

А. И. Галушкин, д-р техн наук, проф., зам. зав. каф.,
 Московский физико-технический институт, г. Долгопрудный, e-mail: neurocomputer@yandex.ru

Мемристоры в развитии высокопроизводительной вычислительной техники

Излагается мнение автора о перспективах развития данного направления сверхпроизводительной вычислительной техники в связи с появлением мемристоров.

Ключевые слова: мемристоры, суперЭВМ, аналоговые вычисления, многослойные нейронные сети, клеточные нейронные сети

Введение

Ниже представлена концепция разработки нейрокомпьютеров с применением мемристоров, ориентированная на большое и очень большое число вычислительных элементов и, как следствие, на большое число мемристоров.

На рис. 1 условно представлена сфера применения мемристорных систем, сформировавшаяся за последние несколько лет. Несомненно, эта сфера будет дополняться и расширяться. В данной работе изложены основные принципы применения мемристоров при создании и применении нейрокомпьютеров, как раздела сверхвысокопроизводительной вычислительной техники.

В настоящее время во всем мире реализуется несколько десятков подобных проектов. Одним из характерных проектов является проект, представленный в работе [2].

В России разработана концепция создания вычислительных систем экзафлопного уровня с использованием классических принципов построения кластерных суперЭВМ, а также концепция развития высокопроизводительных вычислений на базе супернейрокомпьютеров [3].

Необходимо отметить два основных свойства современных суперЭВМ:

- масштабируемость;
- двухслойность архитектур.

С нашей точки зрения для перехода на экзафлопный уровень вычислений необходима реализация по крайней мере еще одного шага: изменения логического базиса алгоритмов решения задач и, соответственно, элементной базы с переходом от носителя информации в виде уровней токов и напряжений в электрических схемах к носителю информации в виде частоты узких импульсов, подобно тому, как это имеет место в реальной нервной системе.

Основными предпосылками принципиального изменения архитектуры вычислительных систем при переходе к системам экзафлопной производительности являются следующие:

- необходимость резкого повышения надежности за счет отказа от фон-Неймановской архитектуры вычислительных систем из элементов булевой логики И, ИЛИ, НЕ, когда в системе из значительного числа этих элементов происходит катастрофический отказ при отказе любого элемента;
- необходимость резкого снижения энергопотребления за счет отказа от существующего позиционного метода представления информации и перехода к другим, обеспечивающим снижение энергопотребления, и как следствие, дополнительное повышение надежности;
- необходимость резкого повышения однородности схемотехники элементов вычислительных систем, что должно привести к повышению эффективности при том резком увеличении интеграции элементов, которое будет иметь место при переходе к экзафлопной производительности.

На рис. 2 отмечены основные этапы развития сверхвысокопроизводительной вычислительной техники за последние десятилетия. Отмечается, что мемристоры являются одним из эффективных путей развития данного вида техники на ближайшее десятилетие.

На рис. 3 условно представлена динамика развития технологий нейрокомпьютеров, начиная с 50-х годов прошлого столетия. В 50-е и 60-е годы прошлого столетия для изготовления нейрокомпьютеров в основном использовали аналоговую технологию при незначительной доле программной реализации на универсальных ЭВМ. С появлением микропроцессоров, а также мощных универсальных цифровых ЭВМ, аналоговые ЭВМ вследствие их ограниченной точности, не-

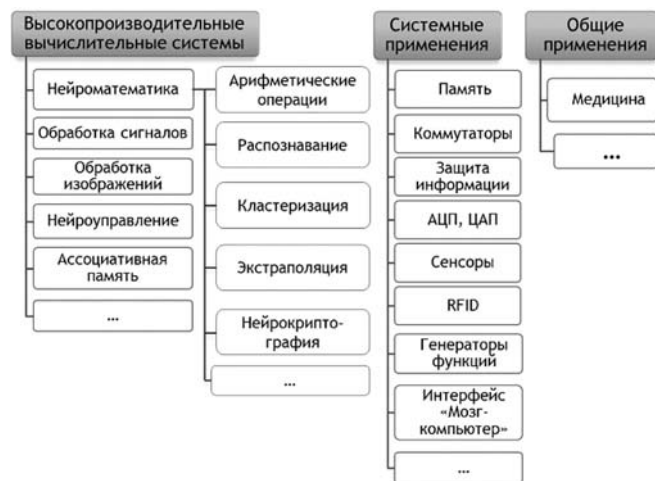


Рис. 1. Функциональные задачи для мемристорных систем (уровень 2014 г.)

1. Предпосылки: архитектура вычислительных систем

В настоящее время высказываются серьезные сомнения в перспективности КМОП-технологии и других технологических решений для решения задачи перехода от петафлопного к экзафлопному уровню производительности вычислительных систем. В работе [1] представлено мнение двух известных экспертов: Харста Саймона и Томаса Стерлинга о том, что существуют серьезные сомнения в достижении к 2020 г. производительности экзафлопного уровня (10^{18}). Основные сомнения касаются энергопотребления, которое для 10-петафлопной системы составляет 10 МВт, а для экзафлопной системы по расчетам должно достигать 2 ГВт. В этой же работе отмечается необходимость рассмотрения кроме КМОП-технологий и других технологий, а также других типов архитектур, отличных от фон-Неймановской.

Время	Производительность	Технология
1990	1 Гигафлоп/с	Транспьютерные системы
1999	1 Терафлоп/с	Кластерные суперЭВМ
2009	1 Петафлоп/с	Графические процессоры
2020	1 Эксафлоп/с	Мемристоры

Рис. 2. Этапы развития сверхвысокопроизводительной вычислительной техники за последние десятилетия



Рис. 3. Развитие технологий нейрокомпьютеров. Мемристоры — детище нанотехнологий

смотря на большое быстродействие, теряют свою роль, оставаясь в период 80-х, 90-х и даже 2000-х годов предметом разработки достаточно большого числа специализированных нейрочипов, наряду с большим числом разработок специализированных цифровых нейрочипов.

Разработка мемристоров возрождает аналоговую обработку, резко понижая энергопотребление, увеличивая скорость обработки при ее контролируемой точности, характерной для высокопараллельных нейросетевых структур с ограниченным числом слоев логической обработки.

2. Предпосылки: теория нейронных сетей, нейроматематика и нейруправление

Российская школа теории нейронных сетей зародилась в середине 60-х годов прошлого столетия [4, 5]. С самого начала этих работ применительно к адаптивным системам автоматического управления нейронная сеть как объект управления была выбрана в структуре, показанной на рис. 4. Нейронная сеть представляет собой высокопараллельную многослойную структуру с настраиваемыми последовательными, перекрестными и обратными связями.

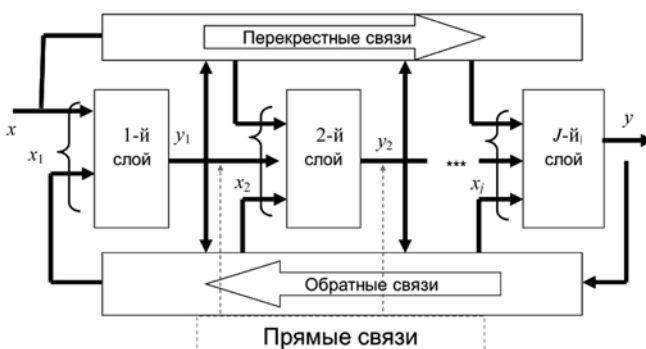


Рис. 4. Многослойная нейронная сеть (концепция 1967 г.)

Ниже представлена основная аксиоматика теории нейронных сетей, характерная для российской научной школы:

- Вероятностная модель мира.
- Нейронные сети — логический базис решения любых задач.
- Разработка нейросетевых алгоритмов, специфичных для конкретной выбранной задачи.
- Отказ от использования субъективных нейросетевых парадигм.
- Нейронная сеть — частный вид многомерного нелинейного динамического объекта управления.
- Любая идея нейросетевого алгоритма должна быть ориентирована на эффективную аппаратную реализацию в соответствие с текущей или перспективной технологией.
- Любая идея в части нейросетевых технологий не эффективна, если не может быть перенесена из одно- или двумерной иллюстрации на многомерную иллюстрацию.

