На правах рукописи



ВУ Хоанг Шон

АЛГОРИТМЫ МНОГОСКОРОСТНОЙ ОБРАБОТКИ РЕЧЕВЫХ СИГНАЛОВ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

Специальность: 2.2.13. Радиотехника,
в том числе системы и устройства телевидения

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Рязань 2026

Работа выполнена на кафедре радиоуправления и связи Федерального государственного бюджетного образовательного учреждения высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина».

Научный руководитель: **Дмитриев Владимир Тимурович,** доктор технических наук, доцент, заведующий кафедрой радиоуправления и связи ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина», г. Рязань.

Официальные оппоненты: **Якименко Игорь Владимирович,** доктор технических наук, профессор, заведующий кафедрой электроники и микропроцессорной техники филиала ФГБОУ ВО «Национальный исследовательский университет «МЭИ» в г. Смоленске, г. Смоленск;

Чучупал Владимир Яковлевич, кандидат физико-математических наук, ведущий научный сотрудник ФГУ «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва.

Ведущая организация: Федеральное государственное казённое военное образовательное учреждение высшего образования «Академия Федеральной службы охраны Российской Федерации»

Защита состоится «18» декабря 2026 г. в 12 часов 00 минут на заседании диссертационного совета 24.2.375.03 в ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина» по адресу: 390005, г. Рязань, ул. Гагарина, 59/1.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина» или на сайте http://www.rsreu.ru.

Автореферат разослан «____» _____ 2026 г.

Учёный секретарь
диссертационного совета
доктор технических наук, доцент Г.В. Овечкин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Современные требования к системам обработки и передачи речевых сигналов (РС), особенно при ограниченных ресурсах канала связи, требуют разработки эффективных методов многоскоростной обработки и кодирования. Классическая теорема дискретизации В.А. Котельникова является основой цифровой обработки сигналов, однако её применимость ограничена стационарными сигналами, в то время как речь — нестационарный сигнал со спектральными характеристиками, динамически изменяющимися во времени. Это стимулирует развитие обобщённых подходов к дискретизации. Практическая реализация теоремы В.А. Котельникова сопряжена с рядом ошибок: усечением ряда отсчётов, наложением спектров из-за нарушения условия ограниченности спектра, джиттером временных интервалов, а также ошибками округления и амплитудными погрешностями. Эти недостатки особенно критичны для речи с резкими фонемными переходами и нестационарной интенсивностью. В таких условиях классическая дискретизация оказывается неадекватной, что обусловливает применение обобщённых методов, в частности — многоканального представления Хургина – Яковлева, реализующего принципы многоскоростной фильтрации на основе кратной децимации и интерполяции. Оно использует банки фильтров для совместной обработки отсчётов сигнала и его производных, обеспечивая более точное восстановление нестационарных особенностей речи. Однако исходная версия представления не гарантирует точного восстановления из-за искажений фазового спектра синтезирующего фильтра, что требует его модификации путём коррекции фазочастотной характеристики (ФЧХ).

К этому же классу методов многоскоростной обработки относится дискретное вейвлет-преобразование (ДВП), обладающее хорошей локализацией во времени и частоте, что делает его эффективным для анализа и сжатия речи. Пороговая обработка вейвлет-коэффициентов применяется для подавления шумов и сжатия, однако выбор порогов часто основан на универсальных эмпирических правилах, не учитывающих статистику речи и условия передачи. Это ведёт либо к потере речевой информации, либо к недостаточному подавлению шумов. Решение возможно за счёт глобальной оптимизации, в частности — алгоритма роя частиц (Particle Swarm Optimization — PSO), который эффективен при настройке многомерных нелинейных систем и легко адаптируется к изменчивым условиям. Вместе с тем, в научной литературе практически отсутствуют исследования, посвящённые совместному применению алгоритма роя частиц PSO, обобщённой дискретизации на основе отсчётов сигнала и его производных и вейвлет-анализа РС.

Степень проработанности темы. Проблематика дискретизации сигналов и её обобщений достаточно полно отражена в трудах российских и зарубежных исследователей. Основы классической теории дискретизации заложены В.А. Котельниковым, К.Э. Шенноном, Э.Т. Уиттекером и Г. Найквистом. Развитие обобщённых теорем отсчётов и многоканальных

схем с использованием производных сигнала представлено в работах Я.И. Хургина, В.П. Яковлева, Дж. Папулиса, М. Веттерли, П.Л. Драготти, и др. Фундаментальные исследования в области ДВП и пороговой обработки выполнены С. Малла, И. Добеши, Р. Коифманом, Д.Л. Доньо и А. Коэном, а в российской научной школе — И.Я. Новиковым, М.А. Скопиной, Е.А. Лебедевой, Ю.А. Фарковым, В.В. Витязевым, Ю.С. Бехтиным и др. Алгоритм роя частиц PSO и его применение для оптимизации параметров обработки сигналов разработаны Дж. Кеннеди, Р. Эберхартом, М. Клерком и получили развитие в работах российских авторов. Вместе с тем анализ литературы показывает отсутствие комплексных исследований, объединяющих обобщённую многоканальную дискретизацию РС, адаптивную вейвлет-пороговую обработку и оптимизацию параметров методом PSO в рамках единого подхода.

Цель и задачи исследования. Целью диссертационного исследования является разработка алгоритмов многоскоростной обработки РС в радиотехнических системах передачи информации на основе представления Хургина – Яковлева и ДВП с оптимизацией параметров пороговой обработки. Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработать математическую модель на основе представления Хургина – Яковлева для многоканальной дискретизации РС.

2. Исследовать возможности реализации методов первичного кодирования РС на основе представления Хургина – Яковлева.

3. Разработать метод адаптивной пороговой обработки вейвлет-коэффициентов на основе энергетического критерия с применением алгоритма роя частиц PSO и интегрировать его в комплексную систему кодирования.

4. Выполнить экспериментальную оценку эффективности разработанных алгоритмов на основе объективных и перцептивных метрик качества в сравнении с существующими методами кодирования речи.

Объект исследования. Алгоритмы многоскоростной обработки нестационарных РС в условиях ограниченной пропускной способности радиотехнических каналов передачи информации.

