



АФАНАСЬЕВ АНДРЕЙ АЛЕКСЕЕВИЧ

**МОДЕЛИ И МЕТОДЫ АНАЛИЗА И ОБРАБОТКИ
РЕЧЕВОГО СИГНАЛА В СИСТЕМАХ СВЯЗИ**

Специальность 05.12.13 – Системы, сети и устройства телекоммуникаций

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени
доктора технических наук

Орел - 2018

Работа выполнена в Федеральном государственном казенном военном образовательном учреждении высшего образования Академия Федеральной службы охраны Российской Федерации, г. Орел.

Научный консультант: **Левин Евгений Калманович**
доктор технических наук, доцент, профессор кафедры Радиотехники и радиосистем ФГБОУ ВО "Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых"

Официальные оппоненты: **Иванов Александр Иванович,**
доктор технических наук, доцент, начальник лаборатории, биометрических и нейросетевых технологий АО «Пензенский научно-исследовательский электротехнический институт», г. Пенза.

Кириллов Сергей Николаевич
доктор технических наук, профессор, заведующий кафедрой Радиоуправления и связи ФГБОУ ВО «Рязанский государственный радиотехнический университет», г. Рязань.

Шаврин Сергей Сергеевич
доктор технических наук, доцент, исполняющий обязанности декана факультета «Сети и системы связи» ФГБОУ ВО «Московский технический университет связи и информатики», г. Москва.

Ведущая организация: Открытое акционерное общество научно-производственное предприятие "Рубин", г. Пенза.

Защита состоится 17.10.2018 г в 14.00 на заседании диссертационного совета Д 212.025.04 при Владимирском государственном университете имени Александра Григорьевича и Николая Григорьевича Столетовых по адресу: 600000, г. Владимир, ул. Горького, д.87, корп.3, ауд. 301-3.

С диссертацией можно ознакомиться в научной библиотеке ВлГУ и на сайте <http://diss.vlsu.ru>

Автореферат разослан 10 июля 2018 г.

Отзывы на автореферат в двух экземплярах, заверенные печатью, просим направлять по адресу: 600000, г. Владимир, ул. Горького, д.87, ВлГУ, РТ и РС, ученому секретарю диссертационного совета Д 212.025.04.

Ученый секретарь диссертационного совета
доктор технических наук, профессор  А.Г. Самойлов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность диссертационных исследований. Постоянный рост объемов информации, циркулирующей в современных инфокоммуникационных системах, в том числе приложениях с низкоскоростной передачей данных, сопровождается сохранением роли речевого сигнала (РС) в межличностном общении.

В связи с тем, что речь является наиболее естественной формой общения телефонный трафик остается одним из основных, а множество инфокоммуникационных служб используют в своем оборудовании средства цифровой обработки РС. Наиболее бурно развивающейся областью цифровой обработки РС являются системы мобильной радиосвязи (сотовой, транкинговой, спутниковой, радио), системы Интернет-телефонии, специальные системы идентификации дикторов и речевого управления.

Большое количество работ отечественных и зарубежных авторов посвящено решению задач создания систем обработки РС, наиболее фундаментальными из них являются труды Михайлова В. Г., Сапожкова М. А., Назарова М. В., Прохорова Ю. Н., Ланнэ А. А., Грэя А. Х., Маркела Дж. Д., Рабинера Л. Р., Ситняковского И. В., Шафера Р. В., Шалимова И. А., Петровского А. А., Ковалгина Ю. А., Итакуры Ф., Рейдера Ч., Макхоула Дж. и других. В настоящий момент задачи повышения качества синтезированного РС по естественности ее звучания и узнаваемости абонента при сохранении требуемой скорости передачи, а также снижения средней скорости передачи при сохранении достоверности передаваемой информации остаются актуальными как с научной, так и с практической точки зрения. Системы обработки РС, функционирующие в инфокоммуникациях, выполняют важную функцию, задавая верхнюю потенциальную границу качества РС при его передаче, очень часто такая обработка производится в условиях наличия акустических шумов различной природы.

Существующие потребности в эффективности использования канальных ресурсов, а также необходимость учета новых особенностей в свете тенденций развития современных инфокоммуникационных систем, определяют перспективность научного поиска в области разработки новых и совершенствования существующих методов и алгоритмов обработки РС, которые непосредственно связаны с созданием эффективных речепреобразующих устройств (РПУ) – кодеков речи.

В диапазоне скоростей передачи до 16 кбит/с доминирующее положение занимают разновидности метода линейного предсказания (ЛП) РС. Повышению эффективности обработки алгоритмами на основе этого метода, в том числе и снижению

скорости передачи РС, препятствует недостаточная степень их адаптации к характеристикам кодируемых параметров, проявляющаяся представлением РС в устройствах телекоммуникаций без учета особенностей его формирования и независимой обработке элементов декомпозиции РС, к которым относятся параметры, описывающие передаточную функцию голосового тракта, и сигнал возбуждения фильтра-синтезатора РС.

В связи с изложенным, **научная проблема** диссертационного исследования формулируется следующим образом. Снижение скорости потоков данных РС при обеспечении требуемого его качества в условиях действия акустических помех обуславливает необходимость разработки моделей и методов анализа и обработки РС, учитывающих в большей степени особенности формирования РС и технологии его представления в системах телекоммуникаций.

Целью диссертационной работы является разработка моделей и методов обработки РС в условиях действия акустических помех, обеспечивающих снижение скорости потока данных при заданном качестве РС за счет более точного учета особенностей его формирования при существующих технологиях представления в системах телекоммуникаций.

Таким образом, **объектом исследования** является РС, подвергаемый обработке в системах связи.

В качестве **предмета исследования** выступают модели, методы и алгоритмы обработки РС, а также способы их реализации в системах связи.

Научная новизна исследования обусловлена тем, что в нем:

1) впервые предложены:

- комплексное техническое решение по формированию методологии понижения скорости кодирования РС при его передаче с сохранением качественных показателей синтезированной речи в условиях воздействия акустических помех;

- метод обработки РС, основанный на учете зависимости элементов декомпозиции РС ЛП, позволяющий уменьшать среднюю скорость передачи в канале связи при сохранении качественных показателей синтезированной речи;

- технологии адаптивной цифровой фильтрации РС в условиях акустического зашумления, отличающиеся применением полиспектрального анализа и позволяющие осуществлять шумоподавление при низких отношениях сигнал-шум;

2) развита теория применения моделей обработки РС со структурно-

параметрической адаптацией, на основе которых разработаны технические решения по выделению сегментов анализа различной длительности на однородных участках;

3) разработан новый подход к выделению пауз в РС при наличии акустического зашумления, основанный на применении технологий полиспектрального анализа;

4) проведено обоснование и применения новых моделей, методов и алгоритмов анализа и обработки РС с учетом принятых технологий его кодирования, учитывающих особенности формирования РС, которые могут быть использованы для совершенствования устройств обработки РС, функционирующих в системах связи.

Теоретическая ценность диссертационного исследования определяется развитием методов анализа РС, разработкой моделей и методов его обработки в системах связи в различных условиях функционирования с учетом ограничений на скорость передачи в КС.