Работы в области нейронных сетей получили развитие в работах [6–8].

В 2007 г. на международной конференции по нейронным сетям (IJCNN — 07, Orlando, USA) была организована российская секция "Overview of Soviet/Russian Neural Network R & D: The Untold Story" под руководством Роберта Хехт-Нильсена. Кроме ряда российских докладов была представлена монография [9] с предисловиями Лотфи Заде, Амари и Р. Хехт-Нильсона, которые отметили высокий уровень российских научных работ в этой области. Работа [10] является итоговой в этой области.

Нейроматематика — новый раздел вычислительной математики, связанный с разработкой нейросетевых алгоритмов решения сложных формализуемых и неформализуемых задач. Основные результаты российской научной школы в области нейроматематики представлены в работах [11–13].

Нейруправление — это новый раздел теории управления, связанный с применением нейрокомпьютеров для идентификации и управления динамическими системами. Нейруправление является направлением теории управления различными сложными системами:

- сильно нелинейными;
- с изменяемыми параметрами;
- с изменяемой структурой;
- многомерными;
- распределенными.

Работы российской научной школы представлены в работах [14, 15] и более десяти монографиях по прикладным задачам нейруправления различными объектами.

3. Предпосылки: реализации нейрокомпьютеров

На рис. 5 представлены этапы развития реализаций нейрокомпьютеров в России.

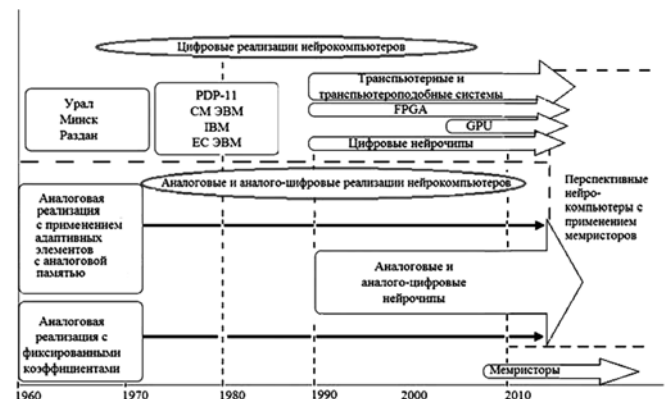


Рис. 5. Развитие реализаций нейрокомпьютеров в России

Реализация аналоговых нейрокомпьютеров с использованием адаптивных элементов с аналоговой памятью в 60-е годы прошлого века оказалась слишком громоздкой и дорогостоящей. Кроме этого существенным недостатком подобных реализаций нейрокомпьютеров являлось достаточно большое время перестройки коэффициентов нейронных сетей.

На рубеже 60-х и 70-х годов прошлого столетия для применения аналоговых нейрокомпьютеров был выбран класс задач, для которых сбор архива машинных данных для обучения был достаточно длительным и дорогостоящим. При этом необходимость в перестройке коэффициентов нейронных сетей либо отсутствовала, либо была необходима через значительный период времени (1 месяц, 1 год и т. д.). Для этих целей были реализованы аналоговые нейрокомпьютеры с коэффициентами, проставляемыми вручную (рис. 6, 7).

Период 70-х, 80-х годов прошлого столетия был периодом программной реализации нейрокомпьютеров на микропроцессорах и универсальных ЭВМ различного типа (PDP-11, IBM, их общегражданских и специальных отказоустойчивых аналогов).

Начало 90-х годов прошлого века характеризовалось сразу несколькими направлениями разработок нейрокомпьютеров, связанных с развитием технологии микроэлектроники:

- программные реализации нейрокомпьютеров на базе транспьютерных систем (десятки, сотни процессоров с сопроцессорами обработки сигналов и изображений);
- нейрокомпьютеры на базе ПЛИС (программируемых логических интегральных схем);
- цифровые, цифроаналоговые и аналоговые нейрочипы.

Эти разработки в значительной степени отражены в монографии [18].

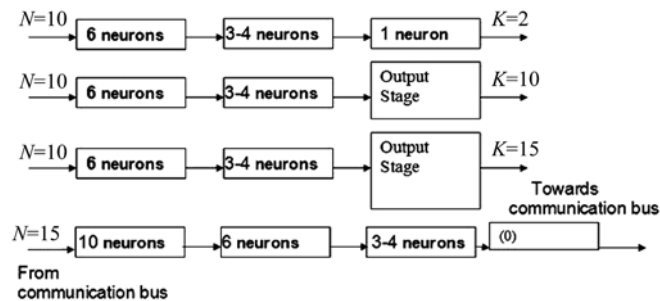


Рис. 6. Функциональная структура нейрокомпьютера (1970 г.)



Рис. 7. Общий вид нейрокомпьютера (1974 г.). Размерность входных признаков — 30, число классов 8

Период 2005—2007 гг. характеризуется активным становлением суперЭВМ на базе графических процессоров и началом разработки для них программных эмуляторов больших нейронных сетей [19].

В настоящее время (см. рис. 5) научно-технический задел для разработки будущих нейрокомпьютеров с применением мемристоров составляют:

- программные реализации нейронных сетей на транспьютерных и кластерных суперЭВМ;
- реализация нейрокомпьютеров на ПЛИС;
- программные реализации нейрокомпьютеров на суперЭВМ с использованием графических процессоров;
- многолетние разработки цифровых, аналого-цифровых и аналоговых нейрочипов.

4. Выбор технологии изготовления мемристоров. Оценка производительности вычислительных систем с применением мемристоров

Мемристоров изготавливают на базе следующих элементов:

- электрохимические элементы;
- оксид титана;
- оксид тантала;
- полимеры;
- ферроэлектрики;
- углеродные нанотрубки;
- кремний;
- аморфный кремний;
- поликремний.

В результате разработок реализованы мемристоровые матрицы и мемристоровые системы с различными физическими свойствами, которые необходимо оценивать с точки зрения их применимости в конкретных структурах вычислительных систем.

Применимость той или другой технологии изготовления мемристоровых систем необходимо оценивать как с точки зрения существующих критериев оценки производительности, так и с точки зрения оценки производительности будущих нейрокомпьютеров с применением мемристоров.

Производительность нейрокомпьютеров с применением мемристоров необходимо оценивать с точки зрения следующих количественных показателей:

- число эмулируемых нейронов;
- число эмулируемых связей;
- число переключений связей в секунду;
- потребляемая мощность;
- число физических переключений до момента отказа (дополнительно).

Расчет производительности мемристоровых систем при анализе технологий необходимо выполнять на нескольких уровнях:

- на уровне элементарных операций;
- на уровне базовых нейросетевых систем на мемристорах;
- на уровне СБИС с учетом внутренней коммуникационной среды;
- на уровне плат с учетом внутрисплатной коммуникационной среды;
- на уровне блоков с учетом внеплатной коммуникационной среды;
- на уровне супернейрокомпьютера с учетом межблочной коммутационной среды.