Предмет исследования. Алгоритмы многоскоростной обработки РС, основанные на представлении Хургина – Яковлева, ДВП с адаптивной пороговой обработкой коэффициентов и оптимизацией параметров порогов методом роя частиц PSO для применения в радиотехнических системах передачи информации.

Методы исследования. В данной диссертационной работе использовались аппарат математического анализа, теория цифровой обработки сигналов, методы вейвлет-анализа и оптимизации, а также методы компьютерного моделирования. Экспериментальная проверка эффективности предложенных алгоритмов осуществлялась на реальных и синтетических РС в условиях, имитирующих ограничения радиотехнических каналов передачи информации.

Научная новизна

1. Разработан алгоритм дискретизации и первичного кодирования РС на

основе представления Хургина – Яковлева с многоканальной реализацией и многоуровневым разложением. (*Соответствует **п. 4 паспорта специальности 2.2.13** Разработка и исследование устройств генерирования, усиления, преобразования и синтеза радиосигналов, сигналов изображения и звука в радиотехнических системах различного назначения, включая системы телевидения. Создание эффективных методов их расчета и основ проектирования*).

2. Предложен метод адаптивной настройки энергетических порогов ДВП для сжатия РС на основе алгоритма роя частиц PSO. (*Соответствует **п. 9 паспорта специальности 2.2.13** Исследование и разработка телевизионных систем и устройств, включая матричные фотоприемники, телевизионные камеры, в том числе с IP-интерфейсами, цифровые кодеки, модемы и устройства передачи и воспроизведения изображений и звука, а также методов и алгоритмов модуляции, демодуляции, сжатия информации, помехоустойчивого кодирования и декодирования сигналов*).

3. Разработана интегрированная модель первичного кодирования речи, объединяющая обобщённую дискретизацию на основе представления Хургина – Яковлева, вейвлет-анализ и оптимизацию энергетических порогов методом роя частиц PSO. (*Соответствует **п. 6 паспорта специальности 2.2.13** Разработка и исследование методов и алгоритмов обработки сигналов и информации в радиотехнических устройствах и системах различного назначения, включая системы телевидения и передачи информации, при наличии помех с целью повышения помехоустойчивости*).

Основные положения, выносимые на защиту

1. Модификация синтезирующих фильтров для представления Хургина – Яковлева, обеспечивающая возможность многоскоростной обработки с точностью восстановления не менее 45 дБ для двухканальных систем и не менее 14 дБ для трёх- и четырёхканальных систем. (*Соответствует **п. 4 паспорта специальности 2.2.13** Разработка и исследование устройств генерирования, усиления, преобразования и синтеза радиосигналов, сигналов изображения и звука в радиотехнических системах различного назначения, включая системы телевидения. Создание эффективных методов их расчета и основ проектирования*).

2. Алгоритмы кодирования речевых сигналов на основе представления Хургина – Яковлева, повышающие качество восстановленной речи по сравнению с системой на основе теоремы В.А. Котельникова на величину до 2,2 балла по шкале MOS (Mean Opinion Score) в системах передачи речи различной скорости при вероятности битовых ошибок в канале до 5%. (*Соответствует **п. 6 паспорта специальности 2.2.13** Разработка и исследование методов и алгоритмов обработки сигналов и информации в радиотехнических устройствах и системах различного назначения, включая системы телевидения и передачи информации, при наличии помех с целью повышения помехоустойчивости*).

3. Алгоритм адаптивной настройки энергетических порогов дискретного вейвлет-преобразования с оптимизацией параметров алгоритмом роя частиц PSO, увеличивающий коэффициент сжатия передаваемой речи на величину до

14% по сравнению с классическими методами пороговой обработки при сохранении заданного уровня качества восстановленной речи. (*Соответствует **п. 9 паспорта специальности 2.2.13** Исследование и разработка телевизионных систем и устройств, включая матричные фотоприемники, телевизионные камеры, в том числе с IP-интерфейсами, цифровые кодеки, модемы и устройства передачи и воспроизведения изображений и звука, а также методов и алгоритмов модуляции, демодуляции, сжатия информации, помехоустойчивого кодирования и декодирования сигналов*).

Достоверность полученных результатов обусловливается использованием в процессе исследований адекватных математических моделей, корректным использованием математического аппарата, подтверждением полученных теоретических результатов имитационным моделированием на реальных речевых сигналах и сравнением полученных данных с результатами других авторов, а также с результатами экспериментальных исследований.

Практическая значимость. Разработанные алгоритмы и модели могут быть использованы при проектировании систем передачи и хранения речевой информации в условиях ограниченной пропускной способности каналов связи, включая радиотехнические, мобильные и спутниковые системы. Результаты работы могут быть интегрированы в программно-аппаратные комплексы для цифровой обработки речи, а также использованы в учебном процессе при подготовке специалистов в области теории сигналов и систем связи.

Внедрение результатов работы. Результаты диссертации внедрены в учебный процесс ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина», в практическую деятельность ПАО «Мобильные ТелеСистемы» (филиал г. Рязани) и ООО «FPT Telecom» (Вьетнам).

Личное участие автора. Автору на правах соавтора принадлежат постановки задач, разработка способов и алгоритмов их решения, отраженные в главах диссертации, а также компьютерное моделирование работы алгоритмов и анализ экспериментальных данных. Степень авторства подтверждается соответствующим заключением кафедры РУС РГРТУ.

Апробация работы. Результаты и положения диссертации докладывались и обсуждались на 1 межрегиональной, 6 всероссийских и 9 международных научно-технических конференциях: 6-я Межрегиональная научная сессия молодых исследователей «Промышленная революция 4.0: взгляд молодёжи» (Тула, 2024 г.); XXVIII, XXIX, XXX Всероссийские научно-технические конференции студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях» (Рязань, 2023, 2024, 2025 гг.); Всероссийская научно-техническая конференция «Актуальные проблемы радиоэлектроники и телекоммуникаций» (Самара, 2025 г.); IX, X Всероссийские научно-технические конференции «Актуальные проблемы современной науки и производства» (Рязань, 2024, 2025 гг.); VII, VIII, IX Международные научно-технические форумы «Современные технологии в науке и образовании» (Рязань, 2024, 2025, 2026 гг.); VIII Международная

научно-практическая конференция «Актуальные проблемы и перспективы развития радиотехнических инфокоммуникационных систем» (Москва, 2024 г.); VIII Международная научно-техническая конференция «Радиотехника, электроника и связь» (Омск, 2025 г.); XI Международная заочная научно-техническая конференция «Информационные технологии. Радиоэлектроника. Телекоммуникации» (Тольятти, 2024 г.); XX, XXI Международные научно-практические конференции «Электронные средства и системы управления» (Томск, 2024, 2025 гг.); XXXII Международная научно-практическая конференция «Радиолокация, навигация, связь» (Воронеж, 2026 г.).