Практическая значимость диссертационного исследования заключается в использовании разработанных и запатентованных методов и алгоритмов обработки РС, которые экспериментально проверены с использованием имитационного моделирования и обеспечивают:

- понижение средней скорости передачи в канале связи при сохранении качества РС;

- функционирование систем обработки РС в условиях шумовых акустических воздействий, при этом осуществлять фильтрацию в задаче шумоподавления со средним снижением показателя *Modified Bark Spectral Distortion* (MBSD) от 4,20 до 2.88 (среднее повышение субъективной оценки от 0,87 до 1,22 балла) в диапазоне входных значений отношения сигнал-шум (ОСШ) от 15 до -5 дБ.

Внедрение и использование разработанных технических решений дает возможность внести значительный вклад в хозяйственно-экономическое развитие и обороноспособность страны.

Обоснованность и достоверность научных положений, основных выводов и результатов диссертации обеспечивается за счет комплексного анализа состояния предметной области исследования, непротиворечивости теоретических выводов с результатами экспериментальной проверки предложенных моделей на основе имитационного моделирования, а также апробацией основных теоретических положений диссертации в печатных трудах и докладах на научных и научно-практических конференциях различных уровней.

Методы исследования. Для достижения цели исследования в работе использованы методы системного анализа, декомпозиции и агрегирования, статистического и корреляционного анализа, исследования операций, теорий информации, множеств и матриц, теории вероятностей и математической статистики, методы многокритериальной оптимизации, теории кодирования и ЦОС, аналитическое и имитационное моделирование.

На защиту выносятся следующие **положения диссертационного исследования:**

1. Структурно-параметрическая адаптация, отличающаяся от известных учетом особенностей формирования РС и технологий его кодирования, позволяет разработать методологию снижения средней скорости передачи данных в КС.

2. Учет зависимости элементов декомпозиции РС при построении модели его обработки при низкоскоростном кодировании на основе метода ЛП, позволяет осуществить снижение мощности пространства представления сигналов возбуждения.

3. Применение полиспектрального анализа при моделировании системы фильтрации РС в задаче шумоподавления позволяет разработать методы шумоподавления в РС для диапазона значений ОСШ от 15 до -5 дБ.

4. Метод выделения сегментов анализа РС различной длительности на однородных участках РС, позволяющий осуществлять его синтез при фиксированных значениях параметров передаточной функции голосового тракта.

5. Метод обработки РС, основанный на учете зависимости элементов его декомпозиции при ЛП, позволяющий уменьшать среднюю скорость передачи данных в КС при сохранении качественных показателей синтезированного РС.

6. Метод адаптивной цифровой фильтрации РС в условиях акустического зашумления, отличающийся применением технологий полиспектрального анализа и позволяющий произвести шумоподавление при отношениях сигнал-шум до 0÷-5дБ.

7. Алгоритмы анализа и обработки РС, позволяющие повысить характеристики функционирования телекоммуникационных устройств в системах инфокоммуникаций в условиях наличия акустических шумов, отличающиеся от известных учетом особенностей формирования и обработки РС, применением результатов полиспектрального анализа с целью шумоподавления, а также учетом взаимозависимости элементов декомпозиции РС при ЛП, устраняющие несоответствие между особенностями формирования РС и существующими в современных системах связи технологиями обработки РС.

Апробация результатов работы. Результаты диссертационного исследования обсуждались и докладывались в период с 2001 по 2017 годы на конференциях различного уровня, основные результаты представлены в [29-55].

Публикации. Основные положения диссертации опубликованы в 56 печатных работах, включая 12 публикаций в рецензируемых научных изданиях из перечня Минобрнауки РФ, издано учебное пособие (в соавторстве), получено 15 патентов на изобретение, 1 патент на полезную модель.

Результаты работы реализованы в 15 изобретениях и 1 полезной модели, в научно-исследовательских работах в интересах ФСО России, внедрены в деятельность ряда практических подразделений Спецсвязи ФСО России, ФГУП НИИ "Квант", НТЦ "Фобос НТ", ООО "Славсервис-Связь; использованы в учебном процессе Академии ФСО России, г. Орел, Рязанского государственного радиотехнического университета, г. Рязань.

Личный вклад автора в публикациях с соавторами кратко характеризуется следующим образом: в публикациях [1, 6, 18, 22, 29, 40, 41, 42, 44, 51, 53] представлен подход к построению моделей обработки РС на основе выделения переменных длительностей сегментов анализа на однородных участках; в [3, 27, 37, 50, 52] сформулированы технические решения по классификации сегментов РС, в [4, 13, 30, 31, 32, 33, 34, 55] разработаны теоретические принципы применения технологий построения систем обработки РС с переменной структурой и показаны прикладные возможности их реализации, в [2, 7, 10, 11, 16, 24, 28, 39, 43, 45, 47, 48, 49, 54] раскрыты и представлены теоретические подходы и практические особенности использования моделей и методов обработки РС на основе учета зависимости элементов декомпозиции РС при ЛП, в [8, 12, 26, 27, 46, 50] разработаны теоретические и практические аспекты полиспектрального анализа в приложении к задаче шумоподавления в РС, в [20, 25] раскрыты особенности реализации процедуры векторного квантования (ВК) параметров ЛП РС, в [5, 14, 15, 17, 19, 21, 23, 35, 36, 38, 56] представлены практические аспекты реализации систем обработки РС в системах связи, в [9] предложены подходы к проведению дальнейших исследований по построению индивидуально ориентированных систем обработки РС.

Структура и объем работы. Диссертационная работа изложена на 332 страницах текста, из них 292 страницы основного текста, содержит 79 иллюстраций и 21 таблицу, состоит из введения, шести разделов, заключения, списка сокращений, списка литературы (241 наименование), 2 приложений.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении дана общая характеристика работы, обоснована ее актуальность, определены научная новизна, теоретическая ценность и практическая значимость проведенного исследования. Сформулирована научная проблема, цель и задачи диссертационного исследования, практическая ценность полученных результатов, приведены основные положения, выносимые на защиту, и сведения об апробациях, публикациях и реализациях результатов работы.

В первом разделе проведен анализ существующих методов обработки РС для систем инфокоммуникаций, применяющихся в комплексах связи, определены требования к системам обработки РС в системах связи, рассмотрено состояние и структурированы проблемы обработки РС, осуществлена постановка задачи диссертационного исследования и проведен анализ методов ее решения.

Исследования в данной области показывают, что существующие системы обработки РС при низкоскоростном кодировании (менее 16 кбит/с) обеспечивают разборчивость речи примерно 85–90 %. Новым этапом развития мобильных персональных абонентских терминалов (МПАТ) является введение средств обработки РС в условиях зашумления. Анализ предоставления инфокоммуникационных услуг выявил факт воздействия акустических шумов, в этих условиях остается проблемой снижение скорости передачи РС при сохранении качества синтезированной речи. Условия функционирования систем обработки РС указывают на возможность совершенствования МПАТ в условиях интенсивных акустических воздействий (УИАВ) с отношением "сигнал-шум" (ОСШ) на входе кодера менее 15 дБ.

