

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

заместителя генерального директора по инновационной деятельности, ООО
«Альфачип», доктора технических наук, члена-корреспондента РАН

Романа Александровича Соловьева

на диссертационную работу Ляхова Павла Алексеевича
«МЕТОДЫ ВЫСОКОПРОИЗВОДИТЕЛЬНОЙ ЦИФРОВОЙ ОБРАБОТКИ
СИГНАЛОВ НА ОСНОВЕ ВЫЧИСЛЕНИЙ В СИСТЕМЕ ОСТАТОЧНЫХ
КЛАССОВ»,

представленную на соискание ученой степени доктора технических наук по
специальности 2.3.5 – Математическое и программное обеспечение
вычислительных систем, комплексов и компьютерных сетей (технические
науки).

1.Актуальность избранной темы

Современный этап развития информационных технологий и систем искусственного интеллекта характеризуется экспоненциальным ростом объёмов обрабатываемых данных и требований к скорости вычислений в режиме, близком к реальному времени. Телемедицина, системы компьютерного зрения, беспилотный транспорт, спутниковый мониторинг и робототехнические комплексы предъявляют к вычислительным системам требования, которые традиционные позиционные системы счисления (ПСС), лежащие в основе классических микропроцессорных архитектур, удовлетворяют всё в меньшей степени. Ключевым ограничением является последовательный поразрядный перенос, формирующий критический путь вычислений и увеличивающий задержки и энергопотребление. Приближение кремниевых технологий к физическим пределам миниатюризации делает поиск альтернативных вычислительных парадигм объективно необходимым.

Система остаточных классов (СОК) обеспечивает естественное распараллеливание модульных операций сложения, вычитания и умножения на арифметико-логическом уровне, поскольку эти операции выполняются независимо над малоразрядными остатками без межразрядных переносов. Однако широкому практическому внедрению СОК в задачи цифровой обработки сигналов (ЦОС) препятствует фундаментальное противоречие: реализация немодульных процедур, таких как обратное преобразование, определение знака, сравнение чисел и коррекция ошибок, требует привлечения информации из всех вычислительных каналов, что разрушает параллелизм и многократно увеличивает аппаратные затраты. Кроме того, до настоящего времени отсутствовала комплексная методология, объединяющая модулярную арифметику с блочной обработкой данных, методом Винограда и

целочисленным масштабированием коэффициентов фильтров в единую технологию.

Разрешение данного противоречия требует не точечной оптимизации отдельных узлов, а разработки научно обоснованных моделей, архитектур и методов, позволяющих нивелировать недостатки СОК и многократно усилить её достоинства. В связи с этим тема диссертации Павла Алексеевича Ляхова обладает безусловной актуальностью как для развития теории модулярных вычислений, так и для обеспечения технологического суверенитета Российской Федерации в области создания высокопроизводительных вычислительных систем.

2. Содержание диссертационной работы

Диссертационная работа объёмом 433 страницы состоит из введения, шести глав, заключения, списка литературы из 427 наименований и трёх приложений. Работа иллюстрирована 116 рисунками и 54 таблицами.

В первой главе проведён критический анализ подходов к улучшению эксплуатационных характеристик систем ЦОС: приближённых вычислений, модификации арифметических блоков, блочной обработки данных, альтернативных систем счисления и проектирования аппаратных ускорителей. Выявлено и формализовано противоречие между скоростью модульных операций и ресурсоёмкостью немодульных процедур в СОК. Сформулирована задача исследования, введены ограничения и предложен многокритериальный функционал оптимальности.

Во второй главе разработана обобщенная модель усовершенствованной ЦОС в СОК и архитектура цифрового фильтра с модулями специального вида. Проведён теоретический анализ ресурсозатрат арифметических устройств на основе unit-gate модели (UGM). Разработаны конструкции усечённых блоков умножения с накоплением, передающих векторы суммы и переноса на следующее звено фильтра без промежуточного суммирования.

В третьей главе предложен метод высокоскоростной вейвлет-фильтрации на основе блочной обработки данных методом Винограда. Разработана методика оценки производительности. Метод обобщён на случай фильтрации с понижающей дискретизацией при произвольном шаге свёртки.