Публикации. По теме диссертации опубликовано 25 работ, в том числе 4 статьи в изданиях, входящих в список рекомендованных ВАК для опубликования результатов кандидатских диссертаций, 16 тезисов и материалов докладов на всероссийских и международных научно-технических конференциях, получено 5 свидетельств о государственной регистрации программ для ЭВМ.

Структура и объем диссертации. Диссертационная работа состоит из введения, трех глав, заключения, приложений и списка литературы. Материал изложен на 185 страницах, включает 41 рисунок и 11 таблиц, а также 4 приложения. Список использованной литературы содержит 134 наименования.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационного исследования, сформулированы его цель и задачи, раскрыта научная новизна, показана практическая значимость полученных результатов, сформулированы основные положения, выносимые на защиту, а также приведены сведения об апробации работы, публикациях, внедрении результатов и структуре диссертации.

В первой главе разработаны теоретические основы K-канальной системы дискретизации и восстановления РС на основе представления Хургина – Яковлева: обоснована применимость метода к нестационарным РС, получены основная формула представления и конкретные аналитические выражения для амплитудно-частотных характеристик (АЧХ) и ФЧХ синтезирующих фильтров при $K = 2, 3, 4$, выбран оптимальный алгоритм вычисления производных, проведена экспериментальная оценка точности восстановления.

Центральным математическим результатом главы является формула восстановления сигнала. Если сигнал $x(t)$ имеет финитный спектр $|f| \le F = 1/(2T)$ и доступны значения его первых $(K-1)$ производных, то сигнал однозначно восстанавливается по отсчётам, взятым с шагом KT — в K раз реже, чем требует теорема В.А. Котельникова:

$$x(t) = \sum_{k=0}^{K-1} \sum_{n=-\infty}^{\infty} \frac{1}{k!} x^{(k)}(nKT)(t-nKT)^k \left[\frac{\sin \pi(t/KT - n)}{\pi(t/KT - n)} \right]^K, \qquad (1)$$

где $h_{n,k}(t) = (t-nKT)^k \left[\frac{\sin \pi(t/KT - n)}{\pi(t/KT - n)} \right]^K$ – импульсная переходная функция

k-го канала определяется типом производной и числом каналов *K*; при *K* = 1 формула вырождается в классическую теорему В.А. Котельникова. Поскольку объём информации на единицу времени в формуле (1) остаётся неизменным — *K*-кратное снижение числа отсчётов сигнала компенсируется передачей $(K-1)$ производных, — представление обеспечивает точное восстановление без увеличения полосы. Обобщённая структурная схема системы, реализующей выражение (1), показана на рисунке 1.

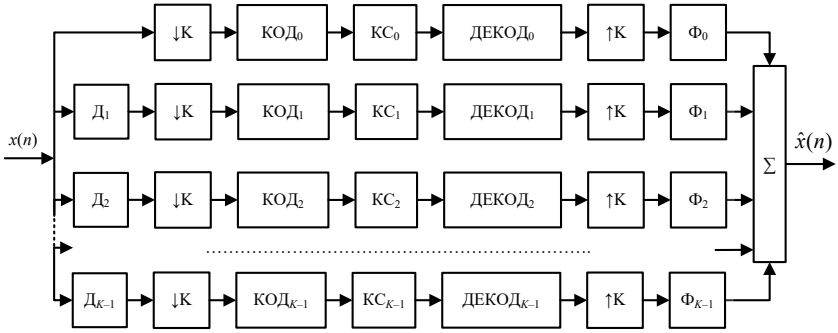


Рисунок 1 – Обобщенная структурная схема системы обработки и передачи РС на основе представления Хургина – Яковлева при *K* канальной реализации

Последовательность преобразований сигнала в данной схеме осуществляется следующим образом. На входе системы передачи исходный сигнал дискретизируется с частотой, определяемой теоремой В.А. Котельникова, в результате чего формируется дискретный сигнал $x(n)$. На передающей стороне исходный сигнал $x(n)$ подаётся на $(K-1)$ дифференциаторов $Д_k$ $(k=1,\ldots,K-1)$, формирующих отсчёты производных порядков $1,2,\ldots,K-1$. Все *K* последовательностей прореживаются в *K* раз дециматором (↓K) и кодируются кодерами ($КОД_k$). На приёмной стороне после декодирования каждая из *K* последовательностей интерполируется (↑K) и подаётся на синтезирующий фильтр $Ф_k$; выходы всех *K* фильтров суммируются, восстанавливая сигнал $\hat{x}(n)$. Альтернативный вариант реализации системы на основе представления Хургина – Яковлева, предусматривающий использование производных более высоких порядков, заключается в многократном применении этой структуры на P этапах разложения.

Ключевой нерешённой проблемой исходной формулировки (1) является неоднозначность ФЧХ синтезирующих фильтров $Ф_k$: они задают АЧХ, но не определяют ФЧХ, что порождает фазовые искажения восстановленного РС. В настоящей работе предложена модификация, состоящая в синтезе конкретных АЧХ и ФЧХ по критерию минимума среднеквадратического отклонения (СКО)

для $K = 2, 3, 4$. Для двухканальной системы ($K = 2$) результат получен в аналитической замкнутой форме, при которой погрешности аппроксимации АЧХ $(\varepsilon_a = 0)$. Для трёхканальной системы $(K = 3)$ аналитические выражения для АЧХ синтезирующих фильтров недостижимы; применяется кусочно-линейная аппроксимация с погрешностью $\varepsilon_a \le 7,69\%$. Для четырёхканальной системы $(K = 4)$ — погрешность $\varepsilon_a \le 4,96\%$. Суммарная погрешность восстановления складывается аддитивно из трёх компонент: $\varepsilon_\Sigma = \varepsilon_\phi + \varepsilon_{der} + \varepsilon_a$, где $\varepsilon_\phi(t)$ – погрешность, обусловленная неопределенностью фазовой характеристики синтезирующих фильтров; $\varepsilon_{der}(t)$ – погрешность вычисления производной; $\varepsilon_a(t)$ – погрешность аппроксимации АЧХ синтезирующих фильтров.

