out[0] в последнем случае не наблюдается. Среднеквадратичное отклонение выходных токов за время действия входных импульсов напряжения, вносимая неодинаковостью мемристоров, оценена при сравнении экспериментальных данных и SPICE-моделирования и составляет 276 нА.

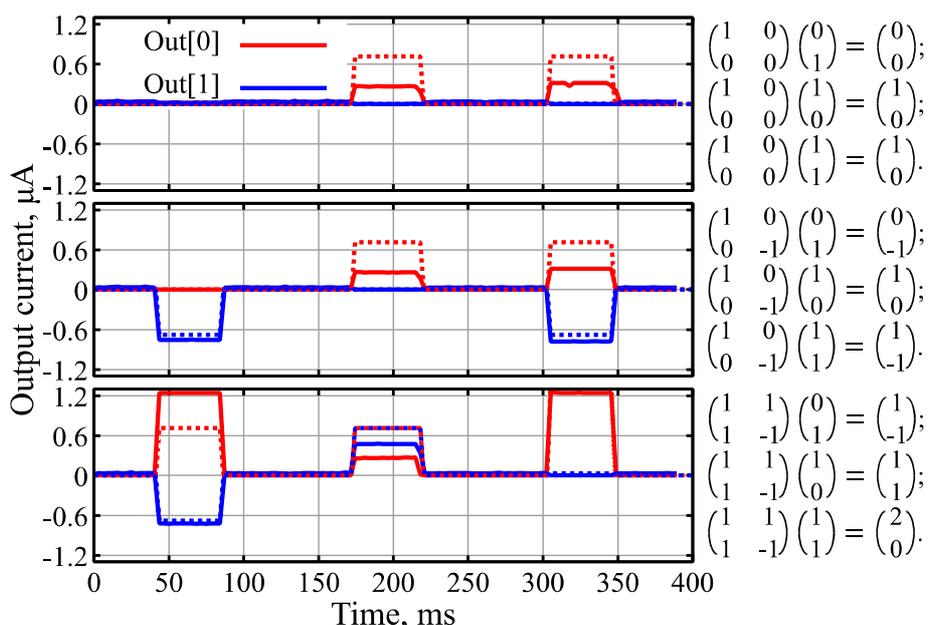


Рис.7. Результат матричного умножения матрицы чисел размером 2x2 на двухкомпонентный вектор: сплошные кривые – экспериментальные, пунктирные кривые - SPICE-моделирование.

Результаты численного моделирования показывают преимущество предложенной запоминающей матрицы по сравнению с известными аналогами с точки зрения затухания сигнала, энергоэффективности и интеграции элементов.

В четвертой главе представлены физико-математические модели ассоциативного самообучения, реализующиеся в запоминающей матрице, и результаты моделирования ее работы в составе импульсной нейронной сети с ассоциативным самообучением. Эти модели обеспечивают специфичную реализацию правил самообучения, связанную с оригинальностью мемристорно-диодной ячейки.

Возможность использования разработанной запоминающей матрицы в качестве массива синапсов аппаратной нейросети была предварительно продемонстрирована в ходе моделирования работы импульсного однослойного персептрона с 49 входами и 10 выходами (рис.8) [A12]. Синаптические веса при этом были заранее вычислены в ходе обучения с учителем программного неимпульсного персептрона и перенесены в состояния мемристоров запоминающей матрицы.

В приведенном примере входная картинка соответствует рукописной цифре «4». В результате моделирования работы аппаратной нейросети в течении одной секунды видно, что нейросеть распознала входное изображение как цифру «4», поскольку частота срабатывания выходного нейрона, ассоциированного с цифрой «4» является наибольшей (14 Гц).