Необходимо высказать некоторое предостережение тем, кто будет оценивать производительность мемристоровых систем. В истории развития вычислительной техники известны примеры, когда оптимистическая оценка производительности на уровне элементарных операций предлагаемой технологии оказывалась несостоятельной при распространении оценки производительности на более общие архитек-

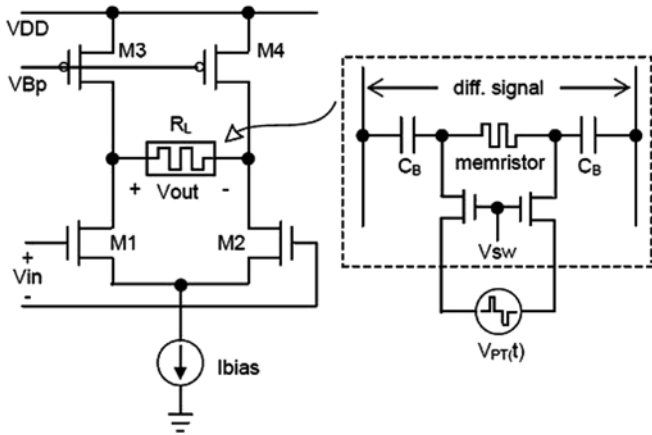


Рис. 8. Реализация синапса нейрона с применением мемристора

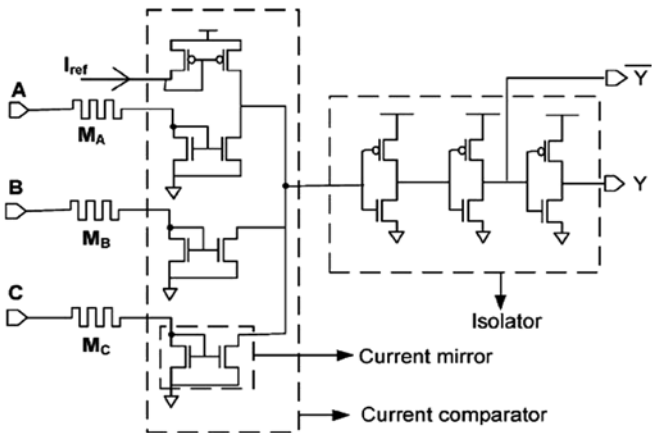


Рис. 9. Схема трехвходового нейрона с применением мемристоров

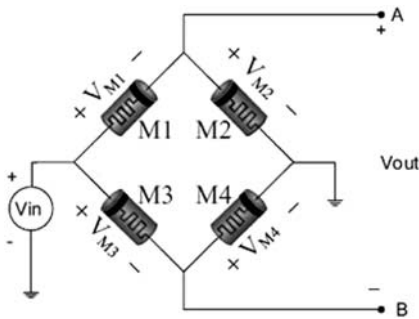


Рис. 10. Схема мемристорного моста

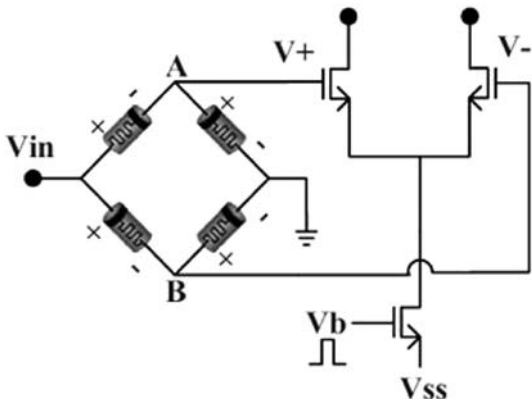


Рис. 11. Схема синапса с мостом из мемристоров

турные элементы, платы, блоки, стойки и т. д. и при расширении круга реализуемых алгоритмов (т. е. в попытке сделать вычислительную систему более универсальной). Так произошло с оптическими ЭВМ, в которых элементарные операции (перемножение матриц) выполнялись очень быстро, однако системы так и не стали универсальными. То же произошло с компьютерными системами с сопроцессорами обработки сигналов и изображений IMSA100, IMSA110 фирмы *Inmos*, в которых операции умножения—сложения выполнялись очень быстро, но в конкретных структурах одномерных и двумерных аппаратно реализованных Z-фильтров.

Исследование производительности мемристорных систем необходимо начинать с разработки инструментальных систем исследования мемристоров и мемристорных матриц, включая следующие этапы.

1. Исследование, разработка и экспериментальная реализация принципов стыковки мемристорных матриц с элементами КМОП-технологии.
2. Разработка инструментальной системы для исследования мемристорных матриц с выходом на персональную ЭВМ, включая драйвер.
3. Разработка контроллера на ПЛИС для стыковки мемристорных систем с персональными ЭВМ, с драйвером.
4. Экспериментальная и теоретическая оценка понижения энергопотребления в мемристорных системах при переходе к представлению сигналов в виде частотно-модулированной последовательности импульсов.

5. Реализация нейрона с применением мемристоров

В схемной реализации нейрона мемристоры выполняют функцию синапсов — перестраиваемых весовых коэффициентов. Возможно несколько вариантов применения мемристоров для этой цели. Ниже представлены несколько известных вариантов, требующих доработки с точки зрения включения нейрона в схему нейронной сети и с точки зрения включения схемы нейронной сети в общую схему СВИС-нейрочипа.

В работе [20] представлен один из вариантов реализации синапса нейрона с применением мемристоров (рис. 8). Вариант принципиальной схемы многовходового нейрона представлен на рис. 9 [21]. На рис. 10 представлена мостовая схема соединения мемристоров, обеспечивающая реализацию положительных и отрицательных значений весовых коэффициентов в нейронной сети [22]. В той же работе в развитие мостовой схемы соединения мемристоров представлена схема синапса с мостом из мемристоров (рис. 11).

В работе [23] представлен другой вариант соединения мемристоров в схеме синапса (рис. 12) и предлагается схема изменения знака весового коэффициента в синапсе, представленная на рис. 13.

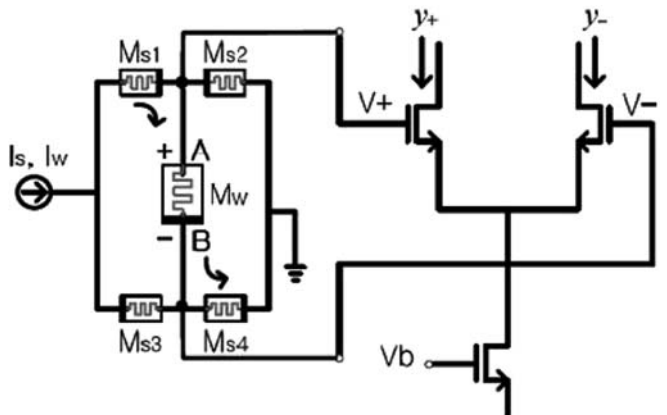


Рис. 12. Вариант реализации синапса с применением мемристоров

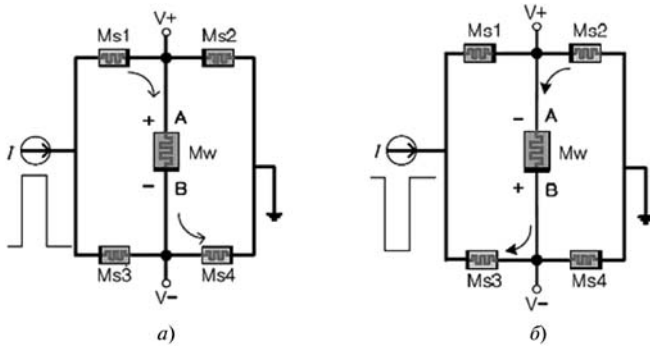


Рис. 13. Предлагаемая схема изменения знака весового коэффициента в синапсе с применением мемристоров: а — положительная конфигурация; б — отрицательная конфигурация

В соответствии с мостовыми схемами соединения мемристоров, представленными в работе [22], в работе [24] показана схема нейрона с несколькими мемристорными входами.

6. Реализация нейронных сетей с применением мемристоров

На рис. 14 представлена простейшая двухслойная нейронная сеть, ориентированная на реализацию с применением мемристоров [22], а на рис. 15 (см. третью сторону обложки) ее реализация с применением мемристоров.

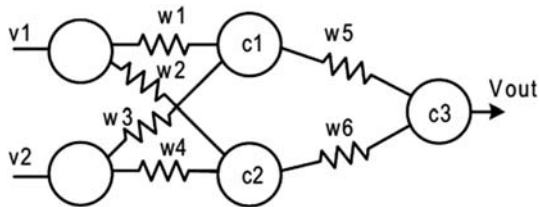


Рис. 14. Типовая двухслойная нейронная сеть

Анализ публикаций 2008—2014 гг. показал, что попытки реализации нейронных сетей с применением мемристоров касаются в основном многослойных нейронных сетей следующих видов:

- с полными последовательными связями;
- с обратными связями (рекуррентные нейронные сети);
- клеточные нейронные сети, ориентированные на обработку изображений.

Рассматриваются нейронные сети, ориентированные на достаточно узкий класс прикладных задач:

- бинарные нейронные сети с бинарными многомерными входными сигналами и бинарными (0, 1) коэффициентами;
- RBF-нейронные сети;
- СМАС-нейронные сети.

