

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.199.01 НА БАЗЕ
ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО
УЧРЕЖДЕНИЯ НАУКИ САНКТ-ПЕТЕРБУРГСКОГО ИНСТИТУТА
ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ
РОССИЙСКОЙ АКАДЕМИИ НАУК ПО ДИССЕРТАЦИИ
НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ КАНДИДАТА НАУК

аттестационное дело № _____

решение диссертационного совета 24.11.2016 г. № 2

О присуждении Шиленкову Егору Андреевичу, гражданину Российской Федерации, ученой степени кандидата технических наук.

Диссертация «Структурно-параметрический синтез ортогонального речевого кодера, адаптивного критериям степени сжатия и разборчивости речи» по специальности 05.13.01 – «Системный анализ, управление и обработка информации (технические системы)» принята к защите 22 сентября 2016 г., протокол №1 диссертационным советом Д 002.199.01 на базе Федерального государственного бюджетного учреждения науки Санкт-Петербургского института информатики и автоматизации Российской академии наук, 199178, Россия, Санкт-Петербург, 14 линия ВО, дом 39, утвержден приказом Рособнадзора номер 2472-618 от 8 октября 2010 года.

Соискатель Шиленков Егор Андреевич, 1986 года рождения, в 2010 г. окончил магистратуру Юго-Западного государственного университета по направлению подготовки «Проектирование и технология электронных средств» (диплом № ВМА 0123456), в 2016 г. окончил очную аспирантуру в Федеральном государственном бюджетном образовательном учреждении высшего образования «Юго-Западный государственный университет». Удостоверение о сдаче кандидатских экзаменов №2249, выдано 3 февраля 2016 г. федеральным государственным бюджетным образовательным учреждением высшего образования «Юго-Западный государственный университет». В настоящее время Шиленков Егор Андреевич работает заместителем начальника 12-го отдела НИИ Радиоэлектронных систем в Федеральном государственном бюджетном образовательном учреждении высшего

образования «Юго-Западный государственный университет». По совместительству работает преподавателем кафедры космического приборостроения и систем связи.

Диссертация выполнена на кафедре космического приборостроения и систем связи Федерального государственного бюджетного образовательного учреждения высшего образования «Юго-Западный государственный университет» Министерства образования и науки Российской Федерации.

Научный руководитель – кандидат технических наук, старший научный сотрудник ПОТАПЕНКО Александр Михайлович, основное место работы: Федеральное государственное бюджетное образовательное учреждение высшего образования «Юго-Западный государственный университет», профессор кафедры космического приборостроения и систем связи.

Официальные оппоненты:

САРЫЧЕВ Валентин Александрович, доктор технических наук, профессор, акционерное общество «Научно-производственное предприятие «Радар ммс», заместитель главного конструктора;

КОРНЕЕВ Юрий Алексеевич, кандидат технических наук, старший научный сотрудник, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения», доцент «Института радиотехники, электроники и связи (Институт №2 ГУАП)».

Ведущая организация: Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», г. Санкт-Петербург в своем положительном заключении, подписанном Столбовым Михаилом Борисовичем, кандидатом технических наук, доцентом кафедры речевых информационных систем Университета ИТМО и утвержденном Матвеевым Юрием Николаевичем, доктором технических наук, заведующим кафедрой речевых информационных систем, указала, что диссертация Шиленкова Егора Андреевича является целостной и логически завершенной научно-квалификационной работой, в которой на основе выполненных исследований автором получено новое решение актуальной научно-технической

задачи, состоящей в повышении эффективности хранения и передачи речевой аудиоинформации во внешних запоминающих устройствах ЭВМ при высокой степени нестационарности свойств сигнала, содержащего русскую речь. Диссертационная работа отвечает требованиям, установленным в п. 9 Положения о присуждении ученых степеней, утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 № 842, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения ученой степени кандидата технических наук по специальности – 05.13.01 «Системный анализ, управление и обработка информации (технические системы)».

Соискатель имеет 20 опубликованных работ, в том числе по теме диссертации - 20 работ, опубликованных в рецензируемых научных изданиях - 20 работ, из них опубликованных в изданиях, рекомендуемых ВАК РФ, - 7.

