

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 212.233.05  
НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО  
АВТОНОМНОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО  
ОБРАЗОВАНИЯ «САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ»,  
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ  
ФЕДЕРАЦИИ ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ  
КАНДИДАТА НАУК

аттестационное дело № \_\_\_\_\_

Решение диссертационного совета от 6 июня 2017 г. № 7/17  
о присуждении Томчуку Кириллу Константиновичу, гражданину Российской  
Федерации, ученой степени кандидата технических наук.

**Диссертация** «Сегментация речевых сигналов для задач  
автоматической обработки речи»

**по специальности** 05.12.13 – «Системы, сети и устройства  
телекоммуникаций»

**принята к защите** 4 апреля 2017 года, протокол № 3/17,  
диссертационным советом Д 212.233.05 на базе Федерального  
государственного автономного образовательного учреждения высшего  
образования «Санкт-Петербургский государственный университет  
аэрокосмического приборостроения», Министерство образования и науки  
Российской Федерации, 190000, Санкт-Петербург, ул. Большая Морская,  
д.67, приказ №741/нк от 08.07.2015 г.

**Соискатель** Томчук Кирилл Константинович, 1986 года рождения,  
гражданин Российской Федерации. В 2009 году соискатель окончил  
Государственное образовательное учреждение высшего профессионального  
образования «Санкт-Петербургский государственный университет  
аэрокосмического приборостроения». В 2012 году окончил очную  
аспирантуру Санкт-Петербургского государственного университета  
аэрокосмического приборостроения, работает в должности старшего

преподавателя кафедры радиотехнических систем Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения», Министерство образования и науки Российской Федерации.

**Диссертация выполнена** на кафедре радиотехнических систем Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения», Министерство образования и науки Российской Федерации.

**Научный руководитель** – кандидат технических наук, Корнеев Юрий Алексеевич, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения», кафедра радиотехнических систем, доцент.

**Официальные оппоненты:**

1. Карпов Алексей Анатольевич, доктор технических наук, доцент, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский институт информатики и автоматизации Российской академии наук», лаборатория речевых и многомодальных интерфейсов, заведующий лабораторией;

2. Тропченко Андрей Александрович, кандидат технических наук, доцент, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», кафедра вычислительной техники, доцент;

дали положительные отзывы на диссертацию.

**Ведущая организация** – Общество с ограниченной ответственностью «Центр речевых технологий», г. Санкт-Петербург, в своем **положительном** заключении, подписанном Поляковым Алексеем Юрьевичем, кандидатом

технических наук, секретарем научно-технического совета, руководителем управления НИОКР, утвержденном Дырмовским Дмитрием Викторовичем, генеральным директором, указала, что диссертационная работа Томчука К.К. является целостной и логически завершенной научно-квалификационной работой. Диссертационная работа отвечает требованиям, установленным п. 9 Положения о порядке присуждения 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 05.12.13 – «Системы, сети и устройства телекоммуникаций».

**Соискатель имеет** 16 опубликованных работ, в том числе по теме диссертации 15 работ, опубликованных в рецензируемых научных изданиях 3. В работах представлено описание системного подхода к автоматической сегментации речевого сигнала, обозначен перечень задач сегментации и раскрыты особенности обеспечения их эффективного решения. Приведено описание разработанных алгоритмов сегментации, модификации алгоритма MFCC-параметризации, рассмотрено применение автоматической сегментации в системе модификации темпа воспроизведения фонограмм. Описаны результаты экспериментальных исследований. 1 работа опубликована без соавторов. В работах, опубликованных вместе с соавторами, личный вклад соискателя заключается в разработке методов и алгоритмов, а также непосредственном проведении экспериментальных исследований. Общий объем научных изданий составляет 7,8 печатных листов. Получены акты о внедрении результатов диссертационной работы от двух организаций: ЗАО «НПП «Иста-Системс» и ФГАОУ ВО «Санкт-Петербургский государственный университет аэрокосмического приборостроения». В диссертации отсутствуют недостоверные сведения об опубликованных соискателем ученой степени работах, в которых изложены основные научные результаты диссертации.