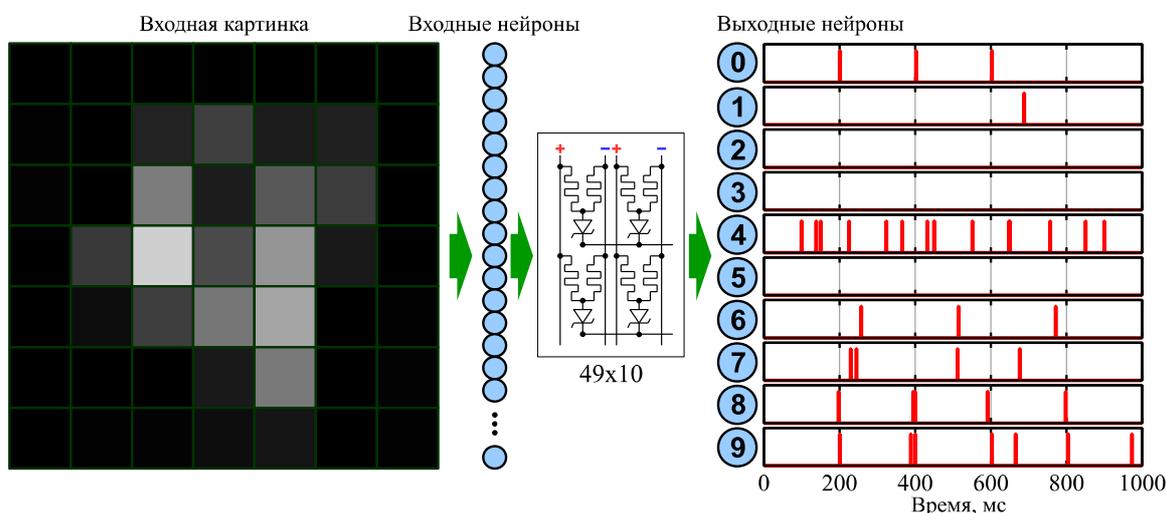


Рис.8. Результат распознавания цифры «4» из датасета MNIST импульсным персептроном с синапсами на основе мемристорно-диодного кроссбара.

Обучение синапсов в запоминающей матрице происходит по правилу Хебба, как и в реальном синапсе: сила связи между одновременно активировавшимися нейронами увеличивается. Изменение веса синапса зависит от разницы между временами срабатывания пресинаптического и постсинаптического нейронов $\Delta t = t_{pre} - t_{post}$. С целью реализации известных

правил индуцированной долговременной потенциации (long-term potentiation - LTP) и пластичности, зависимой от времени импульсов (spike-time dependent plasticity – STDP), разработаны физико-математические модели ассоциативного самообучения LTP и STDP в запоминающей матрице [A5, A10]. Оригинальная электрическая схема ячейки требует специфичной реализации правил самообучения. В обеих моделях самообучения пресинаптический сигнал представляет собой пару импульсов напряжения одинаковой амплитуды, но разной полярности вследствие использования комплементарной пары мемристоров в ячейке. Нейрон при срабатывании отправляет обратно в синапс импульс отрицательного напряжения $post$. Изменение силы синапса при этом произойдет если на синапс придут импульсы от пресинаптического нейрона pre^+ и pre^- (рис.9а,в). Форма пресинаптических импульсов определяет реализуемое локальное правило обучения.

Правило LTP предполагает увеличение веса синапса при совпадении во времени пресинаптических и постсинаптических сигналов. Поэтому в соответствующей модели знак импульсов напряжения пресинаптических сигналов поддерживается постоянным, что вызывает изменение состояние мемристоров в ячейке всегда в сторону уменьшения сопротивления и увеличению веса синапса.

Правило STDP действует аналогично LTP при $\Delta t < 0$. При $\Delta t > 0$, согласно этому правилу, вес синапса должен уменьшаться. В соответствующей модели такое уменьшение реализовано с помощью знакопеременных импульсов напряжения. Итоговое направление изменения состояния пары комплементарных мемристоров и веса синапсов определяется преобладающими знаками напряжения пресинаптических импульсов во время действия постсинаптического импульса и описывается кусочной функцией с 6 областями, например при $-\tau_1 < \Delta t < \tau - \tau_1$:

$$\Delta w(\Delta t) = \left[\frac{2R_0R_d}{\beta V_p(\tau_1 + \Delta t)} - R_d \right]^{-1}, \quad (6)$$

где $R_0 = 0,5 \cdot (R_{off} + R_{on})$, τ – ширина начальной фазы пресинаптического импульса и постсинаптического импульса, τ_1 – ширина второй фазы пресинаптического импульса равна, β – коэффициент, определяющий скорость переключения мемристора, V_p – максимальная разница напряжений между пре- и постсинаптическими импульсами.

Изменение силы синапса будет тем больше, чем ближе по времени произойдет активация пресинаптического и постсинаптического нейронов, что отражает максимум функции синаптической пластичности $\Delta w(\Delta t)$ на рис.9 б,г.