В работе [25] предложена (рис. 16) схема реализации весового коэффициента с мемристорами для клеточной нейронной сети. В работе [26] предложена другая схема применения мемристоров в клеточной нейронной сети.

Разработка и реализация принципиальных схем нейронных сетей различных структур с применением мемристоров является предметом дальнейших исследований в ближайшие годы.

7. Настройка нейронных сетей с применением мемристоров

С точки зрения автора, работы по реализации алгоритмов настройки нейронных сетей с применением мемристоров имеют самый начальный характер и в основном касаются реализации простейших алгоритмов типа правила Хебба [27]. Ориентация мемристоров как детища нанозлектроники на реализацию нейронных сетей с большим и очень большим числом нейронов и настраиваемых коэффициентов делает задачу разработки и реализации алгоритмов настройки важной, сложной, в значительной степени определяемой архитектурой будущих вычислительных систем с применением мемристоров.

Замечание. Отметим, что правило Хебба, в его логическом варианте в России было реализовано диалоговым интерактивным режимом работы в мостовой системе прецизионного измерения сопротивления (рис. 17), выпускаемой во время второй мировой войны сотнями тысяч штук.

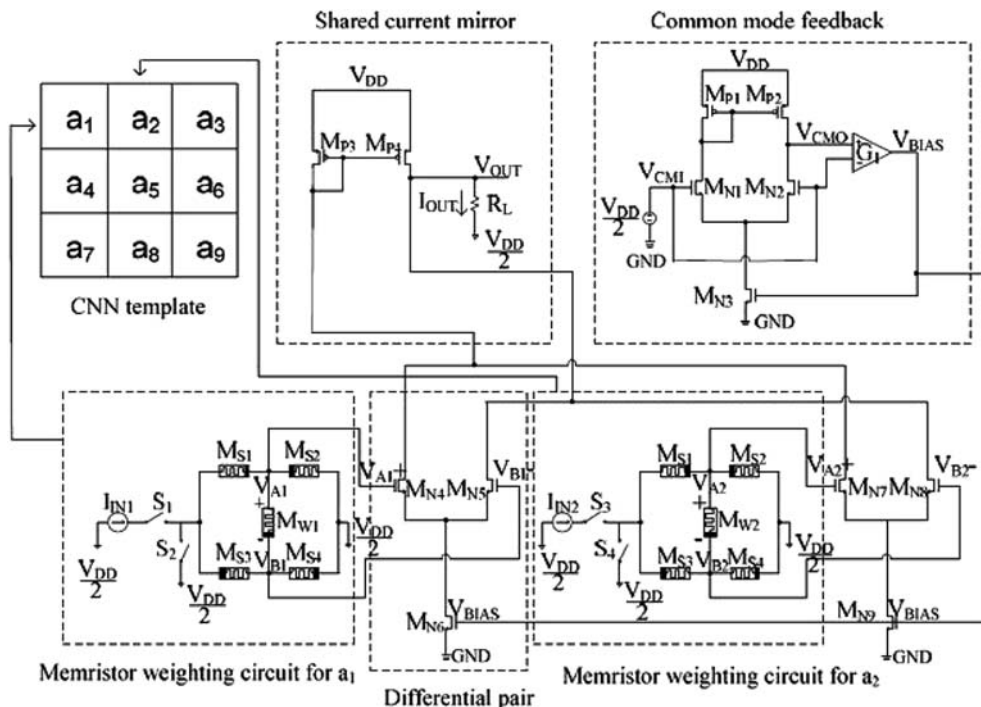


Рис. 16. Схема реализации весового коэффициента с мемристорами для клеточной нейронной сети

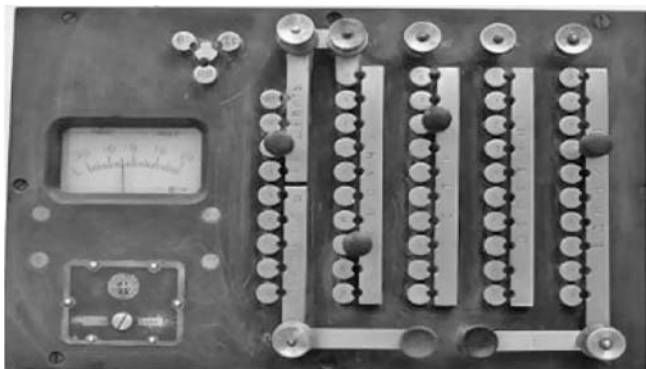


Рис. 17. Отечественный прецизионный измеритель сопротивления, функционирующий по правилу — аналогу правила Хейбба



Рис. 18. Общая структура работ по созданию алгоритмов настройки сетей с применением мемристоров

В работе [28] представлен аналитический подход к сравнению простейших алгоритмов настройки применительно к мемристорным системам. Эти методы требуют развития как с точки зрения алгоритмов адаптации, используемых в MATLAB (Neural Network Toolbox), так и с точки зрения алгоритмов адаптации нейронных сетей, учитывающих ограничения на весовые коэффициенты [5, 6, 8—10]. Условно, если в 60-е годы прошлого века алгоритмы настройки формировали для нейронных сетей как объекта управления:

$$y = f \Sigma a_3 \cdot f \Sigma a_2 \cdot f \Sigma a_1 \cdot x$$

с ограничениями на коэффициенты 1, 2, 3-го слоев нейронной сети, то для мемристорных систем эти алгоритмы должны формироваться как для более сложного объекта:

$$y = f \Sigma a_3(z) \cdot f \Sigma a_2(z) \cdot f \Sigma a_1(z) \cdot x.$$

Здесь x — входной сигнал нейронной сети; y — выходной сигнал нейронной сети; f — функция активации нейрона; $a_1(z)$, $a_2(z)$, $a_3(z)$ — передаточные функции мемристоров, формируемые моделями в системах PSpice, Cadence, MATLAB.

На рис. 18 представлена общая структура работы по созданию алгоритмов настройки нейронных сетей с применением мемристоров. Для нейрокомпьютеров с использованием моделей мемристоров на базе созданных алгоритмов настройки многослойных нейронных сетей с учетом ограничений на коэффициенты необходимо разработать специфические для мемристорных систем алгоритмы:

- адаптации нейронных сетей;
- распараллеливания нейросетевых алгоритмов с контуром адаптации на архитектуру вычислительных систем с применением мемристоров.

В связи с этим представленная в работе [22] схема обучения нейронной сети с применением мемристоров с реализацией блока настройки на хост-ЭВМ вряд ли жизнеспособна для вариантов с большим числом мемристоров (рис. 19, см. третью сторону обложки).

Распараллеливание алгоритмов адаптации и требование повышения быстродействия должны привести к внутрикристальной реализации алгоритма обучения, причем на цифровом и аналоговом уровнях.

8. СБИС — нейрочипы с применением мемристоров

Разработка СБИС-нейрочипов с применением мемристоров является одним из важнейших направлений работ в области создания перспективных нейрокомпьютеров. При этом необходима аналитическая работа для обеспечения качественной разработки мемристорных систем, в том числе:

- анализ разработок адаптируемых систем с аналоговой памятью 60-х—70-х годов прошлого столетия в целях использования в современных разработках мемристорных систем;
- анализ разработок аналоговых и аналого-цифровых нейрочипов последних десятилетий в целях использования в современных разработках мемристорных систем.

При этом анализ разработок аналоговых и аналого-цифровых нейрочипов должен касаться следующих разделов:

- аналоговые нейрочипы;
- аналого-цифровые нейрочипы;
- клеточные нейрочипы;
- нейрочипы с частотно-импульсным представлением сигналов;
- оптические и оптоэлектронные нейрочипы;
- молекулярные нейрочипы;
- специализированные аналоговые и аналого-цифровые нейрочипы:
- АПП;
- СМАС;
- обработки изображений;
- нейроуправления;
- ассоциативной памяти;
- обработки речевой информации;
- другие.