Анализ методов кодирования РС в инфокоммуникационных системах указал на доминирующее положение кодеков РС с гибридным методом кодирования на основе ЛП. Однако в современных стандартах кодирования РС:

- выделение сегментов анализа РС производится без учета особенностей его формирования в голосовом тракте;
- не учитываются зависимости между сигналом возбуждения фильтра синтеза и параметрами передаточной функции голосового тракта, а также условия функционирования системы обработки РС в условиях шумовых акустических воздействий;
- значительное число бит в кадре передачи отводится для представления параметров, описывающих сигнал возбуждения голосового тракта;
- не используется априорная информация о характере сигнала (отсутствует

классификация сегментов анализа по различным признакам;

– при низкоскоростном кодировании РС не используются возможности каскадного векторного квантования (ВК) и его разновидностей с применением новых технологий построения (обучения) кодовой книги (КК).

Указанные недостатки являются следствием противоречия между используемой в современных кодеках РС моделью речеобразования и объективно существующими зависимостями между параметрами передаточной функции синтезирующего фильтра ЛП и сигнала возбуждения.

В диссертации сформирован комплекс требований к перспективным системам обработки РС. К внешним параметрам системы обработки отнесены качество синтезированной речи A , скорость передачи информации V , задержка на обработку τ и вычислительная сложность Q алгоритма. Практика применения алгоритмов кодирования РС на скоростях до 16 кбит/с показывает доминирующее положение методов на основе ЛП с ВК его параметров. Диссертационная работа направлена на создание новых устройств обработки РС, в основе которых лежат адаптивные алгоритмы кодирования РС, обеспечивающие функционирование систем с переменной структурой и параметрами. Для создания такой адаптивной системы обработки целесообразно использовать результаты исследований моделей CELP (Code Excited Linear Prediction), так как в них заложено большинство перспективных решений.

Таким образом, необходимо развить теорию построения моделей и алгоритмов обработки РС и разработать методы его анализа, учитывающие особенности его формирования, принятые технологии низкоскоростного кодирования и возможности современных систем телекоммуникаций.

В связи с этим сформулированы частные задачи диссертационной работы:

– проанализировать актуальные пути развития систем обработки РС, учитывающие современные требования к телекоммуникациям;

– развить теорию анализа РС, на основе которой разработать математические модели, методы и алгоритмы его анализа и обработки, в том числе при наличии акустических шумов, учитывающие особенности формирования РС и принятые технологии его обработки и позволяющие осуществить снижение средней скорости передачи данных в КС;

– разработать методологию снижения средней скорости передачи данных;

– разработать рекомендации по технической реализации и применению новых методов обработки РС в системах инфокоммуникаций.

Системы обработки РС на основе ЛП с ВК, применяемые в современных инфокоммуникационных системах, рассматривают кодируемые параметры РС как независимые. Дальнейшее совершенствование этого класса систем может быть достигнуто на основе перехода к учету зависимостей между указанными параметрами. Такой подход создаст условия для более рационального распределения количества бит, выделяемых для представления кодируемых параметров.

Во втором разделе развиты теоретические аспекты анализа и обработки РС в системах связи, представлены новые подходы к построения систем его обработки на основе систем со структурно-параметрической адаптацией [4, 13, 21, 30, 31, 32, 33, 34, 55], разработан метод выделения сегментов анализа РС различной длительности на однородных участках РС [1, 6, 18, 22, 29, 41], рассмотрены подходы к расчету параметров синтезирующей системы при обработке РС [2, 16, 23, 40, 42, 44], показаны технические решения по определению линейных спектральных частот (ЛСЧ) на однородных сегментах анализа при обработке РС [3, 17, 51, 53]. В диссертации представлена обобщенная математическая модель системы обработки РС, основным отличием которой является введение переменной длины сегмента анализа речевых данных и представление параметров описывающих сигнал возбуждения и значений передаточной функции формирующего фильтра речевого тракта на основе учета их зависимостей между собой. Задача по созданию системы обработки РС при низкоскоростном кодировании аналитически может быть представлена (1):

$$V |_{A \geq A_0; \substack{\rightarrow \\ A_a}} \min, \text{ где } A'_a - \text{параметры адаптации} \quad (1)$$

Использование полного перечня параметров адаптации нецелесообразно, так как их увеличение приводит к повышению доли служебной информации, передаваемой при низкоскоростном кодировании РС, что сокращает информационный ресурс, выделяемый для передачи параметров подсистем кодирования.

Оптимизационную задачу (1) целесообразно упростить и привести к виду (2), где число параметров адаптации сокращается за счет выбора тех, которые оказывают доминирующее влияние на качество кодирования при адаптации структуры системы для заданных подсистем кодирования:

$$V_m |_{A \geq A_0}; \underset{A_a}{\rightarrow} \min, \text{ при } \vec{A}'_a = \{N_{\text{сег}}, N_k, F_{\text{вз}}, F_{\text{от}}, M_S, D_S, B_{SS}, N_{\text{код}} \{LP_i, VQ_j\}\}. \quad (2)$$

Параметры: f_d – частота дискретизации аналогового сигнала, $N_{\text{кд}}$ – число бит кодирования, $N_{\text{кв}}$ – число уровней квантования, $F_{\text{вз}}$ – функция взвешивания, N_k – число отсчетов на сегменте анализа, $N_{\text{код}}$ – число подсистем кодирования. Одновременная перестройка всего спектра параметров адаптации требует большого объема вычислений, что повлияет на алгоритмическую задержку и вычислительную сложность алгоритмов кодирования. Параметры адаптации (2) вычисляются на каждом сегменте анализа, а их учет при обработке может обеспечить существенное повышение качества обработки РС.

Разработанная математическая модель системы для обработки РС дала возможность по-новому сформулировать задачу разработки алгоритмов низкоскоростного кодирования РС. К типовым задачам разработки кодеков РС добавился учет возможности введения структурной адаптации системы за счет априорной и апостериорной информации о РС на этапе его анализа. Так, поиск рационального числа структур $N_{\text{код}}$, номеров субкодеков ЛП_{*i*}, ВК_{*j*}, участвующих в обработке конкретного субсегмента РС, N_k – количество отсчетов на сегменте анализа направлены на поиск рациональной структуры системы обработки для конкретного класса сегментов. Другие параметры связаны с поиском оптимальных параметров системы обработки РС. Определение оптимальных параметров системы реализуется путем поиска, так как ЛП и ВК математически описываются разнородно. Поиск оптимальных векторных квантованных наборов коэффициентов предсказания осуществляется с использованием процедуры "анализа через синтез".

В диссертации показано, что параметры РС изменяются с течением времени достаточно медленно, что позволяет рассматривать РС как стационарный на временных интервалах порядка 2,5– 60 мс, называемых "окнами" или сегментами, параметры передаточной функции голосового тракта также считаются неизменными, что дает возможность ввести в использование понятие "однородного сегмента РС". Таким образом, однородность сегмента РС связана с природой речеобразования и заключается в сохранении параметров передаточной функции голосового тракта при моделировании РС на основе метода ЛП.