Увеличение весового коэффициента синапса соответствует возникновению новой ассоциации (нового знания) в нейросети.

Аналитическое представление функции пластичности позволяет выбирать параметры в модели мемристора для достижения оптимальной скорости обучения нейросети, обеспечивающую сходимость численного расчета нейросети за минимальное время.

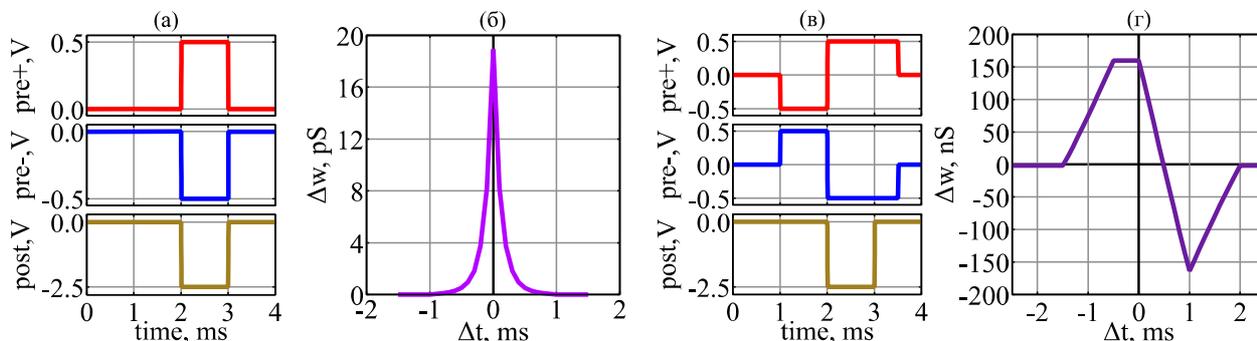


Рис.9. Долговременная потенцияция и пластичность, зависящая от времени импульса: (а), (в) – пре и постсинаптические напряжения ячейки запоминающей матрицы при $\Delta t = 0$, (б), (г) – функция пластичности $\Delta w(\Delta t)$ для LTP и STDP соответственно.

При моделировании долговременной потенцияции (LTP) [A3] выходные импульсы двух нейронов через синаптические связи, представленные комплементарными мемристорно-диодными ячейками, поступают на вход третьего нейрона (рис.10).

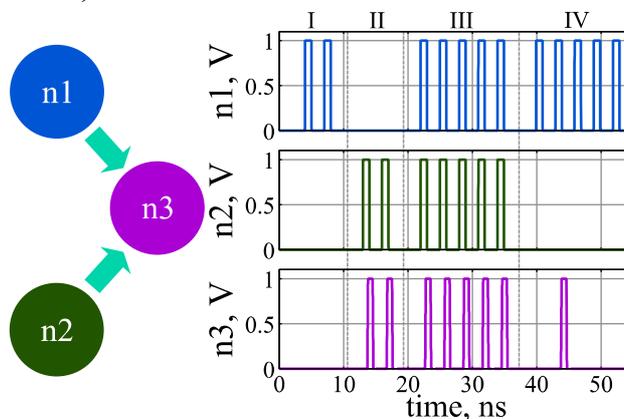


Рис.10. Возникновение новой ассоциации в группе из трех нейронов. Импульсы напряжения, генерируемые нейронами на разных этапах ассоциативного обучения.

Моделирование этой же нейросети (рис.10) с пластичностью, зависимой от времени импульса (STDP), приводит к тому же результату, поскольку ситуации с $\Delta t > 0$ отсутствуют.

Проведено численное моделирование ассоциативного самообучения аппаратной импульсной нейросети в виде однослойного перцептрона с использованием запоминающей матрицы в качестве массива синапсов [A6]. Электрические параметры ячеек при моделировании соответствовали усредненным

характеристикам мемристорно-диодного кроссбара (рис.11) с числом ячеек 4×2 , который изготовлен по магнетронной технологии в одном технологическом модуле путем осаждения проводящих дорожек, полупроводниковых слоев (p-Si/n-Si) диода Зенера и активного слоя мемристоров из смешанных оксидов металлов ($Ti_{0,93}Al_{0,07}O_x$) при одновременном распылении двух мишеней.