Типичным примером домемристорной разработки цифровых нейрочипов с частотно-импульсной модуляцией сигналов является схема, представленная на рис. 20 [29]. Необходимо отметить несколько попыток реализации СБИС-нейрочипов:

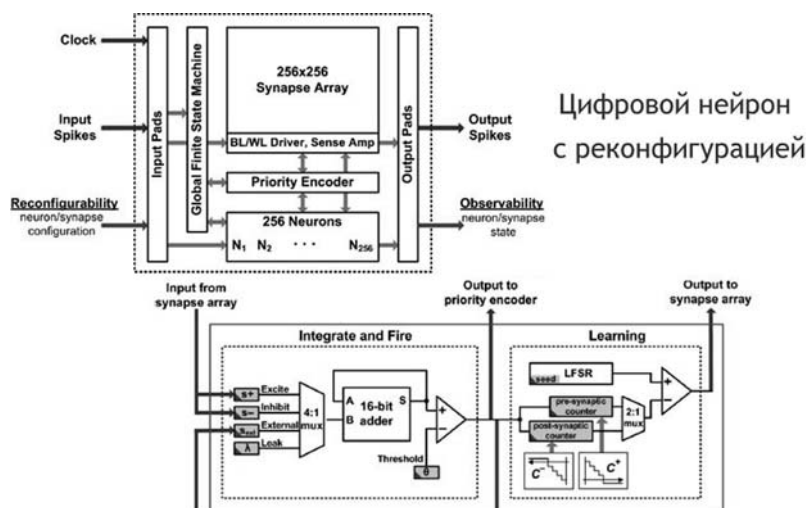


Рис. 20. Пример домемристорных разработок цифровых нейрочипов с частотно-импульсной модуляцией сигналов

пов с применением мемристоров (цифровых и аналого-цифровых) [30–32] — рис. 21–23.

Особое внимание нужно обратить на разработку нейрочипа с применением трехтерминального ферроэлектрического мемристора (рис. 24) [33] (в отличие от наиболее широко распространенных в настоящее время двухтерминальных мемристоров), в значительной степени идеологически повторяющего разработки мемристоров 60-х годов прошлого века [16, 17].

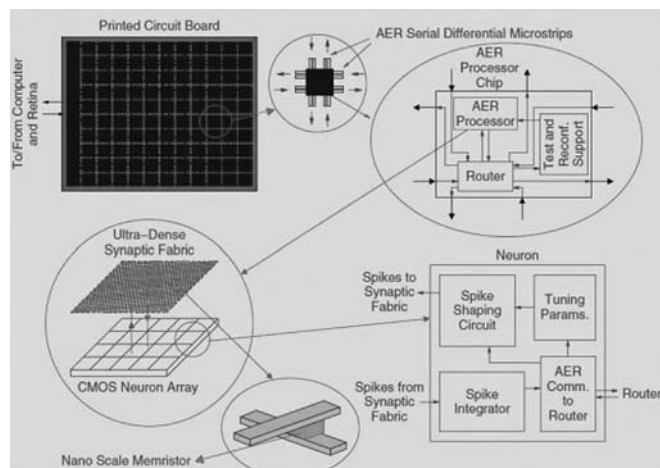


Рис. 21. Вариант архитектуры цифрового нейрочипа с применением мемристоров

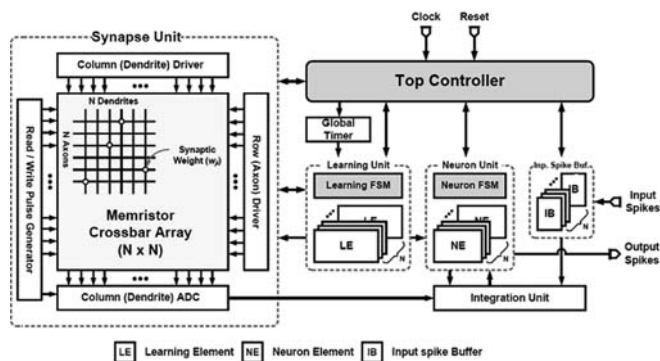


Рис. 22. Вариант архитектуры цифрового нейрочипа с применением мемристоров

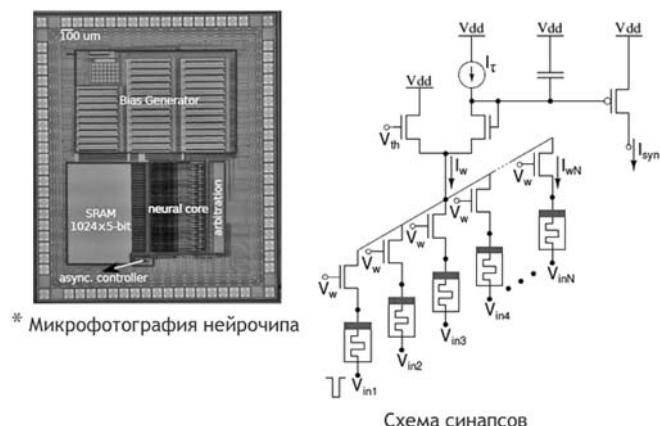


Рис. 23. Аналого-цифровой нейрочип с применением мемристоров

9. Архитектура вычислительных систем с применением мемристоров

На рис. 25 представлена общая архитектура перспективного супернейрокомпьютера с применением мемристоров. Классические варианты хост-ЭВМ и кластерного ядра дополняются в данной архитектуре двумя типами блоков:

- система на базе ПЛИС (БМК) с частотно-импульсным представлением сигналов как программно-аппаратный эмулятор нейронных сетей;
 - аналого-цифровая часть, реализованная на СБИС-нейрочипах с применением мемристоров.
- Результатами работ по разработке технических средств супернейрокомпьютера с применением мемристоров должны быть:
- архитектура и экспериментальные образцы каскадируемых СБИС на базе мемристорматриц для использования в перспективных супернейрокомпьютерах;
 - архитектура и экспериментальные образцы базовых каскадируемых плат в конструктивах PCI, microPCI, VME с использованием СБИС на базе мемристоров для использования в перспективных супернейрокомпьютерах;

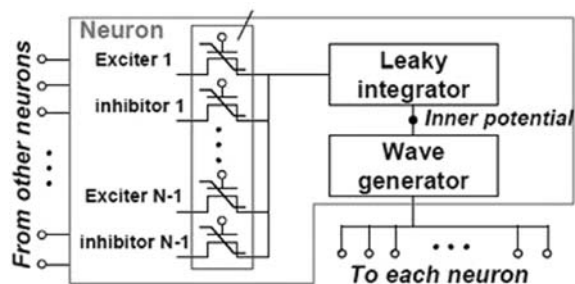
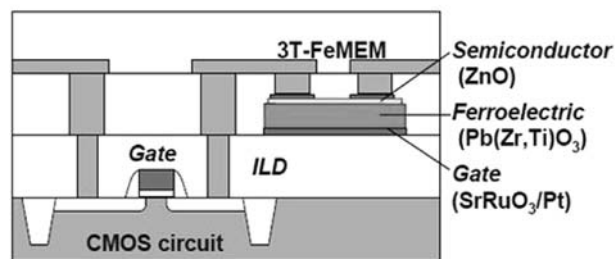


Схема нейрона



Интеграция 3Т-FeMeM и CMOS

Рис. 24. Нейрочип с применением трехтерминального ферроэлектрического мемристора



Рис. 25. Архитектура перспективного супернейрокомпьютера с использованием мемристоров

- архитектура и экспериментальные образцы базовых каскадируемых блоков с использованием плат в конструктивах PCI, microPCI, VME для использования в перспективных супернейрокомпьютерах;
- оформление патентов на архитектуры СБИС, плат и блоков на базе мемристорных матриц.

Структура системного программного обеспечения перспективного супернейрокомпьютера с применением мемристоров должна состоять из следующих разделов:

- общее системное программное обеспечение (ОС, компиляторы и пр.);
- алгоритмы и программы распараллеливания нейронных сетей различной структуры;
- вариант MPI;
- вариант системы типа Foundation для ПЛИС с применением мемристоров;
- алгоритмы и программы настройки нейронных сетей с учетом ограничений мемристоров.

