

УТВЕРЖДАЮ

Первый проректор Белорусского
государственного университета

А.В. Блохин
« 1 » октября 2026 г.



ОТЗЫВ

оппонирующей организации о диссертации

Порхуна Максима Игоревича

«Методы модификации и синтеза звуковых сигналов на основе банков цифровых фильтров для моделирования слухового восприятия и повышения разборчивости речи»,
представленной на соискание ученой степени кандидата технических наук по специальности 05.13.17 – теоретические основы информатики (отрасль науки – технические науки)

Отзыв подготовлен на основании экспертизы диссертации, ее автореферата и опубликованных соискателем работ. Эксперт оппонирующей организации – Хейдоров Игорь Эдуардович, кандидат физико-математических наук, доцент, доцент кафедры радиофизики и цифровых медиа технологий – назначен приказом ректора БГУ от 14.09.2026 № 492-ОД в соответствии с пунктом 42 Положения о присуждении ученых степеней и присвоении ученых званий.

Проект отзыва обсужден на расширенном заседании кафедры радиофизики и цифровых медиа технологий Белорусского государственного университета – 22.09.2026, протокол № 2; соискатель выступил с докладом по диссертации. Присутствовало 15 человек, из них имеющих ученые степени – 11 (докторов наук – 1, кандидатов наук – 10).

1. Соответствие содержания диссертации заявленной специальности и отрасли науки

Объектом исследования являются системы обработки речевой информации для моделирования и улучшения слухового восприятия, предметом – методы обработки звукового сигнала, обеспечивающие моделирование и психоакустически обусловленное улучшение слухового восприятия акустической информации, воспринимаемой тугоухим человеком. Содержание диссертации и выносимые на защиту положения относятся к отрасли технических наук.

Оппонирующая организация отмечает, что содержание диссертации соответствует специальности 05.13.17 – теоретические основы информатики и следующим областям исследований паспорта этой специальности, утвержденного ВАК Республики Беларусь:

– «модели, методы и средства распознавания, классификации и

визуализации образов, в том числе графических и звуковых, их сжатие, фильтрация, модификация и синтез, а также создание интеллектуальных систем классификации и идентификации изображений» (пункт 3 паспорта) – методы модификации и синтеза звукового сигнала (главы 2, 5) и метод аппроксимации частотных характеристик банка гамматон-фильтров (глава 3);

– «системы мультимедиа, виртуальной и дополненной реальности для различных прикладных целей, в том числе для медицинской и технической диагностики, отображения и оценки финансово-экономической информации, а также для автоматизированного и дистанционного управления и обучения» (пункт 9 паспорта) – модели слухового восприятия как средство воспроизведения слуховых патологий для задач слухопротезирования и обучения (главы 2, 3);

– «информационно-аналитические средства, используемые при оценке качества и других показателей изделий, процессов, систем и программных продуктов» (пункт 11 паспорта) – применение разработанной модели потери слуха для объективной первичной оценки алгоритмов обработки сигнала без привлечения слабослышащих лиц.

2. Актуальность темы диссертации

Актуальность темы определяется потребностью в средствах объективной оценки и настройки алгоритмов улучшения слуха. Как показано в главе 1, существующие методы моделирования слуховых патологий опираются на упрощенные представления о тугоухости, ограничены в настройке по аудиограмме конкретного человека, а наиболее точные модели, работающие во временной области, вычислительно сложны; при наличии «мертвых зон» в улитке уха традиционная стратегия усиления неэффективна. Работа выполнена в рамках научно-исследовательских работ «Речевые интерфейсы в системах мультимедиа» (ГБ 16-2008) и «Методы и алгоритмы для прикладных технологий обработки мультимедиа данных» (ГБ 21-2008), а ее тематика соответствует пункту 4 перечня приоритетных направлений научных исследований Республики Беларусь на 2026–2030 годы (постановление Совета Министров Республики Беларусь от 09.10.2025 № 555).

3. Научный вклад соискателя в решение научной задачи и оценка его значимости

Соискателем решена научная задача разработки вычислительно эффективных методов моделирования слухового восприятия, настраиваемых по аудиограмме слушателя, и методов повышения разборчивости речи на основе банков цифровых фильтров.

Оппонирующая организация отмечает, что вклад соискателя не исчерпывается совокупностью частных методов: в диссертации выстроен методически связный инструментальный контур. Модель ухудшения частотной избирательности слуха (глава 2) задает объективную меру последствий тугоухости; реализация слуховых фильтров на базе неравнополосного косинусно-модулированного банка фильтров (глава 3)

делает модели, работающие во временной области, практически реализуемыми; ускорение фазового преобразования (глава 4) снимает вычислительное «узкое место» этой реализации; метод повышения разборчивости речи (глава 5) опирается на разработанную модель как на средство собственной экспериментальной проверки.