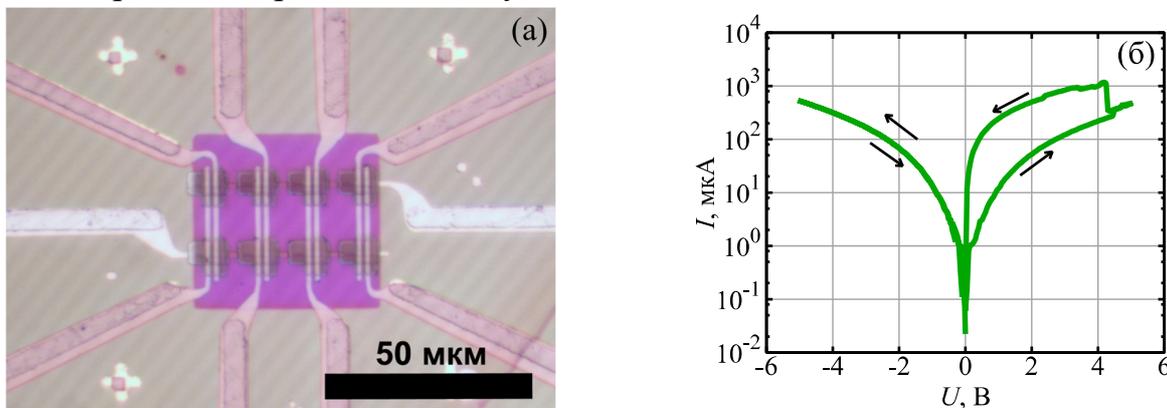


Рис.11. Микрофотография изготовленного мемристорно-диодного кроссбара $TiN/Ti_{0,93}Al_{0,07}O_x/p-Si/n-Si/W$ - (а); усредненная ВАХ его ячеек - (б).

Электрическая схема аппаратной импульсной нейросети (рис.12) построена на основе мемристорно-диодного кроссбара (рис.11) с 4-мя парами входных проводников и двумя выходными шинами [А6].

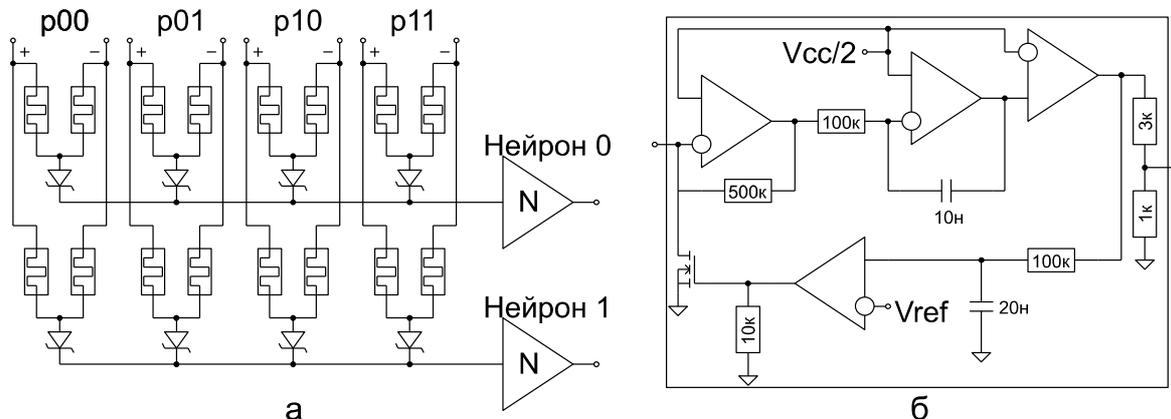


Рис.12. Электрическая схема аппаратной нейросети: а) мемристорно-диодный кроссбар в качестве массива синапсов, б) электрическая схема нейрона.

Соответственно кроссбар содержит восемь ячеек, являющимися синапсами нейросети. Слой персептрона образован двумя нейронами, построенными на основе операционных усилителей. Электрическая схема нейрона состоит из преобразователя ток-напряжение, аналогового интегратора, компаратора, схемы задержки в виде интегрирующей RC-цепи и полевого транзистора.

Нейросеть, состоящая из четырех виртуальных входных нейронов и двух аппаратных выходных нейронов (рис.13) [А6], обучалась по правилу STDP для выполнения задачи распознавания картинок с разрешением 2×2 пикселя.