Для разработанных вариантов архитектуры должна быть проведена общая оценка качества вычислительных систем с использованием мемристоров, в том числе:

- разработка и апробация методики количественной оценки производительности суперкомпьютеров на базе мемристоров;
- разработка методики обеспечения и оценки необходимой точности вычислений при переходе в супернейрокомпьютере к аналоговой обработке;
- оценка понижения энергопотребления в супернейрокомпьютере при переходе к представлению сигналов в виде частотно-модулированной последовательности узких импульсов.

10. Прикладные задачи

Текущее состояние интереса к решению прикладных задач на вычислительных системах с применением мемристоров условно можно представить схемой, изображенной выше на рис. 1.

Необходимо отметить соответствие структуры и параметров реализуемых нейронных сетей с применением мемристоров классу решаемых задач.

Структура нейронных сетей:

- многослойные нейронные сети;
- 2D и 3D клеточные нейронные сети;
- рекуррентные нейронные сети;
- хаотические нейронные сети.

Разрядность коэффициентов нейронных сетей:

- 1;
- K;
- управляемая разрядность;
- плавающая запятая.

Скорость изменения коэффициентов:

- нулевая;
- средняя;
- большая.

Большие нейронные сети — главная задача реализации нейрокомпьютеров с применением мемристоров. Как следствие — выбор для решения на мемристорных системах следующих сложных задач:

- обнаружение атак на информационные ресурсы в больших распределенных вычислительных сетях;
- нейрокриптография;
- уравнения математической физики, включая генерацию адаптивных сеток;
- обработка видеоизображений и медиаинформации (проблема *big data*);
- создание 3D-моделей мозга;
- управление плазмой;

- обработка геномной и протеомной информации;
- любые другие применения нейрокомпьютеров, где необходимо увеличить отношение производительности к стоимости или энергопотреблению.

Заключение

Всю историю развития вычислительной техники сопровождают работы по эмуляции алгоритмов искусственного интеллекта, принятия решений и нейронных сетей.

Российские работы по эмуляции нейронных сетей на каждом этапе развития вычислительной техники были связаны с использованием максимально производительных образцов:

- 60-е годы прошлого столетия — аналоговые реализации;
- 70—80 годы прошлого столетия — программные цифровые реализации;
- 90-е годы прошлого столетия — транспьютерные системы, цифровые, аналого-цифровые и аналоговые нейрочипы, ПЛИС;
- 2000-е годы — различные нейрочипы, графические процессоры;
- в настоящее время — различные нейрочипы, мемристорные системы.

Идеология построения нейрокомпьютеров менялась слабо, менялись и развивались технологии реализации.

Научно-технический задел по созданию отечественных супернейрокомпьютеров эксафлопной производительности с применением мемристоров представлен такими разделами.

- *Теория нейронных сетей* как методика синтеза структур из нейронов различного вида, алгоритмов адаптации весовых коэффициентов в этих структурах в процессе решения различных задач.
- *Нейроматематика* как раздел вычислительной математики, связанный с решением в нейросетевом логическом базисе различных сложных формализуемых и неформализуемых задач.
- *Нейроуправление* как раздел теории управления, связанный с применением нейрокомпьютеров в качестве систем идентификации сложных динамических систем и нейрокомпьютеров для управления.
- *Нейрокомпьютеры и нейрочипы*.

Российский опыт работ в области нейросетевых технологий и их применения обобщен в монографии [34].

Когда сейчас говорят: "мемристоры — будущее искусственного мозга", то это вызывает серьезные сомнения. На самом деле, мемристоры — это очередной этап эволюционного, гармоничного развития вычислительной техники, в котором за несколько предыдущих десятилетий проведена фундаментальная подготовка к построению высокопроизводительных вычислительных систем, реализующих следующие принципы:

- переход от классической фон-Неймановской архитектуры к распределенной;
- возврат в части реализации алгоритмов решения задач к аналоговой обработке в целях увеличения быстродействия при контролируемой точности;
- переход от представления сигналов в виде уровней токов и напряжений к представлению сигналов в виде частоты последовательности узких импульсов;
- переход к архитектурам, в которых функции памяти и обработки хотя бы частично совмещены в отличие от классических ЭВМ с полностью распределенными функциями памяти и обработки;
- переход к нейроматематике — нейросетевым алгоритмам решения задач;
- переход к нейроуправлению — нейросетевым алгоритмам и системам управления сложными динамическими объектами.

Модель мозга и его разделов — задачи искусственного интеллекта — важные, но вторичные задачи вычислительной техники, в том числе и мемристорных систем. Мемристорная система — это просто более эффективный эмулятор нейронных сетей по сравнению с предыдущими типами суперЭВМ. На рис. 26 показано развитие базовых вычислительных платформ для реализации эффективных программных и программно-аппаратных эмуляторов нейронных сетей.

Несмотря на то, что вычислительную технику за весь большой период существования и развития всегда использовали для реализации мыслительных функций, только в период 2007—2009 гг. и далее появились работы, в которых показано, что даже самые современные суперЭВМ типа Blue Gene, K-Fujitsu и других еще крайне ограничены по производительности для моделирования разделов мозга в реальном времени. Эти исследования показывают, как далека современная (и думаем, что даже перспективная на ближайшее время) технология от технологии реализации реального живого мозга, и как далеки мы от настоящего, глубокого понимания принципов работы живого мозга, которые мы сможем использовать при построении искусственных систем.

С этой точки зрения мемристоры — это не шаг к модели мозга, а небольшой шаг в развитии высокопроизводительной вычислительной техники, который позволит создать более эффективные эмуляторы нейронных сетей и более быстро и качественно решать инженерные задачи, следовательно, можно сделать следующие выводы:

— мемристоры — эволюционное развитие микроэлектроники в сторону нанотехнологий;

— нейрокompьютеры с применением мемристоров — эволюционное развитие высокопроизводительной техники.

Мемристорные системы имеют такое же далекое отношение к моделям мозга, как нейрокompьютеры 60-х гг. прошлого века, IBM Computer 70-х и т. д.

На рис. 27 условно представлено мнение автора о взаимоотношении разработок сверхвысокопроизводительной

техники, в том числе и супернейрокompьютеров, и нейрофизиологии. Исходя из данной схемы эти два направления соединяют четыре проблемы:

— эмуляция нейронных сетей для решения технических задач;

— интерфейс "мозг — компьютер";

— нейроимплантаты;

— искусственный мозг.

В части эмуляции нейронных сетей для решения технических задач с точки зрения автора роль нейрофизиологии для развития вычислительных систем практически нулевая. Высокопроизводительные вычислительные системы используют в нейрофизиологии для эмуляции больших нейросетевых образований. В настоящее время эта процедура, реализуемая на ЭВМ IBM Blue Gene и Fujitsu K, показывает, что возможности современных суперЭВМ для моделирования разделов мозга крайне незначительны.

В части интерфейса "мозг — компьютер" нейрофизиология может оказать значительную помощь вычислительной технике в части новых методов управления вычислительными системами.

В части нейроимплантатов вычислительные системы являются основой их технической реализации.

Проблема "Искусственный мозг" является в настоящее время, с точки зрения автора, чисто искусственной, так как технологии реализации вычислительных систем резко отличаются от технологий реализации живого мозга и переносимость результатов с одного на другое является чисто условной.

Необходимо отметить основные перспективные направления работ по созданию нейрокompьютеров с применением мемристоров с ориентацией на большие нейронные сети:

- выбор технологии с точки зрения обеспечения максимальной производительности;
- разработка вариантов архитектуры и схмотехники СБИС-нейрочипов с применением мемристоров;
- разработка вариантов архитектуры и схмотехники нейроплат, нейроблоков и нейростоек нейрокompьютеров с применением мемристоров;
- разработка алгоритмов адаптации нейронных сетей, адекватных архитектурам с применением мемристоров и с учетом моделей мемристоров;
- разработка нейросетевых алгоритмов решения задач, адекватных архитектурам с применением мемристоров;
- разработка алгоритмов распараллеливания нейросетевых структур, адекватных архитектурам с применением мемристоров;
- разработка методики оценки производительности, адекватной архитектуре с применением мемристоров.