Теоретическая составляющая вклада состоит в доопределении частотно- и уровне-зависимого коэффициента расширения *гоех*-фильтров для диапазона потери слуха 65–90 дБ, ранее не охваченного моделью Нэдзимэ–Мура; в установлении условия, при котором построение фильтра-прототипа в виде произведения логистических сигмоидальных функций гарантирует монотонность спадов характеристик банка фильтров; в синтезе структуры сдвоенного фазового звена, сохраняющей критический путь одиночного звена. Вклад соискателя соответствует требованиям, предъявляемым к кандидатским диссертациям, и оценивается как значимый для развития методов психоакустически мотивированной обработки звуковых сигналов.

4. Конкретные научные результаты, за которые соискателю может быть присуждена искомая ученая степень

4.1. Метод модификации и синтеза звукового сигнала для моделирования эффекта ухудшения частотной избирательности слуха.

Новизна состоит в доопределении зависимости коэффициента расширения *гоех*-фильтров от частоты и уровня потери слуха: выражение модели Нэдзимэ–Мура, справедливое лишь до 65 дБ, распространено на диапазон 65–90 дБ линейной интерполяцией со сглаживанием частотной зависимости. В отличие от известных подходов, ограниченных потерей слуха до 65 дБ либо учитывающих лишь степень тугоухости без частотной зависимости, это позволяет моделировать ухудшение частотной избирательности для всех степеней тугоухости, включая IV, с учетом порогов слышимости на каждой частоте.

Практическая значимость: метод обеспечивает автоматический расчет параметров «размывающей» функции непосредственно по аудиограмме. Программная реализация внедрена в ОАО «Исток-Аудио Интернэшнл» и применяется при первичной оценке новых методов обработки сигнала для слуховых аппаратов, что сокращает сроки разработки.

4.2. Метод аппроксимации частотных характеристик банка гамматон-фильтров при помощи неравнополосного косинусно-модулированного банка фильтров.

Новизна состоит в построении фильтра-прототипа на основе мультипликативной модели в виде произведения логистических сигмоидальных функций, что гарантирует монотонность его импульсной характеристики и монотонные спады амплитудно-частотных характеристик банка, аналогичные спадам гамматон-фильтров. Классические методы синтеза прототипа дают пульсации в полосе заграждения на уровне около минус 43 дБ; предложенная модель их не вносит. Рациональным признано восемь сигмоид при квадратичной ошибке $RSS = 6,093$.

Практическая значимость: реализация 100-канального банка требует 2209 операций умножения и 2901 операции сложения против 41,1 тыс. МАС-операций при прямой реализации банка гамматон-фильтров Паттерсона–Холдсворта в виде набора КИХ-фильтров, то есть примерно в 8 раз эффективнее. Центральные частоты банка перестраиваются изменением единственного параметра фазового звена без пересчета коэффициентов фильтров.

4.3. Метод повышения скорости выполнения алгоритма фазового преобразования на основе структуры сдвоенного фазового звена.

Новизна состоит в том, что синтезированная на основе представления фазового звена в пространстве состояний структура вычисляет выходы каскада из двух звеньев за один проход, сохраняя критический путь одиночного звена. Глубина графа вычислений цепочки из N звеньев сокращается вдвое – с $3N$ до $3/2 \cdot N$ слоев, что повышает степень параллелизма по сравнению с прямой формой реализации.

Практическая значимость: эксперименты на реализациях на языке Си для цепочки из 2048 звеньев показали ускорение в 1,31 и 1,81 раза для двух вариантов структуры. Результат снижает время вычисления выходов неравнополосного косинусно-модулированного банка фильтров и применим в устройствах обработки звука реального времени.

4.4. Метод модификации и синтеза звукового сигнала для повышения разборчивости речи.

Новизна состоит в адаптивном выборе целевой полосы при переносе высокочастотных компонент речевого сигнала в область остаточного слуха: ширина целевой полосы определяется из условия равенства ширин исходной и целевой полос в психоакустической шкале барков, а уровни переносимых компонент корректируются по кривым равной громкости. В отличие от метода Лю и соавторов с фиксированной целевой полосой и согласованием ширин в линейном масштабе частот, это сохраняет психоакустический масштаб переносимых компонент.

Практическая значимость: проверка по артикуляционным таблицам слов ГОСТ 16600 показала повышение разборчивости речи в среднем на 6 % (максимально – на 18 %) в условиях моделируемой потери слуха; значимость различия подтверждена парным t -критерием при уровне значимости $p = 0,03$. Метод ориентирован на лиц, для которых стратегия усиления неэффективна, в том числе при наличии «мертвых зон» в улитке уха.

Работоспособность и прикладную применимость перечисленных результатов подтверждает также построенный на их основе метод моделирования эффекта потери слуха во временной области (раздел 3.7).