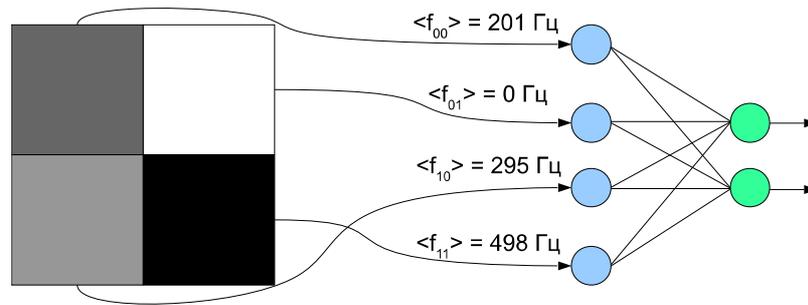


Рис.13. Входное изображение и архитектура аппаратной нейросети: синий цвет – входные виртуальные нейроны, зеленый – выходные нейроны.

Значения яркостей пикселей преобразовывались в среднюю частоту последовательности входных импульсов с помощью микроконтроллера. Выходные импульсы активировавшихся нейронов фиксировались этим же микроконтроллером. Каждый цикл обучения содержал 128 картинок, образованных путем добавления шума к исходным двум эталонным картинкам.

Модельная кривая обучения (рис.14а) показывает долю правильно классифицированных входных картинок от числа циклов (эпох) обучения [А9].

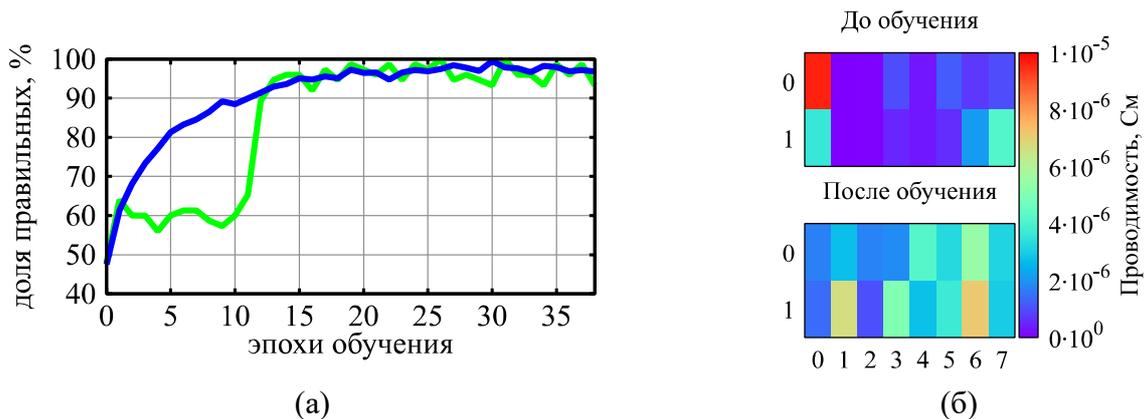


Рис.14. Результат ассоциативного самообучения аппаратной нейросети: а) изменение доли правильных классификаций: модельная (синяя) и экспериментальная (зеленая) кривые; б) изменение проводимости мемристоров кроссбара.

Для проверки адекватности физико-математической модели ассоциативного самообучения STDP на рис.14 кроме модельной кривой изменения доли правильных классификаций картинок показана аналогичная экспериментальная кривая, полученная при обучении изготовленной аппаратной нейросети (рис.11) [А9]. Модельная кривая обучения более сглаженная по сравнению с экспериментальной. Это объясняется тем, что изменение проводимости в модели мемристора происходит плавно, без скачков. Расположение прямоугольников на рис.14б соответствует положению мемристоров на рис.12а. Изначально проводимость мемристоров была низкой за

исключением одного мемристора. В процессе ассоциативного самообучения проводимость этого мемристора уменьшилась, а остальных выросла.

На рис.15 показаны суммарные входные импульсы напряжения нейрона после преобразователя ток-напряжение, полученные в результате SPICE моделирования процесса ассоциативного самообучения аппаратной импульсной нейросети с мемристорно-диодным кроссбаром [A9].

Сначала подаются пресинаптические импульсы напряжения на оба синапса и происходит потенциация слабой связи между нейронами. Соответствующий входной сигнал для нейрона показан светло-синим цветом. Затем импульс подается только на слабый синапс, что порождает слабый входной импульс напряжения для нейрона (темно-синий цвет). Потенциация слабого синапса проявляется в увеличении амплитуды второго импульса и происходит из-за того, что они активируются одновременно, и напряжение на мемристорах, соответствующих этой связи, выше порогового.