Наиболее **значительные научные работы** по теме диссертации:

1. *Томчук, К. К.* Многоуровневая временная сегментация речевых сигналов в задаче модификации темпа воспроизведения фонограмм / К. К. Томчук, А. Ю. Зилинберг, Ю. А. Корнеев // *Вопросы оборонной техники; Серия 16. Технические средства противодействия терроризму.* – М. : НТЦ «Информтехника», 2011. – С. 85–93.
2. *Томчук, К. К.* Иерархическая модель и базовые алгоритмы временной сегментации речевых сигналов / К. К. Томчук, А. Ю. Зилинберг // *Ученые записки Петрозаводского государственного университета.* – Петрозаводск: ПетрГУ, 2013. – №2(131). – С. 114–119.
3. *Томчук К. К.* Применение частотного маскирования при MFCC-параметризации речи на фоне шумов / К. К. Томчук // *Информационно-управляющие системы.* – Санкт-Петербург, 2016. – №3(82). – С. 8–14.
4. *Томчук К. К.* Системные вопросы многоуровневой временной сегментации речевых сигналов / К. К. Томчук, А. Ю. Зилинберг, Ю. А. Корнеев // *Сборник докладов Научной сессии ГУАП.* – СПб. : ГУАП, 2013. – С. 80–83.

**На диссертацию и автореферат поступили отзывы из 9 организаций (все отзывы положительные):**

1. ЗАО «СКБ Орион» (подписали заместитель генерального конструктора, доктор технических наук, профессор Охтилев М.Ю., ведущий инженер-программист, кандидат технических наук Ничипорович О.П., инженер-программист Охтилев П.А.; утвердил генеральный директор - генеральный конструктор Васильев И.Е.). Замечания: **1)** не пояснено влияние векторной интерполяции двух периодов колебаний основного тона, необходимой для модификации темпа речи, на подавление верхних частот речевого сигнала; **2)** в тексте автореферата присутствуют некоторые неточности и стилистические ошибки.

2. Кировский филиал АО «Концерн «Океанприбор» (подписал старший научный сотрудник, кандидат технических наук Бородин М.А.).

Замечания: **1)** не обозначено, возможно ли применение разработанных алгоритмов для других языков, кроме русского; **2)** не проведен анализ быстродействия разработанного алгоритма модификации темпа речи; **3)** не указано, какая модель шума использована при оценке помехоустойчивости разработанных алгоритмов.

3. ФГБОУ ВО «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (подписал профессор кафедры радиосистем и обработки сигналов, доктор технических наук, профессор Волков В.Ю.). Замечания: **1)** нет четкой классификации решаемых задач, что затрудняет чтение автореферата; **2)** из автореферата не вполне понятна математическая модель взаимодействия сигналов и шумов. В связи с этим не ясно использование термина отношение сигнал/шум; **3)** для классификационных задач использованы термины IDR, MR, не свойственные общей теории приема сигналов (например, на стр. 12).

4. ООО «НПП «Цифровые радиотехнические системы» (подписал заместитель директора центра разработок, кандидат технических наук, доцент, Антохин Е.А.). Замечания: **1)** недостаточно полно описаны детали работы предлагаемого алгоритма модификации темпа речи: не приведены конкретные методы обработки, применяемые к сегментам разных типов и не обозначены соответствующие парциальные коэффициенты; **2)** для анализа точности и надежности разработанного алгоритма ОТ-сегментации целесообразно в качестве эталона использовать сигнал с ларингографа, фиксирующего смыкания и размыкания голосовых связок.

