

## ОТЗЫВ

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

### 1. Актуальность работы

Работа Шиленкова Е.А. посвящена важному направлению научных исследований – увеличению степени сжатия данных в каналах передачи информации. Успехи в сфере цифровой обработки позволяют реализовать алгоритмы и методы преобразований речевых сигналов, с помощью которых повышается качество воспроизведения голоса при существенном сокращении количества передаваемых данных, что в диссертации представлено как объект исследования. Разработка таких алгоритмов позволяет создавать оборудование для двустороннего общения между абонентами с низкими временными задержками в канале передачи и настраиваемом уровне качества воспроизводимого сигнала. Под качеством автор понимает разборчивость фраз диктора и выразительность его голоса. Благодаря использованию цифровых сигнальных процессоров, аппаратно поддерживающих обработку речи, существенно повышается количество абонентов телекоммуникационных систем, а за счёт появления программных продуктов возрастает число сервисов онлайн телефонии. Внедрение разрабатываемых методов и алгоритмов в системы связи весьма перспективно с экономической точки зрения.

Одной из важных решаемых задач цифровой обработки голоса является его компрессия, позволяющая сжать поток речевых данных в несколько раз, при этом оставляя уровень качества и задержки сигнала на заданном уровне. Способы достижения потокового сжатия по-разному влияют на данные показатели. В связи с этим применение перспективных подходов к уплотнению речевых сигналов требует

взаимной адаптации и согласования с нестационарными характеристиками голоса диктора, что составляет предмет исследования .

Стремление диссертанта повысить степень сжатия речевого сигнала выразилось в создании структуры вокодера, использующего техники адаптивного предсказания и психоакустического анализа. Отличительной чертой проведенных исследований является ориентация на русскую речь, которая отличается спектральными и временными характеристиками, её произношение и восприятие имеют уникальные лингвистические особенности, которые не учитываются или упрощенно аппроксимируются при реализации мультязычных кодеров.

Таким образом, с учетом вышесказанного, диссертационное исследование Шиленкова Е.А., направленное на разработку алгоритмов и методик обработки русской речи и последующий синтез структуры вокодера, обеспечивающего адаптивное изменение внутренней структуры и параметров процедур обработки для выбора качества и степени сжатия речевой информации, является актуальным и перспективным.

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

Научная задача диссертации Шиленкова Е.А. заключается в повышении эффективности хранения и передачи речевых сигналов в цифровых трактах связи при высокой степени нестационарности свойств речевого сигнала. Под повышением эффективности в диссертационной работе понимается уменьшение объема хранимых и передаваемых аудиоданных при сохранении заданного качества воспроизведения голоса диктора.

При решении научной задачи автором использованы известные методологические подходы к проведению научной работы. Среди них следует выделить два: для математического и визуального решения задач применён метод модельно-ориентированного проектирования, а для обеспечения устойчивости и управляемости системы – методология объектно-ориентированного программирования. Обоснованность полученных при этом результатов подтверждается прототипной реализацией корректно работающего кодера и успешным математическим моделированием с использованием методов цифровой обработки сигналов, теории вероятности, математической статистики, методов кодирования.

Предложенные решения создают условия для создания анализирующего устройства речевого кодирования с вариативной функциональной структурой, позволяющего отождествлять выбор параметров преобразований с изменением качества воспроизведения и степени сжатия.

### **3. Оценка новизны и достоверности**

К основным признакам новизны результатов можно отнести: математические модели, обеспечивающие проведение системного анализа сигнала с русской речью на предмет нахождения в нём неслышимых компонентов, отдельные интерпретации преобразований время-частота и алгоритмы компрессии.

Достоверность результатов обеспечена корректно выполненным математическим моделированием, а также соответствием экспериментальных звуковых сигналов и частотных спектров, генерируемых программным кодеком, основным теоретическим положениям и выводам, представленным в диссертации.

Теоретические положения диссертационных исследований опубликованы в научных статьях, обсуждались и были положительно оценены на всероссийских и международных научно-технических конференциях и круглых столах.