Список литературы

1. Бёрд Киви. К точке критического перехода. 3D-News (Daily Digital Digest). 29.05.2013.
2. McKenzie A., Branch D. W., Forsythe C., James C. D. Toward Exascale Computing through Neuromorphic Approaches // Sandia Report, Sand 2010—6312. September. 2010.
3. Галушкин А. И. Стратегия развития современных супернейрокompьютеров на пути к эксафлопным вычислениям // Информационные технологии. Приложение. 2012. № 3. 32 с.
4. Галушкин А. И. Многослойные системы распознавания образов. М.: Изд. МИЭМ, 1970.
5. Галушкин А. И. Синтез многослойных систем распознавания образов. М.: Энергия, 1974.
6. Галушкин А. И. Теория нейронных сетей. Серия "Нейрокompьютеры и их применение". Кн. 1. М.: ИПРЖР. 2000, 416 с.
7. Галушкин А. И., Цыпкин Я. З. Нейронные сети: история развития теории. Серия "Нейрокompьютеры и их применение". Кн. 5. М.: ИПРЖР. 2001. 840 с.
8. Галушкин А. И. Теория нейронных сетей. Пекин: Изд-во Университета Синьхуа (на китайском языке). 2003.

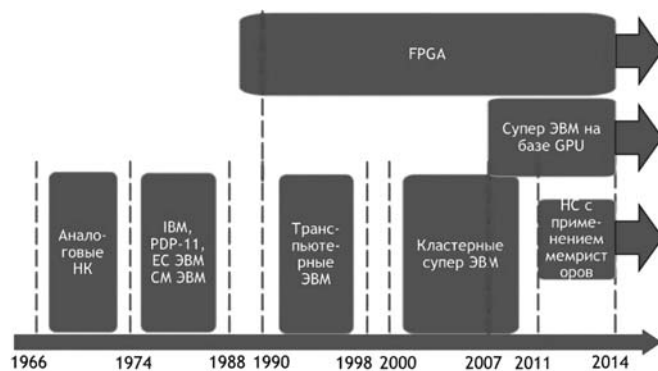


Рис. 26. Создание наиболее эффективных программно-аппаратных эмуляторов нейронных сетей — главная задача нейросетевых технологий



Рис. 27. Структура взаимоотношений высокопроизводительной техники и нейрофизиологии

9. Galushkin A. I. Neural Network Theory. Berlin-Heidelberg: Springer, 2007.
10. Галушкин А. И. Нейронные сети: основы теории. М.: Горячая линия — Телеком, 2010.
11. Нейроматематика. Серия "Нейрокомпьютеры и их применение". Кн. 6. М.: ИПРЖР, 2002.
12. Галушкин А. И. Нейроматематика (проблемы развития) // Нейрокомпьютер. 2003. № 1.
13. Галушкин А. И. Методика решения задач в нейросетевом логическом базисе. Информационные технологии. Приложение. 2006. № 6. 32 с.
14. Терехов В. А., Ефимов Д. В., Тюкин И. Ю. Нейросетевые системы управления. Серия "Нейрокомпьютеры и их применение". Кн. 8. М.: ИПРЖР, 2004.
15. Галушкин А. И. Основы нейрорупления. Информационные технологии. Приложение. 2002. № 10. 32 с.
16. Аналоговые запоминающие и адаптивные элементы / Под ред. Б. С. Сотскова. М.: Энергия, 1973.
17. Трейер В. В., Елизаров А. Б. Электрохимические интегрирующие и аналоговые запоминающие элементы. М.: Энергия, 1973.
18. Галушкин А. И. Нейрокомпьютеры. Серия "Нейрокомпьютеры и их применение". Кн. 1. М.: Радиотехника, 2000.
19. Нейрокомпьютеры: от программной к аппаратной реализации. М.: Горячая линия — Телеком, 2006.
20. Shin S., Kim K., Kang S. M. Memristors and Their Applications for Nanocomputing // IEEE Trans. on Nanotechnology, Mar. 2011.
21. Narika Manem, Jeyavijayan Rajendran, Garrett S. Rose. Stochastic Gradient Descent Inspired Training Technique for a CMOS / Nano Memristive Trainable Threshold Gate Array // IEEE Transactions on circuits and systems — I: regular papers. V. 59, N. 5, May 2012.
22. Adhikari S. P., Yang C., Kim H., Chua L. O. Memristor Bridge Synapse-Based Neural Network and Its Learning // IEEE Transactions on neural network and learning systems. September 2013. V. 23, N. 9.
23. Pd. Sah M., Yang C., Hyongsuk Kim, Leon O. Chua. Memristor Circuit for Artificial Synaptic Weighting of Pulse Inputs // Circuits and Systems (ISCAS). IEEE International Symposium. 2012.
24. Maheshwar Pd. Sah, Changju Yang, Ram Kaji Budhathoki, Hyongsuk Kim. Features of memristor emulator-based artificial neural synapses // Circuits and Systems (ISCAS), IEEE International Symposium. 2012.
25. Young-Su Kim, Keong-Sik Mim. Synaptic weighting circuits for Cellular Neural Networks // Cellular Nanoscale Networks and Their Applications (CNNA). 13th International Workshop. 2012.
26. Walls T. J., Likharev K. K. Self-Organization in Autonomous, Recurrent, Firing-Rate CrossNets With Quasi-Hebbian Plasticity // IEEE Transactions on Neural Networks and Learning Systems / April. V. 25. N. 4. 2014.
27. Hebb D. O. The organization of Behavior: A Neuropsychological Theory. New York: Wiley, June 1949.
28. Gorchetchnikov A., Versace M., Ames H., Chandler B., Leveille J., Livitz G., Mingolla E., Snider G., Amerson R., Carter D., Abdalla H., Qureshi M. S. Review and unification of learning framework in Cog Ex Machina platform for memristive neuromorphic hardware // Proc. of International Joint Conference on Neural Networks, San Jose, California, USA, July 31—August 5, 2011.
29. Seo Jae-sun, Brezzo B., Liu Y., Parker B. D., Esser S. K., Montoya R. K., Rajendran B., Tierno J. A., Chang L., Modha D. S., Friedman D. J. A 45 nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons // Custom Integrated Circuits Conference (CICC), 2011.
30. Serrano-Gotarredona T., Prodromakis T., Linares-Barranco B. A proposal for hybrid memristor-CMOS spiking neuromorphic learning systems // IEEE Circuits and Systems Magazine. 2013. Vol. 13, Is. 2. P. 74—88.
31. Yongtae Kim, Yong Zhang, Peng Li. A digital neuromorphic VLSI architecture with memristor crossbar synaptic array for machine learning // IEEE International SOC Conference (SOCC), 2012.
32. Azghadi M. R., Moradi S., Indiveri G. Programmable neuromorphic circuits for spike-based neural dynamics // IEEE 11th International New Circuits and Systems Conference (NEWCAS), 2013.
33. Kaneko Y., Nishitani Yu., Ueda M., Tsujimura A. Neural network based on a three-terminal ferroelectric memristor to enable on-chip pattern recognition // Symposium VLSI Technology (VLSIT). 2013.
34. Галушкин А. И., Симоров С. Н. Нейросетевые технологии в России (1982—2010 г.). М.: Горячая линия — Телеком, 2011.

A. I. Galushkin, Prof., Deputy Head of Chair,
Moscow Institute of Physics and Technology, Dolgoprudny, neurocomputer@yandex.ru

Memristor in High-Performance Computing Development

The author of this article is not a physicist or a mathematician. He is an engineer and has 50 years of experience in neural network technologies and their application. This is a rather narrow class of over high performance computing. This article presents the author's opinion about the prospects of this trend of super computing with the memristor invention.