В четвёртой главе решена проблема ресурсоёмкости немодульных операций. Разработаны методы и архитектуры обратного преобразования, определения знака, сравнения чисел и коррекции ошибок на основе модифицированной Китайской теоремы об остатках (КТО) с дробными величинами (КТОд). Доказаны теоремы, определяющие минимально необходимую разрядность дробных бит для гарантированно точных вычислений.

В пятой главе представлен алгоритм масштабирования коэффициентов вейвлет-фильтров, учитывающий накопительную погрешность многоуровневого дискретного вейвлет-преобразования (ДВП). Разработан

метод параллельной вейвлет-обработки многомерных сигналов в СОК. Программная верификация выполнена на реальных трёхмерных медицинских томограммах с гарантированным достижением заданного порога пикового отношения сигнала к шуму (ПОСШ).

В шестой главе описана технология сквозного проектирования программно-аппаратных систем ЦОС в СОК. Представлены результаты моделирования на программируемых пользователем вентильных матрицах (ППВМ) и интегральных схемах специального назначения (ИССН). Описаны программно-аппаратная реализация свёрточной нейронной сети (СНС) с вычислениями в СОК и отказоустойчивого модулярного нейрокомпьютера на основе сети Хопфилда.

Структура диссертации логична: от анализа и постановки задачи в первой главе к разработке базовых моделей и методов ускорения во второй и третьей, к преодолению фундаментальных ограничений СОК в четвёртой, к адаптации для многомерных сигналов в пятой и к созданию сквозной технологии проектирования в шестой.

3. Полученные основные научные результаты

К основным научным результатам, выносимым на защиту, относятся:

1. Обобщенная модель усовершенствованной ЦОС в СОК, системно объединяющая параллельную модулярную арифметику, блочную обработку данных, немодульные операции на базе КТОд и целочисленное масштабирование коэффициентов фильтров. Модель обеспечивает снижение аппаратных затрат на немодульные операции в 1,5–3 раза по сравнению с методами на базе обобщённой ПСС и полной КТО.

2. Архитектура цифрового фильтра в СОК с модулями специального вида и усечёнными блоками умножения с накоплением. Передача векторов суммы и переноса на следующее звено фильтра позволяет использовать лишь один параллельно-префиксный сумматор в конце цепочки, сокращая время фильтрации одномерных сигналов на 22-29% и снижая аппаратные затраты в 1,7-4,8 раза по сравнению с ПСС.

3. Метод высокоскоростной вейвлет-фильтрации на основе блочной обработки данных, в котором ресурсоёмкие операции умножения заменяются комбинацией сдвигов и сложений за счёт априорного формирования вспомогательных матриц преобразования. Метод впервые обобщён на случай фильтрации с понижающей дискретизацией при произвольном шаге свёртки. Вычислительная сложность обработки двумерных сигналов снижается до 72,9%, скорость возрастает до 73,62%.

4. Методы и архитектуры для реализации немодульных операций в СОК на базе КТОд, обеспечивающие переход от линейной к логарифмической вычислительной сложности. Экспериментально подтверждено снижение аппаратных затрат до 7,07 раз при сравнении чисел и повышение

производительности до 27,14 раз при определении знака по сравнению с аналогами на базе классической КТО и обобщённой ПСС.

5. Алгоритм масштабирования коэффициентов вейвлет-фильтров и метод параллельной вейвлет-обработки трёхмерных сигналов в СОК. Алгоритм автоматически подбирает минимально достаточную разрядность коэффициентов, гарантируя достижение заданного порога ПОСШ. Распараллеливание по каналам СОК повышает производительность ДВП в 2,89-3,59 раза.

6. Технология сквозного проектирования программно-аппаратных систем ЦОС в СОК, включающая ядро математического моделирования, генератор параметризуемых конфигураций и фреймворк верификации. Практическая эффективность подтверждена ростом производительности нейросетевой обработки на 58,89% и снижением задержки коррекции ошибок в отказоустойчивом нейрокомпьютере в 1,2-3,8 раза.

4. Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации

Каждое из шести выносимых на защиту положений подкреплено как теоретическим обоснованием, так и экспериментальной верификацией. Теоретический фундамент работы составляют строгие доказательства на основе аппарата теории чисел и модулярной арифметики. Соискателем доказаны теоремы 4.1 и 4.2, определяющие минимально необходимую разрядность дробных бит для гарантированно точного обратного преобразования и определения знака числа, что исключает появление зон неопределённости. Обоснован переход от классической КТО к модифицированному методу КТОд. Выведены аналитические зависимости для выбора оптимальных точек интерполяции при блочной обработке данных и для расчёта накопительной погрешности при многоуровневом ДВП. Теоретический анализ ресурсозатрат выполнен на основе UGM для сумматоров Когге-Стоуна, усечённых блоков умножения с накоплением и многоходовых сумматоров в зависимости от вида модулей СОК.

Теоретические выкладки верифицированы масштабным аппаратным моделированием. Синтез и верификация разработанных архитектур проведены в системах автоматизированного проектирования (САПР) Xilinx Vivado, Xilinx ISE, Yosys, OpenLane и Synopsys Design Compiler на широком спектре платформ: ППВМ семейств Artix-7, Spartan-6, Virtex-6, Virtex-7, Kintex-7, Kintex UltraScale, а также на ИССН с использованием библиотеки стандартных ячеек. Результаты моделирования подтвердили адекватность теоретических оценок реальным временным и аппаратным затратам.

Дополнительную обоснованность практическим рекомендациям придаёт программное моделирование в среде MATLAB на наборах реальных трёхмерных медицинских томограмм с использованием специализированных пакетов просмотра и анализа изображений. Экспериментально подтверждено

отсутствие видимых искажений и гарантированное достижение пороговых значений ПОСШ при использовании предложенного алгоритма масштабирования.

Выводы и рекомендации диссертанта логически вытекают из полученных результатов и подтверждены актами о внедрении, патентами и свидетельствами о государственной регистрации программ для ЭВМ. Степень обоснованности не вызывает сомнений.

5. Достоверность и новизна

Достоверность результатов обеспечивается полным совпадением аналитических оценок, полученных на основе UGM, с данными аппаратного моделирования на семи типах вычислительных платформ в пяти использованных САПР. Программные эксперименты на реальных медицинских данных дополнительно подтверждают практическую применимость разработанных методов. Публикационная активность автора свидетельствует об апробации результатов научным сообществом: 129 работ, 836 ссылок в Web of Science, 1250 ссылок в Scopus.

Научная новизна состоит в разрешении фундаментального противоречия между высоким быстродействием модульных операций и экстремальной вычислительной сложностью немодульных процедур в СОК. Впервые предложена обобщенная модель, системно объединяющая четыре взаимодополняющих механизма ускорения. Впервые метод блочной обработки данных обобщён на случай вейвлет-фильтрации с понижающей дискретизацией при произвольном шаге свёртки. Впервые доказана возможность реализации немодульных операций с логарифмической сложностью на основе аппарата дробных величин. Впервые разработан алгоритм автоматического подбора параметра масштабирования коэффициентов вейвлет-фильтров с гарантированным качеством восстановления для трёхмерных сигналов. Впервые в рамках единой технологии предложена архитектура отказоустойчивого модулярного нейрокомпьютера, в котором для коррекции ошибок применена сеть Хопфилда в сочетании с быстрым блоком обнаружения ошибок на базе КТОд. Совокупность указанных результатов образует законченное научное достижение, не имеющее аналогов по полноте и системности.

6. Замечания по диссертационной работе

Диссертация выполнена на высоком научном уровне и содержит значимые результаты, однако считаю необходимым высказать ряд замечаний и пожеланий, не снижающих общей положительной оценки работы.