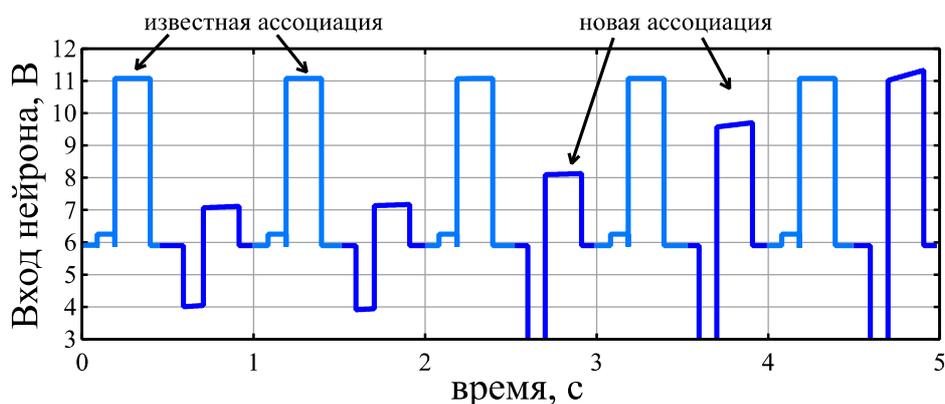


Рис.15. Результат SPICE моделирования процесса генерации новой ассоциации на фоне известной.

Нарастание входного напряжения нейрона, которое пропорционально синаптическому току, является признаком формирования новой ассоциации и вызвано усилением синапса согласно правилу STDP.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Созданы физические основы, электрическая схема и физико-математическая модель работы ячейки запоминающей матрицы на основе комплементарно включенных мемристоров и диода Зенера.

2. Разработаны электрическая схема и физико-математическая модель процессов обработки сигналов в сверхбольшой запоминающей матрице, представляющей собой массив синаптических связей аппаратной нейросети, реализующейся в нейропроцессоре на основе мемристорно-диодного кроссбара и обладающей высокими интеграцией элементов и энергоэффективностью.

3. С помощью численного моделирования показано, что применение диода Зенера в качестве селективного элемента, подключенного электрически последовательно мемристору, позволяет увеличить размер мемристорно-диодного кроссбара запоминающей матрицы вплоть до 10^6 ячеек за счет уменьшения паразитных токов между ячейками, а также уменьшить деградацию выходного сигнала при суммировании входных импульсов напряжения в матрице. Показано, что разработанная запоминающая матрица обладает меньшей деградацией сигнала и большей энергоэффективностью по сравнению с известными аналогами.

4. Предложена топология сверхбольшой запоминающей матрицы на кристалле в планарном исполнении и с применением 3D интеграции. В результате сравнения с известными аналогами показано, что разработанная запоминающая матрица обладает высокой интеграцией элементов.

5. С помощью численного моделирования показана работоспособность малой матрицы размером 2×2 ячеек при взвешивании импульсов напряжения и сложении токов ячеек. Результаты численного моделирования хорошо согласуются с экспериментальными данными, что подтверждает корректность разработанной физико-математической модели работы матрицы и дает основание для использования этой модели при проектировании матриц сверхбольшого размера.

6. Разработаны физико-математические модели ассоциативного самообучения запоминающей матрицы, которые обеспечивают специфичную реализацию правил самообучения LTP и STDP, связанную с оригинальностью мемристорно-диодной ячейки.

7. С помощью численного моделирования продемонстрирована способность запоминающей матрицы с числом ячеек 4×2 к ассоциативному самообучению в составе аппаратной импульсной нейросети. Результаты численного моделирования хорошо согласуются с экспериментальными данными, что подтверждает корректность разработанной физико-математической модели обучения по правилу STDP.

8. Демонстрирована генерация новой ассоциации (нового знания) при численном моделировании аппаратной импульсной нейросети на основе мемристорно-диодного кроссбара в отличие от ассоциативного самообучения в существующих аппаратных нейросетях с синапсами на базе дискретных мемристоров.

Список публикаций по теме диссертации

Публикации в журналах, входящих в международные базы данных:

- A1. Maevsky O.V., Pisarev A.D., Busygin A.N., Udovichenko S.Yu. Complementary memristive diode cells for the memory matrix of a neuromorphic processor // International Journal of Nanotechnology. 2018. V.15. №4/5. P.388-393.