Основные научные результаты опубликованы в 20 научных трудах общим объемом 4,25 п.л., из которых 14 статей объемом 3,1 п.л. выполнены в соавторстве, а 6 статей объемом 1,15 п.л. – лично; получены авторские свидетельства о государственной регистрации программы для ЭВМ. Наиболее значимые работы по теме диссертации:

1. **Шиленков, Е.А.** Способ построения частотно-временной маски для речевого кодирования на основе психоакустической модели слуха /Е.А. Шиленков // Проектирование и технология электронных средств. – 2013. - №4. Владимир, – С. 16-20.
2. **Потапенко, А.М., Шиленков, Е.А.** Методика построения динамических сигнально-кодовых конструкций с целью повышения пропускной способности канала связи / А.М. Потапенко, Е.А. Шиленков // Проектирование и технология электронных средств. – 2012. - №1. Владимир, – С. 12-16.
3. **Ефимова, О.В., Шиленков, Е.А.** Банк фильтров ортогонального речевого кодирования / О.В. Ефимова, Е.А. Шиленков // Материалы Всероссийской научно-практической конференции «Инфокоммуникации и информационная безопасность: состояние, проблемы и пути решения», Курск, 25-26 апреля 2014. – С. 113 - 118.

4. **Шиленков, Е.А.** Определение дескриптора и формата словаря сжатых данных / Е.А. Шиленков // Материалы Всероссийской научно-практической конференции «Инфокоммуникации и информационная безопасность: состояние, проблемы и пути решения», Курск, 25-26 апреля 2014. – С. 287-289.
5. **Шиленков, Е.А.** Методика дескриптора LZSSTACKER / Е.А. Шиленков // Материалы Всероссийской научно-практической конференции «Инфокоммуникации и информационная безопасность: состояние, проблемы и пути решения», Курск, 25-26 апреля 2014. – С. 18 - 22.
6. **Шиленков, Е.А., Хотынюк, С.С.** Линейный предиктор в ортогональном речевом кодере / Е.А. Шиленков, С.С. Хотынюк // Материалы Всероссийской научно-практической конференции «Инфокоммуникации и информационная безопасность: состояние, проблемы и пути решения», Курск, 25-26 апреля 2014. – С. 11 - 17.
7. **Шиленков, Е.А.** Кодирование и сжатие речевого сигнала в ТФОП / Е.А. Шиленков, А.М. Потапенко // Общество, современная наука и образование: проблемы и перспективы: сб. науч. тр. по мат-лам Междунар. заоч. науч.-практ. конф. - 2012: (Ч.10). Тамбов, - С. 159-161.

Оригинальность содержания диссертации составляет не менее 97% от общего объема текста; заимствованного материала, использованного в диссертации без ссылки на автора либо источник заимствования, не обнаружено; научных работ, выполненных соискателем ученой степени в соавторстве, без ссылок на соавторов не выявлено.

На автореферат диссертации поступило 6 отзывов, все отзывы положительные:

1) Омский государственный технический университет. Отзыв составил доктор технических наук, профессор, заведующий кафедрой «Средства связи и информационной безопасности» Майстренко Василий Андреевич. Замечания: нет логики в выборе размерности величин в таблицах и формулах, показанных в процессах сжатия без потерь. В итоге отображаемые результаты трудно воспринимаемы.

2) Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых. Отзыв составил профессор

кафедры радиотехники и радиосистем, доктор технических наук Левин Евгений Калманович. Замечания: Из текста реферата не ясно, относительно каких параметров несжатого потока речи определен коэффициент сжатия разработанного кодера равный 14.

3) Научно-исследовательский институт «Спецвузавтоматика» г. Ростов-на-Дону. Отзыв составил директор ФГАУ НИИ «Спецвузавтоматика», кандидат технических наук, доцент Хади Роман Ахмедович. Замечания: автором заявлено проведение структурно-параметрического синтеза, в процессе исследования им определен перечень параметров для спроектированной структуры, однако отсутствуют конкретные положения при проверке структуры на соответствие начальным требованиям и, в итоге, остаются неясными условия, которым должен удовлетворять результат синтеза.

4) Военно-воздушная академия им. проф. Жуковского и Ю.А. Гагарина. Отзыв составил начальник кафедры № 55, доктор технических наук, доцент Будников Сергей Алексеевич. Замечания: не установлено, какими параметрами разработанная структура (рис. 9) выгодно отличается от существующих аналогов.

5) Пермский национальный исследовательский политехнический университет. Отзыв составили заведующий кафедрой автоматики и телемеханики, доктор технических наук А.А. Южаков, доцент кафедры автоматики и телемеханики, кандидат технических наук И.И. Безукладников. Замечания: в автореферате не отмечено, какими конкретно характеристиками отличается русская речь и как эти особенности влияют на проектируемые кодеки.