- 1) Во второй и шестой главах диссертант оценивает задержки и аппаратные затраты с использованием модели unit-gate и средств САПР, такие как Xilinx Vivado и Synopsys Design Compiler. Однако при разработке

современных топологий сверхбольших интегральных схем задержки межсоединений часто начинают преобладать над задержками логических вентилей. В работе не в полной мере раскрыто, как трассировка и физические ограничения размещения блоков при параллельной архитектуре каналов системы остаточных классов влияют на итоговые характеристики статического временного анализа на современных технологических нормах.

- 2) В разделе 6.5 рассматривается аппаратно-программная реализация нейросетевых вычислений с параллельной обработкой данных в СОК. Учитывая жесткие ограничения по памяти в граничных вычислениях (edge computing), целесообразно было бы провести сравнение предложенного алгоритма масштабирования коэффициентов фильтров (глава 5) с современными методами квантования нейронных сетей до целочисленных форматов пониженной разрядности, например, INT8 или INT4. Требуется уточнения, как накопительная погрешность многократной последовательной свертки в СОК влияет на точность глубоких нейросетевых архитектур.
- 3) В четвертой главе предложены методы реализации немодульных операций на основе дробных величин, которые существенно снижают их вычислительную сложность по сравнению с классической Китайской теоремой об остатках. Тем не менее, при сквозном проектировании сложных систем цифровой обработки сигналов, например, обработки высококадрового потокового видео, частое использование блоков прямого ПСС в СОК и обратного СОК в ПСС преобразования может создать эффект аппаратного «узкого горлышка». В работе хотелось бы видеть более подробный анализ доли энергопотребления и занимаемой площади этих преобразователей в общем бюджете кристалла, смоделированного с использованием библиотеки стандартных ячеек.
- 4) В третьей главе автор предлагает использовать метод Винограда для высокоскоростной фильтрации двумерных сигналов с целью уменьшения числа вычислительно сложных умножений. Хотя этот метод математически обоснован и снижает арифметическую сложность, аппаратная реализация подобных алгоритмических трансформаций в ускорителях сверточных нейронных сетей обычно приводит к значительному увеличению нагрузки на внутреннюю регистровую память и расширению шины данных. В диссертации недостаточно проанализировано влияние предложенного блочного метода на пропускную способность подсистемы памяти устройства.

Перечисленные замечания носят дискуссионный характер и не влияют на общую высокую оценку диссертационной работы.

7. Заключение о соответствии диссертации критериям, установленным Положением о присуждении учёных степеней (пп. 9-14)

Диссертационная работа Ляхова Павла Алексеевича «Методы высокопроизводительной цифровой обработки сигналов на основе вычислений в системе остаточных классов» представляет собой законченную, самостоятельно выполненную научно-квалификационную работу, обладающую внутренним единством, в которой решена крупная научная проблема, имеющая важное теоретическое и хозяйственное значение: разработана комплексная технология и научно обоснованные математические модели параллельной обработки данных в СОК для создания высокопроизводительных систем ЦОС.

По своей актуальности, научной новизне, теоретической и практической значимости, обоснованности и достоверности полученных результатов, а также полноте опубликования основных научных результатов в рецензируемых изданиях диссертация полностью соответствует критериям пп. 9-14 «Положения о присуждении ученых степеней», утверждённого Постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 в действующей редакции, предъявляемым к докторским диссертациям.

Содержание диссертации и выносимые на защиту результаты соответствуют паспорту научной специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» по пунктам 1, 7 и 8.

На основании изложенного считаю, что Ляхов Павел Алексеевич заслуживает присуждения учёной степени доктора технических наук по специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей».

Официальный оппонент заместитель генерального директора по инновационной деятельности, ООО «Альфачип», доктор технических наук, член-корреспондент РАН

Р. А. Соловьев

«07» сентября 2026 г.

Контакты:

Тел: +7 (495) 657-98-80

E-mail: contact@alphachip.ru.

Адрес сайта организации: <https://alphachip.ru>.

Адрес: 124498, г. Москва, Зеленоград, Центральный проспект, корп. 438А.

Подпись официального оппонента Соловьева Романа Александровича, заместителя генерального директора по инновационной деятельности, ООО «Альфачип», доктора технических наук, члена-корреспондента РАН

ЗАВЕРЯЮ»