Для различных звуков длительности устоявшегося процесса формируемой речи значительно превышают 10 ... 30 мс, причем длительность отдельных звуков речи

составляет от 20 до 350 мс. Максимально возможный интервал одновременно анализируемых данных составляет от 60 до 80 мс, что связано с требованиями по задержке при передаче РС, определяемой рекомендацией G.114 Международного союза электросвязи. Если в течение 60-80 мс не изменяется природа формирования РС, то новый анализ начинается при использовании данных о выделенных параметрах предыдущего сегмента. Увеличение длительности кратковременного анализа приводит к тому, что параметры передаточной функции системы обработки и сигнала возбуждения будут сохраняться на всем протяжении анализа, что эквивалентно сокращению средней скорости передачи РС. Так как РС является случайным нестационарным процессом, для которого характерны изменения параметров основного тона, значений коэффициентов, характеризующих передаточную функцию голосового тракта и вида сигнала возбуждения, то понижение средней скорости передачи возможно за счет реализации алгоритмов речевого кодирования с переменной скоростью. На рисунке 1 представлен вокализованный сегмент данных на интервале нескольких периодов основного тона. Одним из принципиальных моментов при формировании конечной границы интервала анализа речевого сегмента является вычисление периода основного тона и изменение значений интервалов корреляции [14]. При этом границы сегмента формируются на основании (3-5):

$$\Delta T \in (t_n \dots t_k) = \text{var}; \Delta T > 20 \text{ мс}; \frac{\Delta T}{T_{\text{от}}} \in \{N\}; \quad (3)$$

$$t_n = \Delta t n, \text{ при } \text{sign}(S(t_n)) + \text{sign}(S(\Delta t(n-1))) = 0; \quad (4)$$

$$t_k = \Delta t k, \text{ при} \quad (5)$$

$$\text{sign}(S(t_k)) = \text{sign}(S(t_n)) \& \text{sign}S(\Delta t(k-1)) = \text{sign}(S(\Delta t(n-1))) \& (t_k - t_n) > 20 \text{ мс},$$

где ΔT – длительность сегмента анализа; t_n – время начала сегмента анализа; t_k – время окончания сегмента анализа; $T_{\text{от}}$ – период основного тона; n – номер отсчета в начале сегмента; k – номер отсчета в конце сегмента; Δt – интервал дискретизации.

Использование выражений (3)–(5) позволяет получить сегменты, отсчеты которых в начале и окончании имеют одинаковые знаки конечной разности первого порядка. При этом начальный и конечный отсчеты во вновь сформированном сегменте будут иметь значения, близкие к изменению знака отсчета РС, что значительно уменьшает возможные искажения на стыках сегментов [15]. На шумоподобных сегментах формирование границы окончания сегмента анализа реализуется в соответствии с вышеизложенным подходом на интервалах соответствующих длительности

20 мс. Деление РС на сегменты одинаковой природы образования достаточно хорошо согласуется с квазистационарным характером РС на интервалах установившихся звуков. В предлагаемой модели формирование сегмента анализа РС переменной длины основано на корреляционных зависимостях речи.

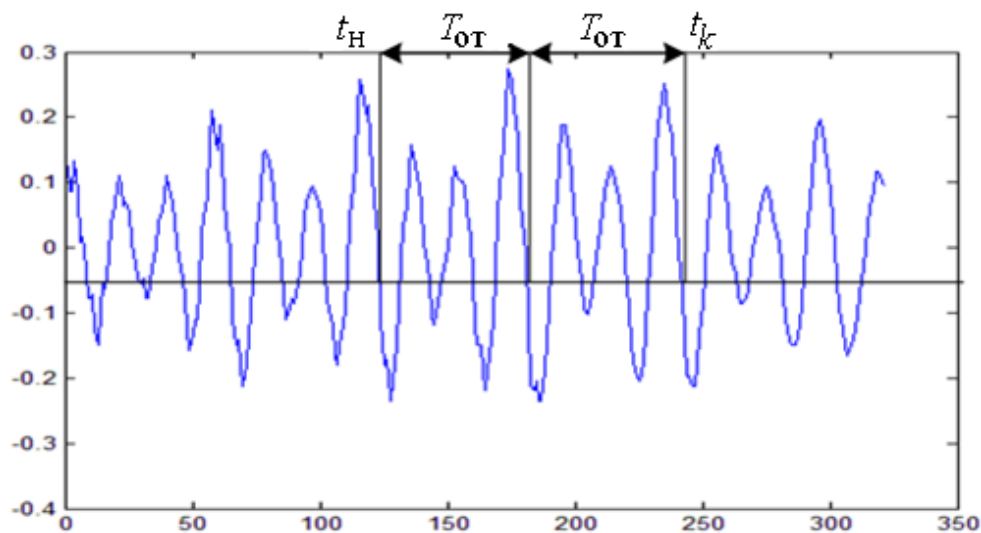


Рисунок 1– Иллюстрация динамического выделения участка квазистационарности

Таким образом, анализ РС позволяет сделать вывод о том, что возможно использование более длительных интервалов анализа, особенно на сегментах имеющих квазипериодическую вокализованную природу образования (рисунок1). Однако применение данных лишь о корреляционных связях между отсчётами не полностью устраняет статистическую избыточность РС при кодировании, при этом может возникать ошибка в определении участков однородности при переходе между звуками одного фонетического ряда [5, 19]. В задачах прикладного характера распределение значений РС априорно является случайным, более того сам вид распределения остаётся неизвестным. Увеличение границ анализируемого сегмента может быть основано на использовании более полной информации о статистических взаимосвязях между мгновенными значениями РС, получаемыми с помощью непараметрических методов при отсутствии данных о характере распределения. При адаптивном изменении границ сегмента анализа РС, количество наблюдаемых отсчётов, составляющих данный участок, априорно является неизвестной величиной. Оно может принимать различные значения в зависимости от особенностей произносимой речи. Соответственно количество наблюдений (отсчётов) составляющих анализируемый участок зависит в данном случае от исхода самих наблюдений. Следовательно, для определения длительности сегмента анализа необходимо применять методы последовательной проверки статистических гипотез. В данном случае предлагается метод,

основанный на последовательном критерии отношения вероятностей (критерий Вальда), который заключается в вычислении отношения вероятностей получения выборок (функций правдоподобия) на каждом этапе обработки. В силу отсутствия информации о виде распределения функция плотности вероятности, необходимая для вычисления правдоподобия, оценивается методом непараметрического ядерного сглаживания Парзена:

$$f_{\text{я}}(X) = \sum_{i=1}^N K\left[\frac{s}{h} - \frac{s_i}{h}\right], \quad (6)$$

где $f_{\text{я}}(X)$ – ядерная оценка плотности, $K(u)$ – ядерная функция, h – ширина окна, N – количество отсчётов составляющих анализируемый сегмент.

В качестве ядерной функции используется гауссов профиль:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, \quad (7)$$

Количество окон сглаживания для каждого начального сегмента анализа напрямую зависит от их ширины и является статистически обоснованным и определяемым с помощью информационного критерия Акаике (AIC).

$$AIC = -2 \ln(L(S, m)) + 2k, \quad (8)$$

где $L(S, m)$ функция правдоподобия для анализируемого сегмента, зависящая от числа окон, ширины окна, k – количество параметров (число окон, m).