6) Военно-космическая академия имени А. Ф. Можайского. Отзыв составили профессор кафедры Космических радиотехнических систем, доктор технических наук, профессор Мальцев Георгий Николаевич и докторант кафедры Космических радиотехнических систем, кандидат технических наук, доцент Козин Игорь Александрович. Замечания: математическая модель содержит операции с неравномерной шкалой квантования, но в результирующей структуре кодера нет разграничений или указаний битовой разрядности параметров процедур, что влечет за собой ошибки масштабирования.

Выбор официальных оппонентов и ведущей организации обосновывается тем, что д.т.н., профессор Сарычев В.А. является известным ученым практиком в сфере мультимедийных информационных продуктов, особое внимание уделяет применению мультимедийных информационных технологий, к.т.н., с.н.с., доцент Корнеев Ю.А. является известным специалистом в области исследования и разработки алгоритмов анализа речевых сигналов, ведущая организация, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», является известной как в России, так и за рубежом, научной деятельностью в направлении распознавания голоса, синтеза речи, записи и анализа, идентификации диктора.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

разработаны алгоритмы повышения степени сжатия речевого сигнала в системах передачи и хранения информации, учитывающие особенности психоакустического восприятия русской речи, реализующие совместную работу процедур линейного предсказания и перцептивного кодирования и обеспечивающие сохранение энергетических параметров голоса;

предложены математическая модель и методики компрессии как инструменты проведения структурно-параметрического синтеза речевого кодера, удовлетворяющего условиям соответствия заданному уровню качества воспроизведения, величине временной задержки и скорости потока данных;

доказана перспективность применения вариативных механизмов частотно-временного анализа, прореживания, переквантования и сжатия без потерь для компрессии энергетических параметров сигнала русской речи;

введены:

1) новые группы параметров управления механизмами речевого кодирования: первые позволяют адаптировать психоакустические модели к шумовой обстановке и восприимчивости слушателя, вторые обеспечивают вариативность процедур сжатия, регулирование которых изменяет степень компрессии, обе группы определяют качество воспроизведения;

2) новый термин линейного предиктора речевого сигнала, отличающийся от используемого в теории линейной модели противоположным свойством – вычленять незначимые дискреты сигнала, предопределяя будущие временные диапазоны редуцированных звуков.

Обработка экспериментальных данных производилась **по стандартным алгоритмам, без их модернизации.**

Теоретическая значимость исследования обоснована тем, что:

доказана эффективность применения разработанных механизмов сжатия цифрового сигнала, основанных на известных особенностях восприятия русской речи, их способность к адаптации предъявленным критериям;

применительно к проблематике диссертации результативно (эффективно, то есть с получением обладающих новизной результатов):

использованы современная методология модельно-ориентированного проектирования, параллельного объектно-ориентированного программирования, математического моделирования, теории вероятности и математической статистики, ортогональных преобразований, цифровой обработки сигналов;

изложены процедуры использования моделей, методик и алгоритмов сжатия с потерями и без, обусловлена степень вариативности их параметров;

раскрыты и конкретизированы: работа функциональных блоков, содержание информационных потоков, структурное взаимодействие и процессы объективной оценки работы ортогонального кодера. Выявлены зависимости изменения качества воспроизведения от выбора процедур сжатия и их параметров;

изучены: опыт международных научных сообществ в разработке кодеков, объективные и субъективные методы оценки качества воспроизведения, методики и концепции уплотнения информации в телекоммуникационных устройствах. Особое внимание уделено отличиям русского языка от языков романской группы;

проведена модернизация существующей математической и методической основы перцептивного кодирования, которая реализована в разработанной программной имплементации. Основные доработки и изменения:

1) разработана модель перцептивного псевдоквадратурного банка фильтров, основная особенность которой заключается в определении частотных диапазонов

сигнала, содержащих основную плотность энергии голоса диктора в текущем временном отрезке, выделяя тем самым выразительные звуки речи при сохранении нулевых фазовых сдвигов;

2) создана модель линейного предиктора, отличие которой состоит в проведении временного уплотнения сигнала путем прореживания редуцированных звуков русского языка, сохраняя при этом разборчивость фраз и понимание реплики в целом. Применение данных моделей обеспечивает преобразование нестационарного речевого сигнала для осуществления психоакустического анализа, позволяющего найти неслышимые диапазоны частот;

3) реализован механизм синусно-косинусного преобразования, возвращающий вещественный энергетический спектр голоса и не привносящий избыточность, обладает свойством масштабируемости при линейной фазочастотной характеристике;

4) методики сжатия без потерь адаптированы для потоковой и пакетной обработки с использованием реструктурированных кодовых деревьев и скользящего окна, что в совокупности с изменением масштаба энергетических параметров осуществляет компрессию.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

разработаны и внедрены (указать степень внедрения) следующие результаты диссертационной работы:

