

МИНОБРНАУКИ РОССИИ

Федеральное государственное автономное  
образовательное учреждение высшего  
образования

«Санкт-Петербургский национальный  
исследовательский университет  
информационных технологий,  
механики и оптики» (Университет ИТМО)

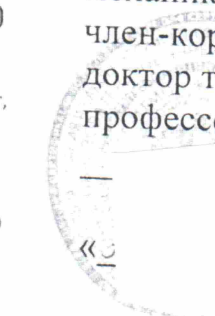
Кронверкский проспект, д. 49, г. Санкт-Петербург,  
Российская Федерация, 197101  
тел.: (812) 232-97-04 | факс: (812) 232-23-07  
od@mail.ifmo.ru | www.ifmo.ru

31.10.2016 № 88-01-18/610

УТВЕРЖДАЮ

Ректор ФГАОУВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»

член-корреспондент РАН,  
доктор технических наук,  
профессор,



### ОТЗЫВ

ведущей организации на диссертационную работу Шиленкова Егора Андреевича «Структурно-параметрический синтез ортогонального речевого кодера, адаптивного критериям степени сжатия и разборчивости речи», представленную на соискание учёной степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (технические системы)

#### 1. Актуальность темы

Обеспечение высокого качества и приемлемой разборчивости речевого сигнала при условии его максимального сжатия является важной задачей. Снижение временных затрат на выполнение процедуры цифрового кодирования способствует повышению производительности подобных систем, в связи с чем указанную процедуру часто реализуют на аппаратном уровне, а применяемые при этом методы психоакустического анализа и линейного предсказания звукового сигнала называют адаптивными. Существующие адаптивные алгоритмы психоакустического анализа и линейного предсказания имеют ряд противоречий, которые затрудняют их эффективное применение в современных системах связи и хранения информации. Один из таких недостатков – невозможность

одновременного их взаимодействия и применения непосредственно для речевого сигнала. Иными словами, известные методы цифровой обработки звука не обеспечивают достаточную гибкость среды компрессии звуковых данных, определяемую требованиями действующих рекомендованных стандартов электросвязи и аудио- и видеопередачи в условиях выбора качества и разборчивости. Особенным образом данный недостаток проявляется при обработке русской речи в силу её нестационарных частотных характеристик.

В связи с этим диссертационная работа Шиленкова Егора Андреевича, посвящённая решению научно-технической задачи повышения эффективности хранения и передачи речевой аудиоинформации во внешних запоминающих устройствах ЭВМ при высокой степени нестационарности свойств речевого сигнала является актуальным научным исследованием.

## **2. Степень обоснованности научных положений, выводов и рекомендаций**

Автором изучены и обобщены результаты исследований отечественных и зарубежных ученых по проблеме речевого кодирования, цифровой обработке звука и перцептивному анализу в целом.

Освещена история развития и современное состояние исследования по построению средств сжатия и потоковой обработки голоса в телекоммуникациях. Показаны возможности использования алгоритмов адаптивного предсказания и психоакустического анализа. Отмечена возможность комплексного использования различных ортогональных по сути методов сжатия для достижения более плотной компрессии речевого сигнала.

Список использованной литературы содержит 133 наименования, в том числе – 40 зарубежных авторов.

На основании анализа имеющейся информации по техникам сжатия звука, собранной в работах известных авторов и мировых лидеров в области цифровой обработки – группы MPEG, диссертант выделил ортогональные процедуры преобразования цифрового сигнала в качестве перспективных для сжатия речи; определил их первичные структуры; выявил признаки избыточности в случае

кодирования голоса, провел адаптацию выражений и методик и систематизировал их в математическую модель.

Для подтверждения теоретических положений автором проведены многоплановые экспериментальные исследования, включающие применение как методов объектно-ориентированного программирования, так и модельно-ориентированного проектирования с последующей квалифицированной статистической обработкой результатов. Несомненно, важным достижением работы является построение средства кодирования речи, обеспечивающее проведение пошагового анализа результата при помощи инструментов объективного (построение энергетических спектров и осциллограмм) анализа и субъективного анализа (прослушивание). Показана достоверная ассоциация выходного речевого потока по сравнению с входным и приемлемым качеством воспроизведения.

Достоверность экспериментальных данных обеспечивается использованием современных средств и методик проведения исследований. В работе диссертант грамотно использует методы математического анализа, соответствующие характеру фактического материала и поставленным целям и задачам исследований.