Решение принимается в пользу минимального значения AIC . Получаемая в результате ядерного сглаживания кривая плотности распределения начального сегмента анализа $f_0(S, m)$, является основой для вычисления значения функции правдоподобия, в случае справедливости гипотезы H_0 (основной гипотезы), для последующих наборов отсчётов, получаемых на каждом этапе увеличения его длительности:

$$L_0(S, m) = \prod_{i=1}^N f_0(s_i, m), \quad (9)$$

где $f_0(s_i, m)$ – значение плотности вероятности в точке s_i .

В случае попадания значения отсчета s_i , между значениями $f_0(S, m)$, вычисление значения плотности в данной точке $f(s_i)$ производится с помощью линейной интерполяции. Выражение (8) с учётом (9) и изменённым знаком, является основой для определения оптимального числа окон ядерного сглаживания:

$$m_{onm} = \arg \max_m \left[\ln \left(\prod_{i=1}^N f_0(s_i, m) \right) - m \right], \quad (10)$$

Определённое таким образом число окон сглаживания используется для оценивания функции плотности распределения при справедливости гипотезы H_1 (альтернативной) на каждом этапе эксперимента. Основная гипотеза, в данном случае имеет вид: альтернативная $H_1: F(S, m) = F_1(S, m)$. При каждом увеличении сегмента на основе вновь полученного набора отсчётов РС производятся следующие операции:

1. Вычисляется значение правдоподобия при справедливости гипотезы H_0 по соотношению (9).
2. Производится ядерная оценка плотности распределения $f_1(s_i, m)$ (6).
3. Аналогично (9) вычисляется значение правдоподобия $L_1(S, m)$ по $f_1(s_i, m)$.

Далее вычисляется статистика критерия Вальда:

$$\ln \frac{L_1(S, m)}{L_0(S, m)} = \ln(L_1(S, m)) - \ln(L_0(S, m)) = Z[S], \quad (11)$$

При этом анализируемые выборки (сегменты) принадлежат к «эффективной» области $C_{эф}$ пространства выборок при выполнении $\ln B|_{n < n'} < Z[S] < \ln A|_{n < n'}$, где n – общее количество отсчётов в анализируемом сегменте, n' – максимально возможное число отсчётов в сегменте, с учётом ограничений G.114 МСЭ.

В случае $Z[S] < \ln B|_{n < n'}$, набор отсчётов принадлежит к "абсолютно эффективной" области $C_{эф}^{abc}$ (сегменту анализа). Оставшийся вариант при $Z[S] > \ln A$, определяет набор отсчётов принадлежащим к «неэффективной» области $C_{н.эф}$ (сегменту анализа). Ограничительные константы A и B , определяются на основании ошибок первого α и второго β рода:

$$\begin{cases} \ln A = \ln \frac{1-\beta}{\alpha} \\ \ln B = \ln \frac{\beta}{1-\alpha} \end{cases}, \quad (12)$$

где $\alpha = 0,002$ и $\beta = 0,001$ задаются в зависимости от требуемых условий анализа.

Критерием принятия решения о смене однородности при формировании сегментов РС является последовательный критерий отношения вероятностей Вальда, в основе которого лежит разделение m -мерного пространства отсчётов РС на три непересекающиеся области. При анализе речевых данных решение о продлении наблюдений, то есть о расширении границы сегмента анализа принимается в случае принадлежности анализируемой выборки двум из трёх областей пространства, "эф-

фективной" либо "абсолютно эффективной". Решение о прекращении увеличении длительности сегмента принимается в случае принадлежности выборки к "неэффективной" области, либо переходе между "абсолютно эффективной" и "эффективной" областями. При этом в случае перехода из "эффективной" в "абсолютно эффективную" область граница сегмента сохраняется на данном этапе. В случае же перехода из "абсолютно эффективной" в "эффективную" область граница сегмента определяется по данным предыдущего этапа.

Вычисляемые последовательно значения параметров передаточной функции голосового тракта на однородном сегменте анализа позволяют утверждать о высокой корреляции между ними. [1, 18]. Для описания передаточной функции синтезирующего фильтра на практике используются другие параметры, связанные с множеством коэффициентов ЛП взаимнообратными функциональными зависимостями, к которым относятся: нули фильтра ЛП, т. е. полюсы z_{*k} передаточной функции; коэффициенты отражения $\{r_i\}$ или коэффициенты частной корреляции; ЛСЧ. Согласно теории устойчивости, точность представления нулей полинома связана с частотными характеристиками полинома $A_M(z)$:

$$A_M(z) = 1 + \sum_{k=1}^M a_k z^{-k} = \prod_{k=1}^M (1 - z_{0k} z^{-1}), \quad (13)$$

где z_{0k} – нули полинома $A_M(z)$.

Нули полинома $A_M(z)$ должны находиться внутри круга единичного радиуса на плоскости Z , т. е. необходимо выполнение условия:

$$|z_{0k}| = |z_{*k}| = |r_{*k} e^{j\hat{\omega}_{0k} T}| = r_{*k} < 1, 1 \leq k \leq M, \quad (14)$$

где r_{*k} – радиус k -го нуля $A_M(z)$, M – порядок передаточной функции.

Однако в системах обработки РС значения z_{*k} не передаются, для описания передаточной функции наиболее предпочтительными являются ЛСЧ, также носящие название линейных спектральных корней (ЛСК). Они характеризуют значения максимумов амплитудно-частотной характеристики речевого тракта на близлежащих однородных сегментах РС, что может быть использовано при обработке РС в системах низкоскоростного кодирования речевых данных с переменной скоростью передачи. В диссертации обосновывается подход к вычислению ЛСК на однородных сегментах анализа, показано, что на данных сегментах значения ЛСК изменяются незначительно и не выходят за границы интервалов локализации, что упрощает их

поиск и подстройку. При анализе РС на однородных участках по значениям ЛСК формируется новый полином $A_M(z)$ с монотонной линейной ФЧХ:

$$\hat{A}_M(z) = z^R A_M(z) = z^R \sum_{k=0}^M a_k z^{-k} = \sum_{k=0}^M a_k z^{R-k}, \quad (15)$$

где $a_0 = 1$, $0,5M \leq R \leq M$.

В частотной области полином (17) может быть представлен в виде двух составляющих – суммой вещественной и мнимой частей:

$$\hat{A}_M(z)|_{z=e^{j\hat{w}}} = \text{Re}\{\hat{A}_M(z)|_{z=e^{j\hat{w}}}\} + j \text{Im}\{\hat{A}_M(z)|_{z=e^{j\hat{w}}}\} \quad (16)$$

$$\text{где} \begin{cases} \text{Re}\{\hat{A}_M(z)|_{z=e^{j\hat{w}}}\} = P(\hat{w}) = \sum_{k=0}^M a_k \cos[(R-k)\hat{w}] \\ \text{Im}\{\hat{A}_M(z)|_{z=e^{j\hat{w}}}\} = Q(\hat{w}) = \sum_{k=0}^M a_k \sin[(R-k)\hat{w}]. \end{cases}$$

Вещественная часть полинома $\hat{A}_M(z)|_{z=e^{j\hat{w}}}$ имеет нули на частотах $\hat{w}_{i|P}$, при которых $\arg\{P(\hat{w})\}$ принимает значения $(2k+1)\pi/2$, $k = 0, 1, \dots, d = R - 0,5$, а мнимая часть – на частотах $\hat{w}_{i|Q}$, при которых $\arg\{Q(\hat{w})\} = k\pi$, при $k = 0, 1, \dots, d = R$.