5. ВУНЦ ВМФ «Военно-морская академия» (подписали главный научный сотрудник, доктор технических наук, профессор Бараненко А.А., старший научный сотрудник, кандидат технических наук, старший научный сотрудник, Светличная А.А., врио начальника, кандидат технических наук, Дегтярев А.Р.; утвердил заместитель начальника по учебной и научной работе, кандидат военных наук, доцент Карпов А.). Замечания: **1)** недостаточно четкая формулировка цели и решаемых задач в автореферате,

что затрудняет их понимание (например, п.3. научной новизны и п.2. выносимого на защиту научного результата, значимость предложенной доработки алгоритма MFCC); **2)** достигнутая эффективность применения предложенных алгоритмов (12%) недостаточно убедительна, так как сопоставима с возможной измерительной ошибкой и погрешностями, обусловленными изменчивостью условий наблюдения; **3)** в автореферате практически отсутствуют сведения по внедрению полученных научных результатов. Указана только применимость для соответствующей области знаний; **4)** вместо обоснования и развития положений, выносимых на защиту, в автореферате приведено описание содержания разделов диссертации; **5)** в автореферате не приведено соответствие диссертации паспорту научной специальности.

6. ООО «Ф-Лайн Софтвер» (подписал ведущий инженер по внедрению программного обеспечения, кандидат технических наук Балакшин П.В.). Замечания: **1)** в автореферате недостаточно внимания уделено вопросам применения результатов сегментации в конкретных речевых задачах; **2)** можно было провести исследование не только на базе фонограмм TIDIGITS.

7. ОАО «ЦНПО «Ленинец» (подписал главный научный сотрудник НИО-200, кандидат технических наук, старший научный сотрудник Большаков А.Н.). Замечания: **1)** в автореферате не объясняется выбор базы фонограмм TIDIGITS для проведения экспериментального исследования; **2)** Из текста автореферата неясно, как производится отслеживание исходного темпа речи диктора в алгоритме модификации скорости произнесения, и можно ли данный алгоритм использовать для нормализации темпа разных дикторов.

8. АО «НПП «ИСТА-Системс» (подписал руководитель Дирекции по исследованиям и разработкам, кандидат технических наук Кузнецов Г.Л.). Замечания: **1)** не поясняется, с чем связано ухудшение показателей MFCC-параметризации некоторыми модификациями при высоких отношениях

сигнал/шум (таблица 1); **2)** в таблице 1 значения в столбце RI несколько отличаются от результатов подстановки значений из соседнего столбца в приведенную автором формулу (5).

9. АО «ВНИИРА» (подписал доктор технических наук, профессор Вельмисов И.А.). Замечание: **1)** недостаточно полно описаны детали работы предлагаемого алгоритма модификации темпа речи: не приведены конкретные методы обработки, применяемые к сегментам разных типов, и не обозначены соответствующие парциальные коэффициенты.

**Выбор официальных оппонентов и ведущей организации обосновывается** их достижениями в области автоматического анализа речевых сигналов и разработки речевых систем, а также наличием соответствующих публикаций, в том числе в ведущих журналах и изданиях, определенных Высшей аттестационной комиссией при Минобрнауки России. Выбор ведущей организации ООО «Центр речевых технологий» обосновывается тем, что данная организация разрабатывает новейшие алгоритмические и аппаратные решения, являющиеся лидерами мирового рынка речевых технологий. При этом компания использует собственные научные разработки, подтвержденные патентами РФ и свидетельствами об официальной регистрации программ для ЭВМ.

**Диссертационный совет отмечает, что на основании выполненных соискателем исследований:**

**разработаны** алгоритмические и программные средства для автоматической многоуровневой временной сегментации речевых сигналов, предназначенные для использования в речевых системах различного назначения;

**предложены** методы повышения шумоустойчивости параметризации речевого сигнала кепстральными коэффициентами;

**доказана** перспективность использования отдельных алгоритмов обработки для разных типов сегментов сигнала с целью повышения эффективности решения речевых задач;

**введены** метрики трендов и разладок структур колебаний основного тона; методика сравнения эффективности работы одностипных алгоритмов сегментации.