### **4. Теоретическая значимость**

В результате решения научно-технической задачи Шиленковым Е.А. получены оригинальные решения, которые могут быть использованы при разработке кодеков в составе телекоммуникационных средств передачи информации. Среди них в теоретическом плане следует отметить:

- методики и алгоритмы сжатия данных, реализующие математическую модель обработки нестационарных речевых сигналов и особенностей его восприятия, адаптируемые для голоса диктора и позволяющие обеспечить управление коэффициентом сжатия и качеством воспроизведения сигнала;
- структурно-функциональную организацию речевого кодера, реализующего операции временного и частотного маскирования, линейного предсказания и



ортогональных вычислений, а также методику сжатия речи, отличающуюся наличием вариативных входных параметров отдельных процедур и позволяющую осуществлять гибкое регулирование между объемом выходных сжатых данных и выбором качества воспроизведения речи.

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

Практическая значимость подтверждается оригинальностью и новизной полученных результатов, перспективность использования которых заключается в более экономном использовании информационных ресурсов, обеспечивающем достижение положительного эффекта при внедрения разработанных методов и алгоритмов в практику.

Использование спроектированной структуры для сжатия сигнала обеспечивает степень компрессии до 14 раз при сохранении приемлемой разборчивости. Данный параметр выгодно отличается от представленных аналогов. Однако выигрыш от внедрения получается только при общении на русском языке.

Практически значимыми представляются выводы автора об итогах применения инструментов анализа к оцифрованной русской речи. Сделано смелое предположение о возможности сократить объём информации, используя невыразительность восприятия русских слов. В силу данного довода автором разработаны процедуры сжатия с потерями, которые обеспечили степень сжатия до четырёх раз.

Результаты работы использованы в отчётах по научно-исследовательским работам и в разработке учебных планов и методических указаний по дисциплинам связи, а также применяются автором при преподавании ряда дисциплин.

Автором получены акты о внедрении выполненных научно-методических разработок.

## **6. Замечания по диссертационной работе**

Проанализировав текст диссертации Шиленкова Е.А., возникает следующий ряд вопросов и замечаний:

1. Проведённые исследования в направлении адаптивного линейного предсказания речевого сигнала имеют трудно интерпретируемые результаты, их использование для разработки кодера освещено недостаточно. Необходимо пояснить выбор автора в пользу применения спектральных операций при сжатии.

2. Автор использует трудно соотносимые формы представления сигнала в разработанных методиках, прибегая к булевой логике, арифметическим выражениям и машинной кодировке чисел в одном ряду. Процесс преобразования данных из формата WAV в дробный дискретный вид и обратно не освещён.

3. В результирующей схеме кодера показан блок квантователя. Он описывается как средство непосредственного уменьшения разрядности кода, но ни модель, ни механизм его работы не раскрываются.

4. Математическая модель цифровой фильтрации является интерпретацией стандартной функции свертки, и не содержит обратных связей, обеспечивающих крутизну передаточной характеристики на уровне современных эквалайзеров. Использование такой фильтрации требует дополнительной аргументации.

5. Диссертация и автореферат не лишены погрешностей стилистического характера, некоторые формулировки перенасыщены оборотами и явно затянуты, в тексте встречаются орфографические ошибки и опечатки.

Сформулированные замечания по диссертационной работе отчасти затрудняют восприятие и понимание хода исследований, однако не носят принципиального характера, не влияют на общую положительную оценку диссертационной работы и не снижают вклад автора.

## 7. Заключение

По содержанию, выводам и результатам диссертация Шиленкова Е.А. является целостной и логически завершенной научно-квалификационной работой. Результаты работы создают научно-методическую основу проектирования средств кодирования потокового речевого сигнала в составе систем инфокоммуникационного взаимодействия абонентов.

В диссертационной работе решена актуальная научная задача разработки, исследования и адаптации методов и алгоритмов кодирования звука, позволяющих

оперативно согласовывать параметры процедур сжатия речевого сигнала для сохранения уровня разборчивости голоса при понижении скорости потока данных.

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.

Шиленков Е.А. заслуживает присуждение ученой степени кандидата технических наук по специальности - 05.13.01 – «Системный анализ, управление и обработка информации (технические системы)».

*Сведения о составителе отзыва:*

Корнеев Юрий Алексеевич, кандидат технических наук, старший научный сотрудник,