Для вычисления производных в блоках $Д_k$ проведён сравнительный анализ трёх методов: конечных разностей (порядок точности $O(\Delta t^2)$), полиномиальной интерполяции по пяти точкам $(O(\Delta t^4))$ и спектрального метода на основе быстрого преобразования Фурье (БПФ), в котором производная k-го порядка вычисляется умножением спектра на $(j2\pi f)^k$ с последующим обратным БПФ. Спектральный метод не имеет методической погрешности аппроксимации дифференциального оператора. Сравнение проведено на 500 синтетических реализациях при размере кадра $M = 256$ отсчётов; точность оценивалась по нормированной СКО (НСКО) σ (в процентах). Зависимости НСКО σ от нормированной частоты дискретизации $F_н = F_д / F_к$, где $F_д$ — рабочая частота дискретизации, $F_к$ — частота Котельникова, показаны на рисунке 2. При нормированной частоте $F_н = 1,0$ спектральный метод обеспечивает $\sigma = 1,8\%$, что на 18% лучше полиномиального метода и на 44% лучше метода конечных разностей. Для производных высших порядков при рабочей частоте $F_д = 8$ кГц НСКО стабилизируется: для второй производной

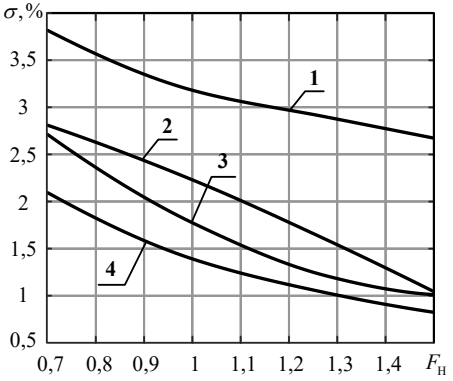


Рисунок 2 – Зависимость НСКО вычисления первой производной от нормированной частоты для четырёх алгоритмов: 1 — конечные разности; 2 — полиномиальная интерполяция; 3 — спектральный метод; 4 — теоретический предел

$\sigma \le 0,01$, для третьей $\sigma \le 0,04$. На основании полученных результатов спектральный метод принят в качестве базового алгоритма для всех конфигураций системы.

Экспериментальная оценка точности восстановления проведена на 100 записях от 10 дикторов (в соответствии с требованиями ГОСТ Р 50840–95) при $M = 256$ отсчётов, $F_д = 8$ кГц, 500 независимых испытаниях для каждой из 9 конфигураций $\left(K \in \{2,3,4\}, P \in \{1,2,3\}\right)$ по идеальному каналу без квантования и шума. Метрики: отношение сигнал/шум (ОСШ, дБ), НСКО и перцептивная оценка ViSQOL (Virtual Speech Quality Objective Listener, шкала MOS, 1…5 баллов). Результаты приведены в таблице 1.

Таблица 1 – Оценки точности восстановления РС для различных конфигураций системы в идеальном канале

Число каналов	Число этапов	Оценка качества		
		НСКО	ОСШ (дБ)	ViSQOL (Балл)
$K = 2$	$P = 1$	$7,6 \cdot 10^{-12}$	85,38	4,73
	$P = 2$	$9,9 \cdot 10^{-9}$	55,93	4,72
	$P = 3$	$3,7 \cdot 10^{-8}$	48,88	4,69
$K = 3$	$P = 1$	$6,3 \cdot 10^{-6}$	26,02	4,71
	$P = 2$	$8,7 \cdot 10^{-6}$	24,47	4,65
	$P = 3$	$9,3 \cdot 10^{-6}$	24,04	4,50
$K = 4$	$P = 1$	$3,6 \cdot 10^{-5}$	17,92	4,69
	$P = 2$	$4,5 \cdot 10^{-5}$	15,66	4,64
	$P = 3$	$5,9 \cdot 10^{-5}$	14,03	4,30

Из таблицы 1 следует: максимальное значение ОСШ = 85,38 дБ достигается при конфигурации $K = 2$, $P = 1$, что обусловлено минимизацией фазовых и амплитудных погрешностей восстановления. В диапазоне $P = 1…3$ обеспечивается ОСШ не менее 45 дБ и оценка качества ViSQOL не ниже 4,69 балла MOS. Для конфигураций $K = 3$ и $K = 4$ при всех значениях P достигаются показатели ОСШ ≥ 14 дБ и ViSQOL ≥ 4,30 балла MOS, что превышает требования к качеству телефонной связи и подтверждает применимость метода для низкоскоростного кодирования. Использование конфигураций $K > 4$ признано нецелесообразным, поскольку накопление погрешностей вычисления производных высших порядков приводит к снижению ОСШ ниже 13 дБ без существенного выигрыша в коэффициенте сжатия. Совокупность результатов таблицы 1 **подтверждает первое положение**, выносимое на защиту.

Во второй главе разработаны алгоритмы первичного кодирования РС, проведён анализ эквивалентности системы банку фильтров, исследована вычислительная сложность, обоснована помехоустойчивость при воздействии аддитивного белого гауссовского шума (АБГШ) и проведены экспериментальные исследования совместно с алгоритмами первичного кодирования различных классов скоростей передачи.

Доказана строгая эквивалентность системы обработки на основе представления Хургина – Яковлева K-канальному банку фильтров анализа-синтеза с максимальной децимацией в K раз. В частности, k-й анализирующий фильтр системы в z-области описывается выражением $H_k(z) = \left(1 - z^{-1}\right)^k$, $k = 1, 2, \ldots, K-1$, что соответствует численному аналогу производной k-го порядка. Обратимость матрицы наложения $\mathbf{H}(z)$ гарантирует существование синтезирующих фильтров $F_k(z)$, при которых выполняется условие полного восстановления. Применение полифазного разложения позволяет вынести децираторы за пределы фильтрации, снизив количество вычислительных операций. В диссертации показано, что для реализации представления Хургина – Яковлева необходимы дополнительные вычислительные операции, не превышающие 4 MIPS.

Аналитически обоснован выигрыш в помехоустойчивости при воздействии АБГШ со спектральной плотностью мощности G_ξ в полосе $(0; F)$.