**Теоретическая значимость исследования обоснована тем, что:**

**доказана** эффективность применения моделей речеобразования и речевосприятия для увеличения шумоустойчивости алгоритмов параметризации и сегментации речевых сигналов;

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

**использованы** методы проектирования и анализа программных средств, общие методы системного анализа, методы статистической обработки результатов эксперимента, цифровой обработки сигналов, спектрального анализа временных рядов, фонетики, психоакустики;

**изложены** идеи и методы дикторонезависимого выделения в речевом сигнале временных интервалов, содержащих характерные типы речевой активности;

**раскрыты** проблемы обеспечения высокой точности, надежности, быстродействия алгоритмов сегментации;

**изучены** факторы, влияющие на высокую вариативность речевого сигнала и дикторозависимость его параметров;

**проведена модернизация** алгоритма параметризации речевого сигнала мел-частотными кепстральными коэффициентами, за счет чего повышена эффективность такой параметризации в условиях воздействия шумов.

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

**разработаны и внедрены** алгоритмы автоматической разметки речевого сигнала на временные сегменты и программный интерфейс для анализа сегментов речевого сигнала, которые используются ЗАО «НПП «Иста-Системс» в ОКР «Создание специализированного аппаратно-программного комплекса для автоматизации процесса выполнения комплексных



исследований по материалам экстремистской направленности», шифр «Фоб», а также в учебном процессе кафедры радиотехнических систем Санкт-Петербургского государственного университета аэрокосмического приборостроения в рамках курсов «Цифровая обработка сигналов» и «Цифровая обработка сигналов и сигнальные процессоры в системах подвижной радиосвязи», о чем имеется 2 акта о внедрении;

**определены** перспективы применения разработанных алгоритмов при решении широкого класса речевых задач;

**создана** система практических рекомендаций для внедрения разработанных программно-алгоритмических средств в системах автоматической обработки речи;

**представлены** перспективные направления дальнейшего развития речевых систем за счет использования временной разметки на различные типы сегментов.

#### **Оценка достоверности результатов исследования выявила:**

**для экспериментальных работ** результаты получены с использованием имитационного моделирования на реальных речевых данных и согласуются с имеющимися практическими результатами в данной области;

**теория** построена на основе собственных исследований и согласуется с известными литературными данными;

**идея базируется** на анализе и обобщении результатов исследований современных отечественных и зарубежных ученых;

**использованы** речевые базы, программные средства аналогичных по назначению решений для сравнительного анализа с разработанными алгоритмами;

**установлено** качественное совпадение результатов, полученных в диссертационной работе, с ранее известными научными положениями в данной области;

**использованы** современные методики сбора и обработки исходной информации.

**Личный вклад соискателя состоит в** непосредственном выполнении всех этапов исследования: определении цели и постановке задач исследования, подготовке исследовательской базы, создании методического, алгоритмического и программного обеспечения, проведении экспериментальных исследований, обработке и интерпретации данных, формулировке выводов.

На заседании 6 июня 2017 года диссертационный совет принял решение присудить **Томчуку Кириллу Константиновичу** ученую степень кандидата технических наук по специальности 05.12.13 – «Системы, сети и устройства телекоммуникаций».

При проведении тайного голосования диссертационный совет в количестве 16 человек, из них 8 докторов наук по специальности 05.12.13, участвовавших в заседании, из 22 человек, входящих в состав совета, проголосовали: за присуждение ученой степени – 16, против присуждения ученой степени – 0, недействительных бюллетеней – 0.

Председатель диссертационного совета Д 212.233.05  
доктор технических наук, профессор



Крук Евгений Аврамович

Ученый секретарь диссертационного совета Д 212.233.05  
кандидат технических наук, доцент

Овчинников Андрей Анатольевич

«6» июня 2017 года