ЛСК могут быть получены из решения уравнений:

$$\begin{cases} P'(\hat{w}) = \sum_{k=0}^R b_k \cos^k[\hat{w}] = 0; \\ Q'(\hat{w}) = \sin[\hat{w}] \sum_{k=0}^{R-1} c_k \cos^k[\hat{w}] = 0. \end{cases} \quad (17)$$

Коэффициенты b_k , c_k зависят от значений КЛП. Задача нахождения значений ЛСК сводится к определению нулей тригонометрических полиномиальных функций (17), для определения нулей $\hat{w}_{i|P'}$ и $\hat{w}_{i|Q'}$ функций $P'(\hat{w})$ и $Q'(\hat{w})$ с заданной точностью на интервале $[0, \pi]$ необходимо применять численные методы решения уравнений. Определение нулей функции $P'(\hat{w})$ заключается в проверке с фиксированным шагом поиска диапазона $[0, \pi]$ на изменение знака функции в проверяемых точках. При этом значение шага поиска не должно превышать меньшего из минимальных расстояний между всеми соседними нулями функции $P'(\hat{w})$:

$$\Delta\hat{w} \leq \min\{\min|\hat{w}_{iP'} - \hat{w}_{i+1P'}|\}, \text{ где } \hat{w}_{i|P'}, \hat{w}_{i+1|P'} - i\text{-й и } (i+1)\text{-й нули функции,}$$

$i = 1, 2, P_1$; P_1 – количество нулей (порядок) функции $P'(\hat{w})$.

Вычисление нулей функции $P'(\hat{\omega})$ на найденных интервалах обычно проводится методом дихотомического поиска до достижения заданной точности. Выражение (17) можно привести к алгебраическому виду, произведя замену $x = \cos \hat{\omega}$. Интервал на частотной оси $[0, \pi]$ в области x отобразится в интервал $[-1, 1]$:

$$\begin{cases} P'(x) = \sum_{i=0}^R b_i x^i = 0 \\ Q'(x) = \sum_{i=0}^{R-1} c_i x^i = 0 \end{cases} \quad (18)$$

Нули двух полиномиальных функций определяются поочередно по диапазону $[-1, 1]$. Предлагаемый подход к определению направления поиска заданного интервала подразумевает наличие информации о значении опорных точек. Значение опорной точки принадлежит одному из множеств $\{\hat{\omega}_{i-1}, \hat{\omega}_i\}$ или $\{\hat{\omega}_{i-1}, \hat{\omega}_{i+1}\}$, причем на разных сегментах значения $\{\hat{\omega}_i\}$ различны. В устройствах обработки и передачи РС фиксированы как порядки функций $P'(\hat{\omega})$ и $Q'(\hat{\omega})$, так и количество их нулей, а в процессе анализа изменяются только коэффициенты полинома ЛП и соответственно коэффициенты функций $P'(\hat{\omega})$ и $Q'(\hat{\omega})$. Учитывая, что значения нулей функций принадлежат множеству $(0, \pi)$, предлагается использовать в качестве исходных значений статистические параметры нулей функций, определяемые по средневзвешенным значениям соответствующих нулей на подсегментах РС единого однородного сегмента РС. Такой метод позволяет определить интервал нахождения соответствующего нуля полиномиальной функции. В трудах Ланнэ А.А. показано, что при вычислении значений ЛСЧ функций $P'(\hat{\omega})$ и $Q'(\hat{\omega})$ значение последующего спектрального корня одного полинома всегда больше значения предыдущего корня другого полинома. Для нахождения корня $P'(\hat{\omega})$ целесообразно выбирать значение предыдущего найденного корня соответствующего полинома $Q'(\hat{\omega})$.

Предлагаемый подход позволяет определить значения ЛСК $P'(\hat{\omega})$ на однородных сегментах РС, используя их значения на подсегментах:

1) Начальная установка коэффициентов функции: $\{g_k\} = \{b_i\}$ и инициализация значений опорных точек: $\{\hat{\omega}_{\text{оп}}\} = \text{sort}[\{\hat{\omega}_{F_1 \text{ оп}}\}, \{\hat{\omega}_{F_2 \text{ оп}}\}]$, где $\text{sort}[*]$ – операция сортировки по убыванию.

2) Инициализация счетчика общего количества нулей $i = 1$.

3) Определение значения текущего проверяемого значения на подсегментах $\hat{w}_b = \hat{w}_{i_{оп}}$. Если $i = 1$, то перейти к пункту 5 алгоритма.

4) Проверка корректности выбора значения текущей точки.

Если значение текущей точки больше значения предыдущего найденного нуля $\hat{w}_b > \hat{w}_{i-1}$, то в качестве текущей точки принимается предыдущее значение $\hat{w}_b = \hat{w}_{i-1}$.

5) Вычисление значения функции в текущей точке по выражению

$$P(\hat{w}_b) = \sum_{k=0}^{M/2} g_k \hat{w}_b^k \quad (19)$$

6) Определение направления поиска i -го нуля функции, которое осуществляется следующим образом:

если $F(\hat{w}_b) > 0$ и i – нечетное, $\hat{w}_h = \hat{w}_b - \Delta\hat{w}_{шаг}$, $m = -1$;

если $F(\hat{w}_b) > 0$ и i – четное, то $\hat{w} = \hat{w}_b$, $\hat{w}_b = \hat{w} + \Delta\hat{w}_{шаг}$, $\hat{w}_h = \hat{w}$, $m = 1$;

если $F(\hat{w}_b) < 0$ и i – нечетное, то $\hat{w} = \hat{w}_b$, $\hat{w}_b = \hat{w} + \Delta\hat{w}_{шаг}$, $\hat{w}_h = \hat{w}$, $m = 1$;

если $F(\hat{w}_b) < 0$ и i – четное, $\hat{w}_h = \hat{w}_b - \Delta\hat{w}_{шаг}$, $m = -1$;

где m – идентификатор направления поиска.

7) Вычисление значения функции в краевых точках шага поиска. Значение функции в точках определяется выражениями:

$$P(\hat{w}_b) = \sum_{k=0}^{M/2} g_k \hat{w}_b^k, \quad (20)$$

$$P(\hat{w}_h) = \sum_{k=0}^{M/2} g_k \hat{w}_h^k. \quad (21)$$

8) Определение наличия нуля функции на шаге поиска по правилам:

– если $P(\hat{w}_b)P(\hat{w}_h) > 0$, то текущий интервал не содержит нуля функции, и следует перейти к пункту 12 алгоритма;

– если $P(\hat{w}_b)P(\hat{w}_h) < 0$, то текущий интервал содержит нуль функции, и следует перейти к пункту 10 алгоритма.

9) Определение новых границ поиска нуля функции по условиям:

если $m = -1$, то $\hat{w}_b = \hat{w}_h$, $\hat{w}_h = \hat{w}_h - \Delta\hat{w}_{шаг}$; если $m = 1$, то $\hat{w}_h = \hat{w}_b$, $\hat{w}_b = \hat{w}_b + \Delta\hat{w}_{шаг}$.