Эффективность алгоритмов первичного кодирования при совместном использовании с представлением Хургина – Яковлева оценивалась по разностной метрике $\Delta Q = Q_{ХЯ} - Q_{КОТ}$, где $Q_{ХЯ}$ и $Q_{КОТ}$ — значения оценки качества ViSQOL систем на основе представления Хургина – Яковлева и теоремы В.А. Котельникова при идентичной скорости передачи. Исследование проводилось на 50 РС, 100 независимых испытаниях для каждого сочетания $\{K, P$, кодек, $P_{ош}\}$ с тремя кодеками: MMBE (1,2 кбит/с), G.723.1 (6,3 кбит/с), G.726 (24 кбит/с). Исследование разбито на два этапа. На первом этапе (идеальный канал, $P_{ош} = 0\%$) ΔQ обусловлен некогерентным суммированием шума квантования в K каналах. На рисунке 3 показаны зависимости ΔQ от числа этапов разложения P для трёх кодеков.

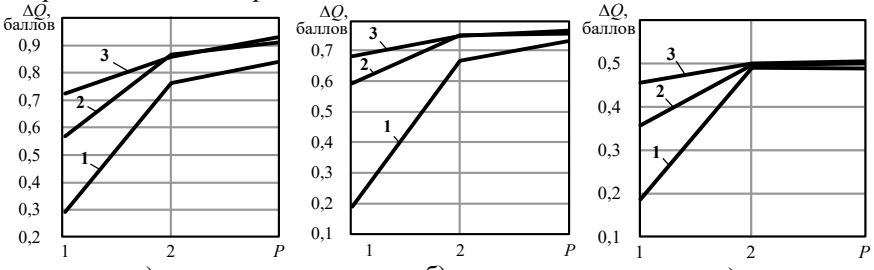


Рисунок 3 – Выигрыш в качестве восстановления РС при различных K и P: а) MMBE (1,2 кбит/с); б) G.723.1 (6,3 кбит/с); в) G.726 (24 кбит/с), кривые: 1 — $K = 2$, 2 — $K = 3$, 3 — $K = 4$

Как видно из приведенных зависимостей, выигрыш в качестве восстановления речи остается положительным во всех рассмотренных случаях, что

свидетельствует о стабильном превосходстве системы на основе представления Хургина – Яковлева над классической системой по воспринимаемому качеству восстановленной речи при эквивалентных условиях передачи. Это происходит за счёт того, что многоканальная обработка с передачей производных позволяет снижать влияние шумов квантования в первичных кодеках речи.

Далее на втором этапе, проведено исследование помехоустойчивости системы в условиях наличия ошибок в канале связи. Вероятность битовых ошибок в канале $P_{ош}$ варьировалась в диапазоне 0,1…5%. Диапазоны значений ΔQ приведены в таблице 2.

Таблица 2. Диапазоны выигрыша ΔQ в качестве восстановленной речи при наличии ошибок в канале связи $(P_{ош} = 0{,}1...5\%)$ для различных конфигураций системы передачи

Кодек	Число каналов K	$P = 1$	$P = 2$	$P = 3$
ММВЕ (1,2 кбит/с)	$K = 2$	0,36…0,77	0,80…1,53	0,89…1,78
	$K = 3$	0,60…1,07	0,89…1,76	0,94…1,86
	$K = 4$	0,79…1,06	0,77…1,52	0,99…1,96
G.723.1 (6,3 кбит/с)	$K = 2$	0,00…0,65	0,67…1,63	0,87…1,74
	$K = 3$	0,53…0,71	0,71…1,81	0,79…1,93
	$K = 4$	0,61…1,47	0,80…1,85	0,91…2,20
G.726 (24 кбит/с)	$K = 2$	0,29…0,34	0,54…1,01	0,55…1,19
	$K = 3$	0,27…0,64	0,53…1,16	0,52…1,27
	$K = 4$	0,50…0,99	0,54…1,17	0,58…1,34

Таблица 2 позволяет сформулировать три ключевых вывода. Во-первых, ΔQ строго положительно во всём исследованном пространстве параметров $\{K, P,$ кодек, $P_{ош}\}$ – отсутствие регрессии подтверждает робастность метода. Во-вторых, ΔQ монотонно возрастает с ростом K, P и $P_{ош}$: чем интенсивнее помеха, тем эффективнее фазовая структура синтезирующих фильтров распределяет ошибки по каналам. В-третьих, максимум $\Delta Q = 2{,}2$ балла достигается для кодека G.723.1, конфигурации $K = 4$, $P = 3$ при $P_{ош} = 5\%$: алгоритм G.723.1 (с кодированием кратковременного предсказания и параметров возбуждения) наиболее чувствителен к ошибкам передачи своих параметров, и именно в этом случае некогерентное суммирование шума даёт наибольший эффект. Важно подчеркнуть, что выигрыш достигается без снижения скорости передачи: K-кратное снижение числа отсчётов компенсируется передачей $(K - 1)$ производных, суммарный битрейт остаётся неизменным. Изложенные результаты **подтверждают второе положение**, выносимое на защиту.

В третьей главе разработан метод сжатия РС на основе ДВП с адаптивной оптимизацией вектора энергетических порогов алгоритмом роя частиц PSO, а также исследована интегрированная система на основе представления Хургина – Яковлева и ДВП.

Реализована покадровая схема с J-уровневым ДВП по алгоритму С. Маллата. В результате разложения каждый кадр представляется вектором вейвлет-коэффициентов $\mathbf{C}_m = \left[cA_J^{(m)}; cD_J^{(m)}; cD_{J-1}^{(m)}; \ldots; cD_1^{(m)} \right]^T$: коэффициенты аппроксимации $cA_J^{(m)}$ содержат основную энергию РС, тогда как подавляющее большинство коэффициентов деталей $cD_j^{(m)}$ несут пренебрежимо малую энергию и могут быть обнулены без ощутимой потери качества восстановления — свойство естественной разреженности РС в вейвлет-области. Сравнительное исследование 17 вейвлет-базисов четырёх семейств (Добеши, симлет, койфлет, биортогональных) по интегральному критерию эффективности установило: вейвлет Добеши db4 при $J = 4$ уровнях оптимален. Оптимум $J = 4$ инвариантен к типу вейвлета и физически обусловлен спектральной структурой РС.