5. Обоснованность и достоверность результатов, опубликованность и апробация

Выводы диссертации обоснованы: они опираются на аналитический обзор состояния вопроса, корректные аналитические выкладки и программные реализации предложенных методов в средах MATLAB, Python

и на языке Си. Достоверность подтверждена экспериментами с участием испытуемых по методике артикуляционных измерений ГОСТ 16600 (15 человек – глава 2, 10 человек – глава 5) со статистической проверкой значимости, вычислительными экспериментами с десятикратным повторением измерений и расчетами вычислительной сложности. Внедрение результатов подтверждается актами, приведенными в приложении А.

По теме диссертации опубликовано 14 работ, в том числе 5 статей в рецензируемых научных журналах, соответствующих требованиям пункта 19 Положения о присуждении ученых степеней и присвоении ученых званий («Electronics Letters», «Информатика», «Информатика и автоматизация», «Доклады БГУИР»); суммарный объем публикаций – 5,94 авторского листа. Результаты апробированы на 20, 21, 22, 24, 25 и 26-й международных конференциях «Цифровая обработка сигналов и ее применение» (г. Москва, 2018–2020, 2022–2024) и на 54–56-й научных конференциях аспирантов, магистрантов и студентов БГУИР. Каждое из защищаемых положений обеспечено публикациями в рецензируемых изданиях.

6. Замечания оппонировавшей организации

1. Ключевым достоинством разработанных методов заявлена настройка по аудиограмме конкретного человека, однако экспериментальная проверка выполнена на нормально слышащих испытуемых с использованием типовых аудиограмм, соответствующих II–IV степеням тугоухости (глава 2, с. 49–54; глава 5, с. 96–97). Проверка на индивидуальных аудиограммах слабослышащих лиц не представлена, что сужает доказательную базу именно индивидуальной настройки.

2. Эффективность метода повышения разборчивости речи оценивалась через собственную модель потери слуха соискателя (глава 5, с. 96), что создает зависимость оценки от оцениваемого инструмента; целесообразно дополнить ее проверкой по независимой модели слухового восприятия либо с участием слабослышащих слушателей. Отсутствие у испытуемых периода адаптации соискатель отмечает (с. 97), однако количественно вклад адаптации в прирост разборчивости не оценивает.

3. Выигрыш «примерно в 8 раз» получен сопоставлением только с прямой реализацией банка гамматон-фильтров в виде набора КИХ-фильтров, тогда как сравнение с известной реализацией Слэйни на основе БИХ-фильтров вынесено за рамки исследования (глава 3, с. 68). Увеличение групповой задержки разработанного банка фильтров отмечено, но численно не охарактеризовано (с. 67, рисунок 3.11), хотя задержка критична для заявленного приложения – слуховых аппаратов реального времени.

4. Предложенные структуры сдвоенного фазового звена увеличивают число операций умножения (шесть и семь против двух у прямой формы), а ускорение в 1,81 раза измерено на универсальном процессоре средствами профилировщика среды разработки (глава 4, с. 83–84). Оценка на целевых платформах (сигнальные процессоры, арифметика с фиксированной запятой) и анализ накопления ошибок округления в новой структуре не выполнены.

Высказанные замечания носят частный, уточняющий и дискуссионный характер, не затрагивают существа полученных научных результатов и не влияют на положительную оценку диссертации.

7. Соответствие научной квалификации соискателя искомой ученой степени

Содержание диссертации свидетельствует, что соискатель владеет современным аппаратом цифровой обработки сигналов, теории цифровых банков фильтров и психоакустики, методами численной оптимизации и машинного обучения, а также средствами их программной реализации. Соискатель самостоятельно спланировал и провел экспериментальные исследования с участием испытуемых, корректно применил методики артикуляционных измерений и статистической проверки гипотез и обеспечил внедрение результатов.

Основные результаты и защищаемые положения получены соискателем лично; в публикациях с соавторами его вклад определяется рамками излагаемых в диссертации результатов. Оппонирующая организация считает, что научная квалификация Порхуна Максима Игоревича соответствует ученой степени кандидата технических наук, на которую он претендует.

8. Рекомендации по практическому использованию результатов

1. Организациям, разрабатывающим и производящим слуховые аппараты и иные средства слухопротезирования (в том числе ОАО «Исток-Аудио Интернэшнл», где результаты уже используются), – применять метод моделирования эффекта ухудшения частотной избирательности слуха как штатный инструмент первичной оценки алгоритмов обработки сигнала без привлечения слабослышащих лиц, а структуру сдвоенного фазового звена – при интеграции банков фильтров в устройства обработки звука реального времени.