После переопределения границ перейти к пункту 7 алгоритма, иначе к п. 10.

10) Уменьшение найденного интервала определения i -го нуля функции за N шагов методом дихотомии.

11) Вычисление нуля функции на конечном интервале за один шаг метода линейной интерполяции:

$$\hat{w}_i = \hat{w}_H - \frac{(\hat{w}_B - \hat{w}_H) \cdot P(\hat{w}_H)}{F(\hat{w}_B) - F(\hat{w}_H)} \quad (22)$$

12) Проверка условия нахождения всех ЛСК: при выполнении завершить работу, если не выполнено перейти к пункту 3.

13) Определение математических ожиданий значений нулей функций $P'(\hat{w})$ на подсегментах единого сегмента одинаковой природы образования.

В третьем разделе рассмотрены теоретические основы присутствия зависимости элементов декомпозиции РС при его обработке на основе метода ЛП [7, 10, 49], разработана аналитическая модель системы обработки РС, учитывающая зависимости элементов его декомпозиции [11, 45, 47, 48], проведено обоснование возможности применения нового метода обработки РС, основанного на учете зависимости элементов декомпозиции [16, 24, 28, 39, 43, 54]. Использование принципов структурно-параметрической адаптации системы обработки РС и углубленный детальный анализ параметров ЛП при их кластерном представлении позволили разработать новые технические решения по созданию устройств обработки РС. Повышение порядка передаточных функций фильтров анализа и синтеза приводит к "обелению" сигнала остатка ЛП, который является наилучшим сигналом возбуждения. В задаче параметрического цифрового спектрального анализа на основе авторегрессионной модели линейное разностное уравнение формирующего фильтра определяется (23):

$$y(n) = - \sum_{m=1}^M a_m y(n-m) + u(n), \quad (23)$$

где $y(nT)$ – выходной сигнал, T – интервал дискретизации, a_m – коэффициенты фильтра.

Амплитудно-частотная характеристика фильтра определяется выражением

$$A(wT) = \frac{1}{\sqrt{(1 + \sum_{m=1}^M a_m \cos mwT)^2 + (\sum_{m=1}^M a_m \sin mwT)^2}}, \quad (24)$$

а спектральная плотность мощности определяется выражением

$$G(wT) = \frac{\sigma^2 \{u(n)\} T}{(1 + \sum_{m=1}^M a_m \cos mwT)^2 + (\sum_{m=1}^M a_m \sin mwT)^2}, \quad (25)$$

где W – круговая частота дискретного преобразования Фурье.

Повышение порядка модели в выражениях (23) – (25) приводит к более точным оценкам относительно анализируемого сигнала $\vec{S}$. В идеале $e^2(n) \rightarrow 0$ при $M \rightarrow \infty$.

На практике при реализации ЛП значение M ограничено, что приводит к возникновению сигнала $e(n)$, являющегося сигналом возбуждения фильтра синтеза модели ЛП. При этом сигнал $u(n)$ становится квазидетерминированным относительно множества $\{a_m\}$ и связан с ним соответствующими зависимостями.

В работе представлена аналитическая модель системы обработки РС, отличающаяся учетом зависимости элементов его декомпозиции при кодировании. Использование зависимостей элементов декомпозиции РС при ЛП позволяет усовершенствовать процесс его обработки. Наличие зависимостей между параметрами передаточной функции голосового тракта на сегменте и сигналом возбуждения объясняется особенностями постановки и решения обратной задачи ЦСА при фиксированном порядке формирующего фильтра. Представленная в работе математическая модель (рисунок 2) процедур обработки РС отличается учетом зависимости элементов декомпозиции РС. Пространства состояний кодера $\{\vec{C}_x\}$ из пространства анализируемых сегментов $\{\vec{S}\}$ формируется посредством отображения F_a .

$$F_a: \{\vec{S}\} \rightarrow \{\vec{C}_x\}. \quad (26)$$

Используемые в современных гибридных кодерах отображения F_{2_i} оптимизируют кодеки в условиях постоянства мощностей подпространств $\{Y_i\}$; а в разработанной системе необходим совместный поиск значений y_i и отображений F_{2_i} . В связи с этим состояния $\{\vec{C}_x\}$ кодера представляет собой декартово произведение:

$$\{\vec{C}_x\} = \{No_{F_{2_1}}\} \times \{No_{F_{2_2}}\} \times \dots \times \{No_{F_{2_i}}\} \times \{y_1\} \times \{y_2\} \times \dots \times \{y_i\}, \quad (27)$$

где $No_{F_{2_i}}$ – номер возможного отображения F_{2_i} , $y_i = \overline{1, \gamma}$, $i = \overline{1, 3}$.

Исключение i -х подпространств из структуры пространства представлений кодируемых параметров обеспечивается при $y_i = 1$. Учет зависимостей параметров голосового тракта ЛСЧ и сигнала возбуждения определяет введение оператора F_{B3} , реализующего процедуру выбора или установления соответствия между вектором ЛСЧ и конечным набором векторов возбуждения из i -й КК сигналов возбуждения

$$F_{B3}: \{Y_1\} \rightarrow \{Y_{2_i}\}. \quad (28)$$

Множество $\{\vec{C}_x\}$ конечно, а его мощность α определяется на основании компромисса между получаемым выигрышем при оптимизации и требуемыми затратами информационных ресурсов при передаче информации о текущем состоянии кодера. Координаты $\vec{X}_{1_0}$ и $\vec{X}_{2_0}$ случайного вектора считаются зависимыми, следовательно, их совместная плотность распределения имеет вид:

$$W(\vec{X}_{1_0}, \vec{X}_{2_0}, X_{3_0}) = W_1(\vec{X}_{1_0}) \cdot W_2(\vec{X}_{2_0} | \vec{X}_{1_0}) \cdot W_3(X_{3_0}). \quad (29)$$

С учетом случайного характера вектора $(\vec{X}_{1_0}, \vec{X}_{2_0}, X_{3_0})$ есть возможность усреднить величину искажения РС по всем анализируемым речевым векторам:

$$D(\vec{S}_0, \hat{\vec{S}}_0) = M[D(\vec{S}_0, \hat{\vec{S}}_0)] = \int_{\{\vec{X}_{1_0}\}} \int_{\{\vec{X}_{2_0}\}} \int_{\{X_{3_0}\}} D(\vec{S}_0, \hat{\vec{S}}_0) W(\vec{X}_{1_0}) \cdot W(\vec{X}_{2_0} | \vec{X}_{1_0}) \cdot W(X_{3_0}) \times \\ \times d\vec{X}_{1_0} (d\vec{X}_{2_0} | \vec{X}_{1_0}) dX_{3_0}. \quad (30)$$