Классические методы Донохо–Джонстон и SUREShrink основаны на предположении статистической однородности шума и гауссовости вейвлет-коэффициентов. РС не удовлетворяют этим условиям: их энергетическое распределение по уровням разложения существенно неравномерно и меняется от кадра к кадру, поэтому применение единого порога ко всем уровням систематически недооптимально. Задача выбора порогов сформулирована как задача нелинейной оптимизации: максимизировать коэффициент сжатия $\rho(\boldsymbol{\alpha})$ при ограничениях на качество $Q_1(\boldsymbol{\alpha}) \geq Q_1^{\min} = 9{,}5$ дБ и $Q_2(\boldsymbol{\alpha}) \geq Q_2^{\min} = 3{,}5$ балла по метрикам SegSNR (сегментированное отношение сигнал/шум) и ViSQOL соответственно. Вектор $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_J]^T \in [0,1]^{J+1}$ содержит $J+1 = 5$ раздельных порогов. Целевая функция имеет вид:

$$F(\boldsymbol{\alpha}) = \mathrm{w}_1 f_\rho(\boldsymbol{\alpha}) + \mathrm{w}_2 f_{Q_1}(\boldsymbol{\alpha}) + \mathrm{w}_3 f_{Q_2}(\boldsymbol{\alpha}), \tag{4}$$

где $f_\rho(\boldsymbol{\alpha})$, $f_{Q_1}(\boldsymbol{\alpha})$, $f_{Q_2}(\boldsymbol{\alpha})$ — нормализованные значения коэффициента сжатия, SegSNR и ViSQOL; а весовые коэффициенты $\mathrm{w}_1, \mathrm{w}_2, \mathrm{w}_3 \geq 0$ удовлетворяют условию $\mathrm{w}_1 + \mathrm{w}_2 + \mathrm{w}_3 = 1$. Функция (4) является негладкой и многоэкстремальной вследствие кусочно-постоянного характера операции обнуления коэффициентов ДВП, что исключает применение градиентных методов. Алгоритм PSO не требует вычисления градиента, допускает включение ограничений в структуру целевой функции и устойчиво сходится при размерности $D = J+1 = 5$ (одноканальная система) и $D = 2(J+1) = 10$ (двухканальная интегрированная система). Параметры алгоритма: размер роя $N_p = 200$ частиц; максимальное число итераций $I_{\max} = 100$; коэффициент инерции изменяется линейно от $w_{\max} = 0{,}9$ до $w_{\min} = 0{,}3$, обеспечивая баланс между глобальным поиском на ранних итерациях и локальной оптимизацией на завершающих; когнитивный и социальный коэффициенты $c_1 = c_2 = 2{,}0$;

максимальная скорость частицы $v_{\max} = 0{,}15$. Критерием ранней остановки является отсутствие улучшения глобального оптимума $F(\boldsymbol{\alpha})$ в течение 20 последовательных итераций при изменении целевой функции менее 10^{-6}.

Для верификации предложенного метода PSO-оптимизации энергетических порогов проведено сравнение с тремя классическими методами пороговой обработки вейвлет-коэффициентов. Исследование выполнено для одноканальной системы ДВП $(D = J + 1 = 5)$. При указанных ограничениях PSO-метод обеспечивает коэффициент сжатия ρ = 87,0 %, что превышает метод Донохо–Джонстона на 14,0 %, метод SUREShrink — на 9,5 %, фиксированный энергетический порог — на 7,1 %. Превосходство обусловлено раздельной оптимизацией порога для каждой из J+1 частотных полос, что учитывает реальную неравномерность энергетического распределения РС по уровням разложения — в отличие от классических методов, применяющих единый порог ко всем уровням. Полученные результаты **подтверждают третье положение**, выносимое на защиту.

Совместное применение представления Хургина – Яковлева и ДВП с PSO-оптимизацией реализовано в интегрированной системе, структурная схема которой представлена на рисунке 4.

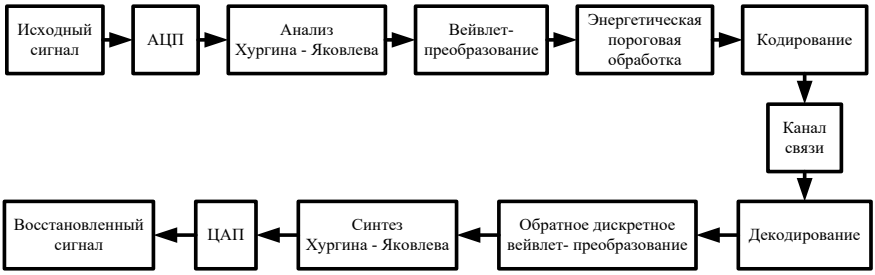
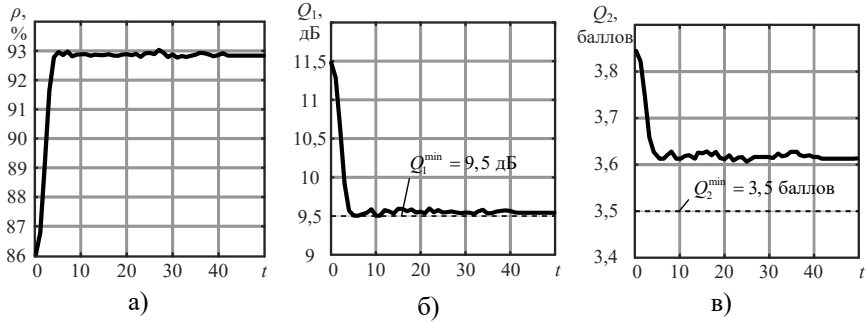


Рисунок 4 – Структурная схема интегрированной системы обработки и передачи сигналов на основе представления Хургина – Яковлева и вейвлет-преобразования