2. Организациям здравоохранения, осуществляющим диагностику нарушений слуха и слухопротезирование (республиканский научно-практический центр оториноларингологии, центры коррекции слуха), – рассмотреть метод повышения разборчивости речи при настройке слуховых аппаратов пациентам с выраженной высокочастотной тугоухостью и «мертвыми зонами» в улитке уха. С учетом замечания 3 внедрению целесообразно предпослать клиническую апробацию с участием слабослышащих пациентов и предусмотренным периодом адаптации.

3. Учреждениям образования – использовать результаты в учебном процессе по дисциплинам цифровой обработки сигналов. Такое использование реализовано на кафедре электронных вычислительных средств БГУИР для специальности 1-40 02 02 «Электронные вычислительные средства» и может быть распространено на профильные специальности других учреждений образования.

4. Организациям, разрабатывающим системы распознавания и кодирования речи, выявления речевых и голосовых нарушений, –

использовать метод аппроксимации банка гамматон-фильтров неравнополосным косинусно-модулированным банком фильтров как вычислительно эффективный модуль частотного анализа, согласованный с работой слуховой системы. Предложенный подход к объективной оценке алгоритмов может быть учтен при разработке методических документов по оценке характеристик средств слухопротезирования.

9. Выводы

Диссертация Порхуна Максима Игоревича «Методы модификации и синтеза звуковых сигналов на основе банков цифровых фильтров для моделирования слухового восприятия и повышения разборчивости речи» представляет собой завершенную научно-квалификационную работу, в которой решена актуальная научная задача разработки вычислительно эффективных методов моделирования слухового восприятия и методов повышения разборчивости речи на основе банков цифровых фильтров. По уровню научной новизны, обоснованности и достоверности результатов, их научной и практической значимости, а также по полноте опубликования диссертация соответствует требованиям, установленным главой 3 Положения о присуждении ученых степеней и присвоении ученых званий, утвержденного Указом Президента Республики Беларусь от 17.11.2004 № 560, а ее содержание соответствует специальности 05.13.17 – теоретические основы информатики (технические науки).

Порхун Максим Игоревич заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.17 – теоретические основы информатики за следующие научные результаты:

1. Метод моделирования эффекта ухудшения частотной избирательности слуха, основанный на частотном размытии амплитудного спектра банком расширенных гоех-фильтров и отличающийся доопределением частотно- и уровне-зависимого коэффициента расширения на диапазон потери слуха 65–90 дБ, что обеспечивает моделирование для всех степеней тугоухости, включая IV;

2. Метод аппроксимации частотных характеристик банка гамматон-фильтров неравнополосным косинусно-модулированным банком фильтров, отличающийся построением фильтра-прототипа в виде произведения логистических сигмоидальных функций, что обеспечивает монотонные спады характеристик банка и снижает вычислительную сложность реализации 100-канального банка примерно в 8 раз;

3. Метод повышения скорости выполнения алгоритма фазового преобразования на основе структуры сдвоенного фазового звена, сохраняющей критический путь одиночного звена и вычисляющей выходы каскада из двух звеньев за один проход, что ускоряет вычисление фазового преобразования в 1,81 раза;

4. Метод повышения разборчивости речи, основанный на переносе неслышимых высокочастотных компонент в область остаточного слуха с коррекцией их уровней по кривым равной громкости и отличающийся

адаптивным выбором ширины целевой полосы из условия равенства ширин полос в психоакустической шкале барков, что повышает разборчивость речи в среднем на 6 %.

Отзыв на диссертацию принят на расширенном заседании кафедры радиофизики и цифровых медиа технологий Белорусского государственного университета 22.09.2026, протокол № 2, по результатам открытого голосования простым большинством голосов членов расширенного заседания кафедры, имеющих ученые степени.

На заседании присутствовали:

к.т.н., доцент Максимович Е.С., к.т.н., доцент Демидчик В.И., к.ф.-м.н., доцент Малый С.В., к.ф.-м.н., доцент Хейдоров И.Э., к.т.н., доцент Яскевич В.Э., старший преподаватель Беленькая С.С., старший преподаватель Курило В.С., старший преподаватель Полешук Н.Н., заведующий учебной лабораторией, к.т.н., с.н.с. Любецкий Н.В., д.ф.-м.н., профессор Афоненко А.А., к.ф.-м.н., доцент Людчик О.Р., к.т.н., доцент Семенович С.Н., к.т.н., доцент Стецко И.П., к.ф.-м.н., доцент Скакун В.В., инженер Черепанова Ю.В.

Результаты голосования: «за» – 11; «против» – 0; «воздержались» – 0.

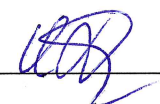
Председатель расширенного
заседания кафедры радиофизики
и цифровых медиа технологий
к.т.н., доцент

 Е.С. Максимович

Секретарь расширенного
заседания кафедры радиофизики
и цифровых медиа технологий

 Ю.В. Черепанова

Эксперт, подготовивший проект отзыва
к.ф.-м.н., доцент

 И.Э. Хейдоров