Keywords: memristor, supercomputer, multilayer, neural network, analog computation

References

- Kivi B. K tochke kriticheskogo perehoda. *3D-News (Daily Digital Digest)*. 29.05.2013 g.
- McKenzie A., Branch D. W., Forsythe C., James C. D. Toward Exascale Computing through Neuromorphic Approaches // *Sandia Report, Sand 2010 — 6312*. September 2010.
- Galushkin A. I. Strategija razvitiya sovremennykh supernejrokomputеров na puti k jekzaflopnyim vychislenijam. *Informacionnye tehnologii*. Prilozhenie. 2012. N. 3. 32 p.
- Galushkin A. I. *Mnogoslojnye sistemy raspoznavanija obrazov*. M.: MIJeM, 1970.
- Galushkin A. I. *Sintez mnogoslojnykh sistem raspoznavanija obrazov*. M.: Jenergija, 1974.
- Galushkin A. I. *Teorija nejronnyh setej*. Kn. 1. Serii "Nejrokomputery i ih primenenie". M.: IPRJR, 2000.
- Galushkin A. I., Cypkin Ja. Z. *Nejronnye seti: istorija razvitiya teorii*. Kn. 5. Seriy "Nejrokomputery i ih primenenie". M.: IPRJR, 2001, 840 c.
- Galushkin A. I. *Teorija nejronnyh setej*. Pekin: Izdatel'stvo Universiteta Sin'hua. (na kitajskom jazyke). 2003.
- Galushkin A. I. *Neural Network Theory*. Berlin Heidelberg: Springer. 2007.
- Galushkin A. I. *Nejronnye seti: osnovy teorii*. M.: Gorjachaja linija — Telekom, 2010.
- Nejromatematika. Kn. 6 Seriya "Nejrokomputery i ih primenenie". M.: IPRJR, 2002.
- Galushkin A. I. *Nejromatematika (problemy razvitiya)*. *Nejrokomputer*. 2003. N. 1.
- Galushkin A. I. Metodika reshenija zadach v nejrosetevom logicheskom bazise. *Informacionnye tehnologii*. Prilozhenie. 2006. N. 6. 32 p.
- Terehov V. A., Efimov D. V., Tyukin I. Yu. *Nejrosetevye sistemy upravlenija*. Kn. 8. Seriya "Nejrokomputery i ih primenenie". M.: IPRJR, 2004. 32 p.
- Galushkin A. I. Osnovy nejroupravljenija. *Informacionnye tehnologii*. Prilozhenie. 2002. N. 10.
- Analogovye zapominajushhie i adaptivnye jelementy / pod red. B. S. Sotskova. M.: Jenergija, 1973.

17. Trejer V. V., Elizarov A. B. *Jelektrohimičeskie integrirujuščie i analogovye zapominajuščie jelementy*. M.: Jenergija, 1973.
18. Galushkin A. I. *Nejrokomп'jutyry*. Kn. 1. Seriya "Nejrokomп'jutyry i ih primenenie". M.: IPRJR, 2000.
19. *Nejrokomп'jutyry*: ot programmnoj k apparatnoj realizacii. M.: Gorjachaja linija — Telekom, 2006.
20. Shin S., Kim K., Kang S. M. Memristors and Their Applications for Nanocomputing. *IEEE Trans. on Nanotechnology*, Mar. 2011.
21. Manem H., Rajendran J., Garrett S. Rose Stochastic Gradient Descent Inspired Training Technique for a CMOS / Nano Memristive Trainable Threshold Gate Array. *IEEE Transactions on circuits and systems — I: regular papers*. May 2012. V. 59, N. 5.
22. Adhikari S. P., Yang C., Kim H. Memristor Bridge Synapse-Based Neural Network and Its Learning. *IEEE Transactions on neural networks and learning systems*. 2012. V. 23, N. 9.
23. Maheshwar Pd. Sah, Changju Yang, Hyongsuk Kim and Leon O Chua. Meristor Circuit for Artificial Synaptic Weighting of Pulse Inputs. *Circuits and Systems (ISCAS)*, 2012, IEEE International Symposium.
24. Maheshwar Pd. Sah, Changju Yang, Ram Kaji Budhathoki, Hyongsuk Kim. Features of memristor emulator-based artificial neural synapses. *Circuits and Systems (ISCAS)*, 2012 IEEE International Symposium.
25. Young-Su Kim, Kyeong-Sik Min. Synaptic weighting circuits for Cellular Neural Networks. *Cellular Nanoscale Networks and Their Applications (CNNA)*, 2012 13th International Workshop.
26. Walls T. J., Likharev K. K. Self-Organization in Autonomous, Recurrent, Firing-Rate CrossNets With Quasi-Hebbian Plasticity. *IEEE Transactions on Neural Networks and Learning Systems*. 2014. V. 25, N. 4.
27. Hebb D. O. *The organization of Behavior: A Neuropsychological Theory*. New York: Wiley, June 1949.
28. Gorchetchnikov A., Versace M., Ames H., Chandler B., Leveille J., Livitz G., Mingolla E., Snider G., Amerson R., Carter P., Abdalla H., Qureshi M. S. Review and unification of learning framework in Cog Ex Machina platform for memristive neuromorphic hardware // *Proc. of International Joint Conference on Neural Networks*. San Jose, California, USA, July 31—August 5, 2011.
29. Jae-sun Seo et al. A 45 nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons. *IEEE. Custom Integrated Circuits Conference (CICC)*, 2011.
30. Serrano-Gotarredona T., Prodromakis T., Linares-Barranco B. A Proposal for Hybrid Memristor-CMOS Spiking Neuromorphic Learning Systems. *IEEE Circuits and Systems Magazine*. V. 13, Is. 2.
31. Yongtae Kim, Yong Zhang, Peng Li. A digital neuromorphic VLSI architecture with memristor crossbar synaptic array for machine learning. *IEEE International SOC Conference (SOCC)*, 2012.
32. Azghadi M. R., Moradi S., Indiveri G. Programmable neuromorphic circuits for spike-based neural dynamics. *IEEE 11th International New Circuits and Systems Conference (NEWCAS)*, 2013.
33. Kaneko Y., Nishitani Yu., Ueda M., Tsujimura A. Neural network based on a three-terminal ferroelectric memristor to enable on-chip pattern recognition. *Symposium VLSI Technology rVLSIT*. 2013.
34. Galushkin A. I., Simorov S. N. *Nejrosetevye tehnologii v Rossii (1982—2010 g.)*. M.: Gorjachaja linija — Telekom, M. 2011.

УДК 519.856.2

В. В. Федосов, канд. техн. наук, доц., г. Москва, e-mail: vlr.fdsv@gmail.com,
А. В. Федосова, канд. физ.-мат. наук,
 Национальный университет, Богота, Колумбия
 (Universidad Nacional de Colombia, e-mail: afedosova@unal.edu.co)

Использование нейросетей для графической оценки загрязнения территории выбросами группы источников

Взамен математического описания выбросов загрязнений источниками предложено использовать нейрофункции, генерируемые по данным рисованных контурных графиков облаков выбросов.

Нейрофункции эффективно выводят графику выборочного или общего загрязнения территории и пригодны для дальнейших расчетов или оптимизации.

Ключевые слова: источники промышленных выбросов, облака выбросов, загрязнение территории, контурные графики, нейронные сети, фильтрация выбросов

Введение

Промышленная экология рассматривает загрязнения территории или среды источниками вредных выбросов (например промпредприятиями) как сложные системы, содержащие параметры источников (мощности, размещение, дальность разброса), выбросов (типы, масса и размеры частиц, наборы компонентов или состав), среды (ветер, влажность, температура), динамику и пр. Выбросы в значительной степени формируются за счет пылевых, капельных или газовых частиц. Несмотря на обширные базы данных и трудоемкость создания таких систем, конечные результаты расчетов загрязнений территорий следует признать только оценочными.

При моделировании ключевым является функциональное описание выбросов источниками. Его концепция, математический аппарат, введение усложнений (упрощений) определяются целями задачи и квалификацией исследователя.

Последующие расчеты облаков выбросов ведут к построению соответствующих карт загрязнения, обсуждению адекватности моделирования и практической ценности результатов. Такой подход достаточно апробирован и ценен уже тем, что выявляет неясные заранее наложения загрязнений выбросами группы источников. Он применим как для прямых, так и для оптимизационных расчетов загрязнений.

Между тем авторская математика часто субъективна, достаточно трудоемка и ограничена в вариативности описания облаков выбросов. Формулы содержат значительное число условных коэффициентов. При наличии в модели множества источников выбросов их работа описывается единообразно, хотя на практике облака даже близко расположенных