В данной системе (D = 10, K = 2, P = 1) каждый субполосный сигнал блока Хургина – Яковлева независимо обрабатывается блоком ДВП с пороговым сжатием, а вектор $\boldsymbol{\alpha}$ охватывает пороги обоих каналов по всем уровням ДВП. Динамика оптимизации показана на рисунке 5: $\rho(t)$ интенсивно растёт с 86 % до 92,9 % в первые 10 итераций, начиная с $t \approx 40$ устанавливается стационарный режим ρ = 92,9 %; алгоритм завершает работу на итерации t = 45. $Q_1(t)$ стабилизируется на уровне 9,50 ± 0,05 дБ (активное ограничение); $Q_2(t) = 3{,}6$ балла при запасе 0,1 балла относительно $Q_2^{\min} = 3{,}5$ (неактивное ограничение). Достигнутый коэффициент сжатия ρ = 92,9 % превышает одноканальный PSO-метод (ρ = 87,0 %) на 5,9 процентных пункта —

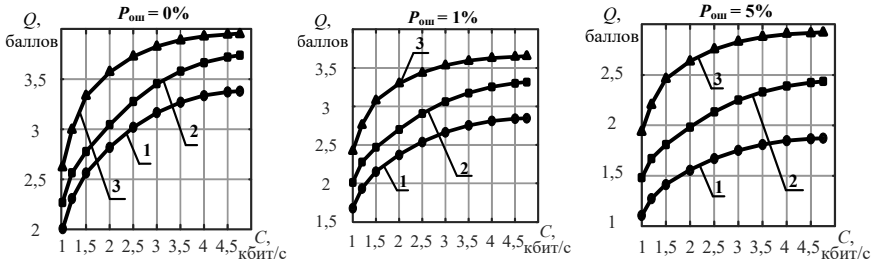
синергетический эффект: совместная оптимизация порогов обоих каналов учитывает их взаимодействие через сумматор синтезирующего блока, чего раздельная оптимизация достичь не может.



а) б) в)

Рисунок 5 – Результаты многокритериальной оптимизации методом PSO: зависимости от номера итерации *t*
(а) степени сжатия; (б) SegSNR; (в) ViSQOL

Экспериментальное исследование интегрированной системы при кодировании кодеком LBRAMR (скорость C = 1,0…5,0 кбит/с; вероятность ошибок $P_{ош}$ = 0...5%) показало: выигрыш по ViSQOL до 0,9 балла по шкале MOS по сравнению с системой на основе теоремы В.А. Котельникова без увеличения скорости передачи (рисунок 6). Анализ вычислительной сложности подтвердил полную аппаратную реализуемость на платформах TMS320C6748, ADSP-21569, Zynq-7020 и Cortex-M55 без компромисса по качеству обработки.



а) б) в)

Рисунок 6 – Зависимость качества восстановленной речи от скорости *C* (кбит/с) при: а) $P_{ош}$ = 0 %; б) $P_{ош}$ = 1 %; в) $P_{ош}$ = 5 %, кривые: 1 – система Котельникова, 2 – ДВП+PSO, 3 – интегрированная система

В заключении приводятся результаты диссертации, подтверждающие основные положения, выносимые на защиту.

В приложении содержатся титульные листы свидетельств о государственной регистрации программ для ЭВМ и акты внедрения результатов диссертационной работы.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в журналах, рекомендованных ВАК

1. Ву Хоанг Шон. Применение трехканальной модификации алгоритма Хургина – Яковлева в алгоритмах первичного кодирования речевых сигналов / В. Т. Дмитриев, Ву Хоанг Шон // Вестник Рязанского государственного радиотехнического университета. – 2024. – № 88. – С. 3-14.

2. Ву Хоанг Шон. Разработка многоканальной системы на основе модификации алгоритма Хургина – Яковлева при использовании отсчетов сигнала и его трех первых производных / В. Т. Дмитриев, Ву Хоанг Шон // Вестник Рязанского государственного радиотехнического университета. – 2025. – № 91. – С. 24-35.

3. Ву Хоанг Шон. Разработка и исследование первичных кодеков речи при многоэтапном разложении на основе модификации алгоритма Хургина – Яковлева / В. Т. Дмитриев, Ву Хоанг Шон // Цифровая обработка сигналов. – 2025. – № 3. – С. 76-85.

4. Ву Хоанг Шон. Оптимизация энергетических порогов вейвлет-преобразования для сжатия речевых сигналов на основе алгоритма роя частиц / В. Т. Дмитриев, Ву Хоанг Шон // Вестник Рязанского государственного радиотехнического университета. – 2026. – № 95. – С. 3-15.

Публикации в других изданиях

5. Ву Хоанг Шон. Актуальность разработки низкоскоростных первичных кодеков речевых сигналов. // Новые информационные технологии в научных исследованиях и образовании – НИТ-2023: материалы XXVIII Всероссийской научно-технической конференции студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2023. – С. 138-140.

6. Ву Хоанг Шон. Модификация алгоритма Хургина – Яковлева при трёхканальной реализации / В. Т. Дмитриев, Ву Хоанг Шон // Современные технологии в науке и образовании – СТНО-2024: сборник трудов VII Международного научно-технического форума: в 10 т. – Рязань: РГРТУ, 2024. – Т. 1. – С. 139-143.

7. Ву Хоанг Шон. Сравнение эффективности первичных кодеков речевых сигналов на основе модификации алгоритма Хургина – Яковлева / В. Т. Дмитриев, Ву Хоанг Шон // Электронные средства и системы управления – МНПК ЭСиСУ-2024: материалы докладов XX Международной научно-практической конференции. – Томск: В-Спектр, ТУСУР, 2024. – Ч. 1. – С. 33-35.

8. Ву Хоанг Шон. Восстановление речевых сигналов на основе модифицированного алгоритма Хургина – Яковлева при четырехканальной реализации / В. Т. Дмитриев, Ву Хоанг Шон // Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем – РАДИОИНФОКОМ-2024: Сборник научных статей по материалам VIII Международной научно-практической конференции. – Москва, МИРЭА, 2024. – С.159-164.

9. Ву Хоанг Шон. Выигрыш качества воспроизведения речевых сигналов с алгоритмом модифицированного Хургина – Яковлева в трех- и четырехканальных реализациях // Новые информационные технологии в научных исследованиях и образовании – НИТ-2024: материалы XXIX Всероссийской

научно-технической конференции студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2024. – С. 132-134.

10. Ву Хоанг Шон. Разработка и исследование помехоустойчивой системы передачи речевой информации на основе алгоритма Хургина – Яковлева / В. Т. Дмитриев, Ву Хоанг Шон // Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем – АПСНИП-2024: Сборник научных статей по материалам IX Международной научно-практической конференции. – Рязань: РГРТУ, 2024. – С.51-58.