1) Математическая модель дискретных преобразований во временной и частотной области, реализующая анализ нестационарного сигнала; алгоритмы прореживания с помощью наложения частотно-временной маски и редуцированных предсказаний линейного предиктора; модифицированные методики сжатия ортогонального речевого потока без потерь. Данные результаты использованы в Юго-Западном государственном университете г. Курск на кафедре космического приборостроения и систем связи, в частности включены в качестве теоретических материалов в учебный курс «Цифровая обработка сигналов» по направлению подготовки «Инфокоммуникационные технологии и системы связи»

и «Теория радиотехнических сигналов» по направлению подготовки «Информационная безопасность телекоммуникационных систем».

2) Применение адаптированных дискретных ортогональных преобразований позволило существенно снизить количество вычислений при выполнении алгоритмов определения преграды в рамках выполнения НИР «Исследование научно-технических путей построения интеллектуальной системы управления движением автономной гусеничной платформы», проводимой в НИИ РЭС ЮЗГУ г. Курск, совместно с программными средствами имплементации алгоритмов сжатия и энергетического анализа дискретного сигнала;

определены возможности и перспективы практического использования полученных результатов диссертации при проектировании телекоммуникационного оборудования, в том числе специального назначения.

создано программное обеспечение речевого кодера, реализующее механизмы сжатия, которое позволяет осуществить анализ выходного звука на соответствие исходным данным и провести объективную и субъективную оценку выходного сигнала;

представлены предложения и направления для дальнейших научных исследований, в основу которых могут быть положены результаты структурно-параметрического синтеза кодера.

Оценка достоверности результатов исследования выявила:

для экспериментальных работ согласованность и воспроизводимость теоретических положений и выводов с результатами, осуществляемых разработанной программной имплементацией ортогонального речевого кодера, реализованной на сертифицированном современном оборудовании. Достоверность подтверждена проведением корреляционного анализа аудиокодеков и материалов научных работ по данному направлению, корректным применением методологического аппарата, содержащего использованные в исследовании теории и методы. Результаты диссертации прошли апробацию на международных конференциях, круглых столах и печатных докладах;

теория построена на широко известных методах исследований и техниках проектирования, достоверных данных и фактах, сопоставима с опубликованными экспериментальными данными по теме диссертации;

идея базируется на гипотезе об избыточности речевого сигнала для восприятия, основанной на закономерностях русского языка, при систематизации знаний отечественных и зарубежных авторов в области сжатия речи;

использованы программные и аппаратные средства аналогичных устройств и прототипов, для сопоставления с результатами структурно-параметрического синтеза функционально и параметрически;

установлено параметрическое отождествление результатов решения научно-технической задачи повышения эффективности хранения и передачи речевой аудиоинформации в условиях высокой степени нестационарности свойств сигнала с аналогичными показателями современных речевых кодеров. Функциональное сравнение выявило преимущество разработанного кодера в силу его вариативности. Объективная и субъективная оценка показали выигрыш в степени сжатия речи при учете особенностей языка.

Личный вклад соискателя состоит в:

1) проведении анализа современных технологий в области речевого кодирования и выборе гибридного подхода, включающего линейное предсказание и ортогональные процедуры, для синтеза структуры кодера;

2) постановке задачи структурно-параметрического синтеза для реализации ортогонального речевого кодера, адаптивного определенным критериям степени сжатия и разборчивости речи;

3) разработке математической модели блока полифазной фильтрации и линейного предиктора и построении механизмов их реализации;

4) создании модели и алгоритма наложения психоакустической маски;

5) структурировании системы сжатия без потерь по словарным и оконным методам;

6) реализации программной имплементации речевого кодера, позволяющей использовать методы и алгоритмы исследования для проведения экспериментов.

Диссертационный совет считает, что Шиленков Е.А. в своей диссертационной работе решил актуальную научную задачу разработки алгоритмов сжатия голосовых сообщений с учетом особенности русской речи.

На заседании 24.11.2016 г. диссертационный совет принял решение присудить Шиленкову Е.А. ученую степень кандидата технических наук.

При проведении тайного голосования диссертационный совет в количестве 21 человека, из них 6 докторов наук по специальности рассматриваемой диссертации, участвовавших в заседании, из 26 человек, входящих в состав совета, проголосовали: за 18, против 0, недействительных бюллетеней 3.

Председатель диссертационного совета

доктор технических наук,

член-корреспондент РАН

Юсупов Рафаэль Мидхатович

Ученый секретарь диссертационного совета

кандидат технических наук, доцент

Фаткиева Роза Равильевна

24.11.2016 г.