### **3. Новизна и достоверность основных научных результатов**

Детальное рассмотрение диссертационной работы Шиленкова Е.А. позволяет сделать вывод о том, что автором выбран и решён круг задач, представляющих новизну, получен ряд оригинальных научных результатов. Содержание работы характеризуется узкой направленностью исследований и разработок, достоверность которых подтверждается соответствующими документами о внедрении результатов, наличием детально представленных и подробно описанных методик и алгоритмов. Основные результаты диссертации пригодны для дальнейшей технологической проработки.

Полученные диссертантом результаты непротиворечивы, аргументировано обоснованы, базируются на грамотном применении положений и методов математической логики, теорий: множеств и графов, вероятностей и



математической статистики, систем и сетей массового обслуживания, проектирования устройств ЭВМ и систем, цифровой дискретной обработки, а также подтверждаются совпадением теоретических выводов с результатами имитационного моделирования.

#### **4. Теоретическая значимость**

Теоретическая значимость работы обусловлена ее новизной и заключается:

- в постановке и решении проблемы создания алгоритмов и методов речевого кодирования, позволяющих достичь компактного представления сигнала, содержащего русскую речь;

- в развитии математической модели, основанной на учёте нестационарности свойств сигнала, содержащего русскую речь, и особенностей его восприятия, позволяющей оценить изменение качества (разборчивости) сигнала в зависимости от состава и параметров процедур преобразования при различной степени сжатия;

- в расширении основ кодирования речевого сигнала, базирующихся на непрерывном анализе его свойств и динамическом управлении составом и параметрами процедур сжатия, позволяющих обеспечить структурно-параметрическую адаптацию речевого кодера при заданном качестве.

#### **5. Практическая значимость и внедрение результатов**

Практическая значимость полученных диссертантом результатов определяется тем, что они способствуют эффективному использованию аппаратных средств адаптивного сжатия речи и их внедрению в производство телекоммуникационного оборудования. Среди общего числа особенно выделяются:

- структурная схема речевого кодера, отличающегося применением линейного и ортогонального алгоритмов сжатия, а также модифицированных методик сжатия без потерь. Кодер позволяет сократить объем передаваемых или хранимых аудиоданных при сохранении заданного качества воспроизведения;

- программное обеспечение речевого кодека, позволяющее провести проверку корректной работы алгоритмов и методик сжатия речевого потока, а также оценить качество выходного звука на разборчивость и выразительность.

Результаты диссертационной работы Шиленкова Е.А. обсуждались на международных и российских конференциях и круглых столах, опубликованы в свободной печати, в том числе в известных журналах, входящих в перечень ВАК. Объем и содержание представленных публикаций по теме исследования в целом отражают основные положения, вынесенные на защиту.

Научные результаты, полученные в работе Шиленкова Е.А., можно считать новыми. Они создают основу для проектирования устройств кодирования сжатия цифровых сигналов, предназначенных для работы в составе системы хранения и передачи русской речи. На протяжении диссертационного исследования автор стремился сохранить однородность и прозрачность научных результатов и положений.

## **6. Замечания по диссертации**

Из недостатков работы можно отметить следующее:

1. Отсутствует оценка качества и разборчивости речевых сигналов, подвергшихся процедурам кодирования и декодирования. Поэтому выигрыш от их использования по сравнению с известными и стандартными процедурами неясен.

2. Современные средства сжатия медиа-контента активно используют вейвлет-преобразования, однако автором данный аспект ни разу не затронут.

3. Из текста диссертации не ясно, каким образом произведена оценка задержки прохождения сигнала в синтезированном кодере и как она оценивается.

4. В диссертации и автореферате имеются трудно воспринимаемые стилистические обороты и словосочетания, что вызывает дискомфорт в понимании логики исследований.

Однако, указанные недостатки не носят принципиального характера и не снижают научной и практической значимости диссертационной работы.

## 7. Заключение

Диссертация Шиленкова Егора Андреевича является целостной и логически завершённой 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 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 присуждение ученой степени кандидата технических наук по специальности – 05.13.01 «Системный анализ, управление и обработка информации (технические системы)».

Заключение рассмотрено на заседании кафедры речевых информационных систем федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», протокол заседания № 3 от «25» октября 2016 года.

Заключение составлено:

Столбов Михаил Борисович

к.т.н., доцент кафедры речевых информационных систем  
Университета ИТМО