11. Ву Хоанг Шон. Анализ вычислительных затрат при обработке речевых сигналов на основе алгоритма Хургина – Яковлева // Промышленная революция 4.0: взгляд молодежи: тезисы докладов 6-й Межрегиональной научной сессии молодых исследователей. – Тула: ТулГУ, 2024. – С. 32-33.

12. Ву Хоанг Шон. Определение выигрыша помехоустойчивости многоканальной системы передачи на основе модификации алгоритма Хургина – Яковлева / В. Т. Дмитриев, Ву Хоанг Шон // Информационные технологии. Радиоэлектроника. Телекоммуникации – ITRT -2024: материалы XI Международной заочной научно-технической конференции. – Тольятти: ПВГУС, 2025. – С.445-451.

13. Ву Хоанг Шон. Система обработки и передачи информации на основе модификации алгоритма Хургина – Яковлева при повторном разложении / В. Т. Дмитриев, Ву Хоанг Шон // Современные технологии в науке и образовании – СТНО-2025: сборник трудов VIII Международного научно-технического форума: в 10 т. – Рязань: РГРТУ, 2025. – Т. 1. – С. 171-176.

14. Ву Хоанг Шон. Сравнение помехоустойчивости радиосистем на основе модификации алгоритма Хургина – Яковлева и теоремы В.А. Котельникова / В. Т. Дмитриев, Ву Хоанг Шон // Актуальные проблемы радиоэлектроники и телекоммуникаций: материалы всероссийской научно-технической конференции. – Самара: ООО «АРТЕЛЬ», 2025. – С. 184-186.

15. Ву Хоанг Шон. Сжатие речевых сигналов на основе вейвлет-пакетного разложения / В. Т. Дмитриев, Ву Хоанг Шон // Радиотехника, электроника и связь – РЭиС-2025. VIII Международная научно-техническая конференция. – Омск, 2025. – С. 101-103.

16. Ву Хоанг Шон. Оценка эффективности кодеков речи на основе вейвлет-пакетного разложения / В. Т. Дмитриев, Ву Хоанг Шон // Электронные средства и системы управления – МНПК ЭСиСУ-2025: материалы докладов XXI Международной научно-практической конференции. – Томск: В-Спектр, ТУСУР, 2025. – Ч. 1. – С. 39-42.

17. Ву Хоанг Шон. Алгоритмы пороговой обработки в вейвлет – разложении для сжатия речевых сигналов // Новые информационные технологии в научных исследованиях и образовании – НИТ-2025: материалы XXX Юбилейной всероссийской научно-технической конференции студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2025. – С. 125-126.

18. Ву Хоанг Шон. Метод сжатия речевых сигналов на основе

дискретного вейвлет-преобразования с энергетическим пороговым отсечением / В. Т. Дмитриев, Ву Хоанг Шон // Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем – АПСНИП-2025: Сборник научных статей по материалам X Международной научно-практической конференции. – Рязань: РГРТУ, 2025. – С.53-62.

19. Ву Хоанг Шон. Метод первичного кодирования речи на основе модифицированного алгоритма Хургина – Яковлева и дискретного вейвлет-преобразования / В. Т. Дмитриев, Ву Хоанг Шон // Современные технологии в науке и образовании – СТНО-2026: сборник трудов IX Международного научно-технического форума: в 10 т. – Рязань: РГРТУ, 2026. – Т. 1. – С. 224-229.

20. Ву Хоанг Шон. Экспериментальная оценка эффективности алгоритмов первичного кодирования речевых сигналов на основе представления Хургина – Яковлева / В. Т. Дмитриев, Ву Хоанг Шон // Радиолокация, навигация, связь: сборник трудов XXXII Международной научно-технической конференции. – Воронеж: Издательский дом ВГУ, 2026. – Т. 2. – С. 134-141.

Свидетельства о регистрации программы для ЭВМ

21. Дмитриев В. Т., Ву Хоанг Шон. Программа моделирования модификации алгоритма Хургина – Яковлева при трехканальной реализации: свидетельство о гос. регистрации программы для ЭВМ № RU 2024619863. – Зарегистр. 27.04.2024; опубл. 27.04.2024. – Правообладатель: ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В. Ф. Уткина».

22. Дмитриев В. Т., Ву Хоанг Шон. Программа для исследования синтезирующих фильтров Хургина – Яковлева: свидетельство о гос. регистрации программы для ЭВМ № RU 2025618510. – Зарегистр. 04.04.2025; опубл. 04.04.2025. – Правообладатель: ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В. Ф. Уткина».

23. Дмитриев В. Т., Ву Хоанг Шон. Программа моделирования системы передачи информации при восстановлении сигнала по его прореженным отсчетам и отсчетам его трех первых производных: свидетельство о гос. регистрации программы для ЭВМ № RU 2025619213. – Зарегистр. 14.04.2025; опубл. 14.04.2025. – Правообладатель: ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В. Ф. Уткина».

24. Дмитриев В. Т., Ву Хоанг Шон. Программа моделирования многоэтапного разложения речевых сигналов на основе модификации алгоритма Хургина – Яковлева: свидетельство о гос. регистрации программы для ЭВМ № RU 2025663352. – Зарегистр. 27.05.2025; опубл. 27.05.2025. – Правообладатель: ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В. Ф. Уткина».

25. Дмитриев В. Т., Ву Хоанг Шон. Программа сжатия речевых сигналов на основе энергетических порогов в дискретном вейвлет-преобразовании с оптимизацией методом роя частиц: свидетельство о гос. регистрации программы для ЭВМ № RU 2026611916. – Зарегистр. 22.01.2026; опубл. 22.01.2026. – Правообладатель: ФГБОУ ВО «Рязанский государственный радиотехнический университет имени В. Ф. Уткина».

ВУ Хоанг Шон

АЛГОРИТМЫ МНОГОСКОРОСТНОЙ ОБРАБОТКИ РЕЧЕВЫХ СИГНАЛОВ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать ___. _____. 2026. Формат бумаги 60×84 1/16.
Бумага офсетная. Печать трафаретная. Усл. печ. л. 1,0.
Тираж 100 экз. Заказ _____
Рязанский государственный радиотехнический университет.
390005, г. Рязань, ул. Гагарина, 59/1.
Редакционно-издательский центр РГРТУ