- A2. Pisarev A., Busygin A., Udovichenko S., Maevsky O. 3D memory matrix based on a composite memristor-diode crossbar for a neuromorphic processor // *Microelectronic Engineering*. 2018. V.198. P.1-7.
- A3. Pisarev A.D., Busygin A.N., Udovichenko S.Yu., Maevsky O.V. A biomorphic neuroprocessor based on a composite memristor-diode crossbar // *Microelectronics Journal*. 2020. V. 102. Article 104827.

*Публикации в журнале, входящем в международную базу
Chemical Abstracts Service и в перечень ВАК:*

- A4. Pisarev A.D., Busygin A.N., Bobylev A.N., Gubin A.A., Udovichenko S.Yu. Simulation and experimental evaluation of signal processing routines in a composite memristor-diode crossbar the basis of a biomorphic neuroprocessor // *Nanoindustry*. 2020. V.13. № S5-3(102). P. 724-732.
- A5. Udovichenko S.Yu., Pisarev A.D., Busygin A.N., Bobylev A.N. Biomorphous neuroprocessor – prototype of a new generation computer being a carrier of artificial intelligence. Part 2 // *Nanoindustry*. 2021. V. 14. № 1(102). P. 68-79.

Публикация в журнале из перечня ВАК, входящем в международные базы:

- A6. Bobylev A.N., Busygin A.N., Gubin A.A., Pisarev A.D., Udovichenko S.Yu. Manufacture and testing of a pulsed hardware neural network with memristor synapses for a biomorphic neuroprocessor // *Nanobiotechnology Reports*. 2021. V.16. N.6 P. 761-766.

Бобылев А.Н., Бусыгин А.Н., Губин А.А., Писарев А.Д., Удовиченко С.Ю. Изготовление и тестирование аппаратной импульсной нейросети с мемристорными синапсами для биоморфного нейропроцессора // *Российские нанотехнологии*. 2021. Т.16. № 6. С. 793-798.

Патенты:

- A7. Маевский О.В., Писарев А.Д., Бусыгин А.Н., Удовиченко С.Ю. Запоминающее устройство на основе комплементарной мемристорно-диодной ячейки // Патент №2649657. Дата публикации патента 04.04.2018.
- A8. Писарев А.Д., Бусыгин А.Н., Удовиченко С.Ю., Бобылев А.Н., Маевский О.В. 3D запоминающая матрица на основе комплементарной мемристорно-диодной ячейки // Патент №2697623. Дата публикации патента 15.08.2019.

Публикации в других изданиях:

- A9. Бусыгин А.Н., Бобылев А.Н., Губин А.А., Писарев А.Д., Удовиченко С.Ю. Численное моделирование и экспериментальное исследование аппаратной импульсной нейросети с мемристорными синапсами // *Вестник Тюменского*

государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. 2021. Т.7. №2 (26). С. 223-235.

- A10. Удовиченко С.Ю., Бусыгин А.Н., Писарев А.Д. Ассоциативное самообучение аппаратной импульсной нейросети, созданной на основе мемристорно-диодного кроссбара // Сборник тезисов XII конференции Нанотехнологического общества России. 2021.
- A11. Ибрагим А.Х., Бусыгин А.Н., Удовиченко С.Ю. Компьютерная симуляция биоморфного нейропроцессора // Сборник тезисов II Международной научно-технической конференции «Smart energy systems 2021» (SES-2021). С.35,48.
- A12. Удовиченко С.Ю., Писарев А.Д., Бусыгин А.Н., Ибрагим А.Х., Компьютерная симуляция биоморфного нейропроцессора — носителя искусственного интеллекта // Сборник тезисов международной научной конференции «Применение технологий виртуальной реальности и смежных информационных систем в междисциплинарных задачах FIT-M 2020», 2020, С. 14-19.
- A13. Писарев А.Д., Бусыгин А.Н., Бобылев А.Н., Губин А.А., Удовиченко С.Ю. Моделирование и аппаратное тестирование процессов обработки сигналов в комбинированном мемристорно-диодном кроссбаре, являющемся основой биоморфного нейропроцессора // Сборник тезисов 6-ой международной научной конференции «Электронная компонентная база и микроэлектронные модули». Наноиндустрия (Спецвыпуск). 2020. Т. 13. № S4 (99). С. 591-593.
- A14. Удовиченко С.Ю., Писарев А.Д., Бусыгин А.Н., Бобылев А.Н. Высокая интеграция элементов в логической и запоминающей матрицах нейропроцессора с помощью комбинированного мемристорно – диодного кроссбара // Сборник тезисов IX конференции Нанотехнологического общества России. 2018. С. 6-7.
- A15. Udovichenko S.Yu., Pisarev A.D., Busygin A.N., Bobylev A.N. Memory and universal logic matrices for neuroprocessor. // First International Workshop “Nanoelectronic Memristive Devices for Quantum and Neuromorphic Computing”. Book of Abstracts. 2018. P. 19.
- A16. Удовиченко С.Ю., Маевский О.В., Писарев А.Д., Бусыгин А.Н., Комплементарная мемристорно-диодная ячейка для запоминающей матрицы нейроморфного процессора. // Сборник тезисов VIII конференции Нанотехнологического общества России. 2017. С. 37-40.

Список цитируемой литературы

1. Filippov V.A., Bobylev A.N., Busygin A.N., Pisarev A.D., Udovichenko S.Yu. A biomorphic neuron model and principles of designing a neural network with memristor synapses for a biomorphic neuroprocessor // Neural Computing and Applications. 2020. V.32. Pp. 2471-2485.

2. Li C., Belkin D., Li Y. et al. Efficient and self-adaptive in-situ learning in multilayer memristor neural networks // *Nature Communications*. 2018. V. 9. P. 2385.
3. Ielmini D. Brain-inspired computing with resistive switching memory (RRAM): Devices, synapses and neural networks // *Microelectronic Engineering*. 2018. V.190. P. 44–53.
4. Yao P., Wu H., Gao B, et al. Fully hardware-implemented memristor convolutional neural network // *Nature*. 2020. V. 577. P. 641-661.
5. Cai F., Correl J.M., Lee S.H., et al. A fully integrated reprogrammable memristor–CMOS system for efficient multiply–accumulate operations // *Nature Electronics*. 2019. V. 2. P. 290–299.
6. Wang .Z., Li C., Song W., et al. Reinforcement learning with analogue memristor arrays // *Nature Electronics*. 2019. V. 2. P. 115–124.
7. Wang Z., Wang X., Novel A. Memristor-Based Circuit Implementation of Full-Function Pavlov Associative Memory Accorded with Biological Feature // *IEEE Transactions on Circuits and Systems I:Regular Papers*. 2018. V. 65(7). P. 2210-2220.
8. Minnekhanov A.A., Emelyanov A.V., Lapkin D.A. et al. Parylene based memristive devices with multilevel resistive switching for neuromorphic applications // *Scientific Reports*. 2019. V. 9. Article 10800.
9. Demin V.A. Nekhaev D.V., Surazhevsky I.A. et al. Necessary conditions for STDP-based pattern recognition learning in a memristive spiking neural network // *Neural Networks*. 2021. V. 134. P. 64-75.
10. Antonov I. N., Belov A. I., Mikhaylov A. N., et al. Formation of Weighting Coefficients in an Artificial Neural Network Based on the Memristive Effect in Metal–Oxide–Metal Nanostructures // *Journal of Communications Technology and Electronics*. 2018. V. 63(8). P. 950–957.
11. Surazhevsky I.A., Demin V.A., Ilyasov A.I., et al. Noise-assisted persistence and recovery of memory state in a memristive spiking neuromorphic network // *Chaos, Solitons and Fractals*. 2021. V. 146. Article 110890
12. Biolek D., Di Ventra M., Pershin Y.V. Reliable SPICE Simulations of Memristors, Memcapacitors and Meminductors // *Radioengineering*. 2013. V. 22(4). P. 945-968.
13. Писарев А.Д., Бусыгин А.Н., Бобылев А.Н., Губин А.А., Удовиченко С.Ю. Исследование электрофизических свойств комбинированного мемристорно-диодного кроссбара, являющегося основой для аппаратной реализации биоморфного нейропроцессора // *Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика*. 2020. Т.6. №3. С. 93-109.
14. Pisarev A., Busygin A., Bobylev A., Gubin A., Udovichenko S. Fabrication technology and electrophysical properties of a composite memristor-diode crossbar used as a basis for hardware implementation of a biomorphic neuroprocessor // *Microelectronic Engineering*. 2021. V. 236. Article 111471.