Полученная величина $D(\vec{S}_0, \hat{\vec{S}}_0)$ характеризует среднее значение избранной меры искажения РС при обработке алгоритмом кодирования на основе метода ЛП. При классификации сегментов РС на $\{\vec{C}_x\}$ непересекающихся классов с характеристиками $c_x = 1, 2, \dots, C_x$ – номер класса, $P^{(C_x)}$ – вероятность принадлежности текущего сегмента речи к классу c_x при $\sum_{c_x=1}^{C_x} P^{(C_x)} = 1$; $W^{(C_x)}(\vec{X})$ – плотность распределения вероятности вектора $\{\vec{X}\}$ в классе c_x ; $D^{(C_x)}(\vec{S}_0, \hat{\vec{S}}_0)$ – средняя величина выбранной меры искажения для сегментов РС класса C_x , зависимость математического ожидания:

$$\bar{D}(\vec{S}_0, \hat{\vec{S}}_0) = M[D^{(C_x)}(\vec{S}_0, \hat{\vec{S}}_0)] = \sum_{c_x=1}^{C_x} P^{(C_x)} D^{(C_x)}(\vec{S}_0, \hat{\vec{S}}_0). \quad (31)$$

В случае конечности множества состояний кодера $\{\vec{y}_i^{(C_x)}\}$ мощности $\alpha^{(C_x)}$ можно считать состояния кодера, принадлежащие этому классу сегментов, равновероятными: $P(\vec{y}_\theta^{(C_x)}) = \frac{1}{\alpha^{(C_x)}}$, где $\theta = 1, 2, \dots, \alpha^{(C_x)}$ – номер текущего состояния кодера для класса C_x . При данном способе формирования множества состояний кодера его мощность будет определяться

$$\alpha = \sum_{c_x=1}^{C_x} \alpha^{(C_x)}. \quad (32)$$

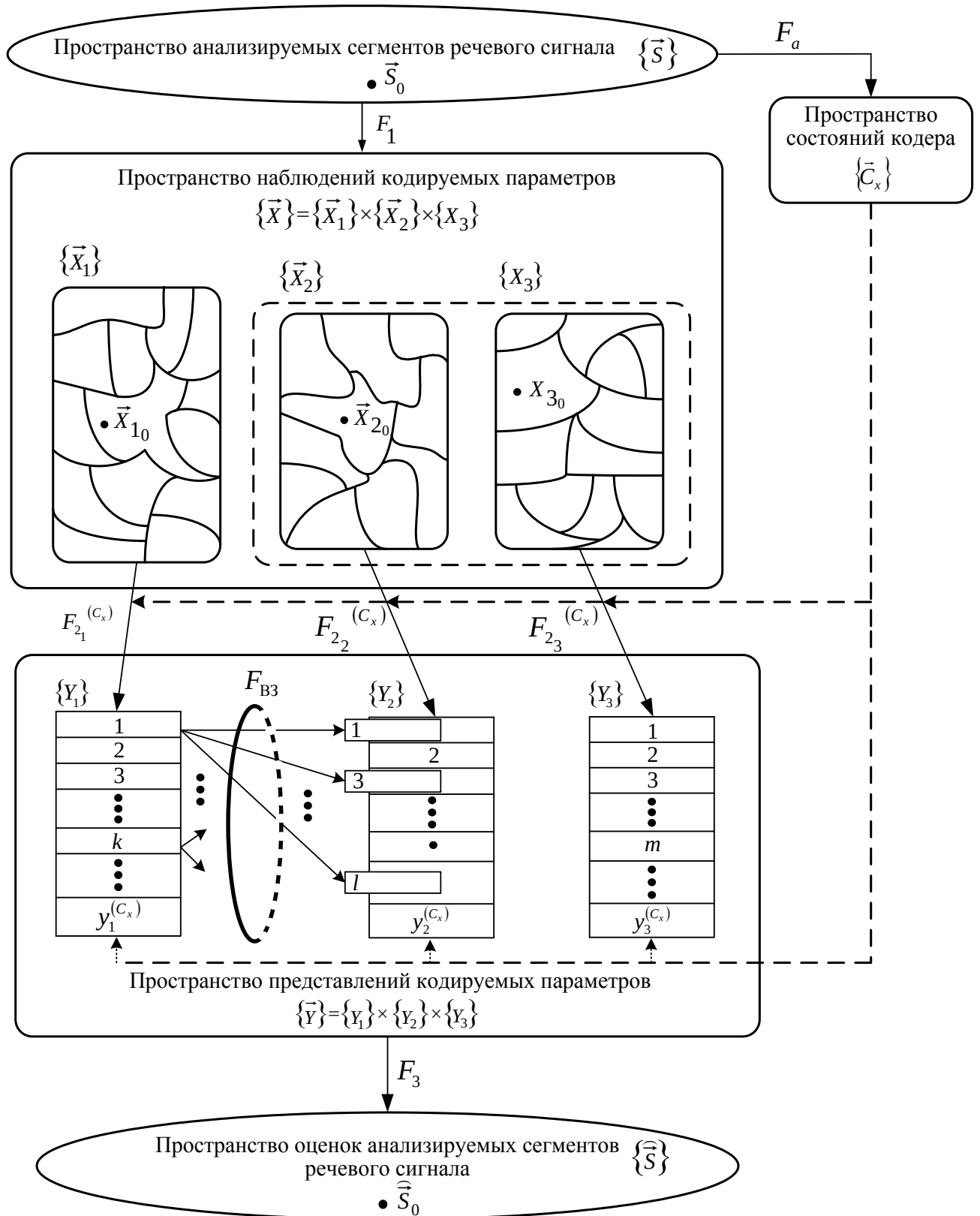


Рисунок 2 – Математическая модель процедур преобразования РС в системе гибридного кодирования с учетом зависимостей элементов декомпозиции

Для C_x -го класса сегментов средняя величина D определяется как

$$D^{(C_x)}(\vec{S}_0, \hat{\vec{S}}_0) = \frac{1}{\alpha^{(C_x)}} \sum_{\theta=1}^{\alpha^{(C_x)}} D(\vec{S}_0, \hat{\vec{S}}_0). \quad (33)$$

Среднее искажение кодируемых кадров речи при обработке будет определяется выражением

$$D(\bar{S}_0, \hat{\bar{S}}_0) = M \left[D^{(C_x)}(\bar{S}_0, \hat{\bar{S}}_0) \right] = \sum_{C_x=1}^{C_x} P^{(C_x)} D^{(C_x)}(\bar{S}_0, \hat{\bar{S}}_0) = \sum_{C_x=1}^{C_x} P^{(C_x)} \frac{1}{\alpha^{(C_x)}} \times$$

$$\times \sum_{\theta=1}^{\alpha^{(C_x)}} \frac{1}{\gamma} \sum_{Y_1=1}^{Y_1} \sum_{Y_2=1}^{Y_2} \sum_{Y_3=1}^{Y_3} D\{\bar{X}_0, F_3[F_{2_1}(\bar{X}_{1_0}), F_{2_2}(\bar{X}_{2_0}), F_{2_3}(\bar{X}_{3_0})]\} \times P_W(\bar{X}_{1_0}, \bar{X}_{2_0} | \bar{X}_{1_0}, \bar{X}_{3_0}) \quad (34)$$

Практически значимой особенностью разрабатываемой системы является возможность построения гибридных кодеров РС учитывающих зависимость элементов декомпозиции РС в процессе его обработки на основе метода ЛП, использующих ВК данных параметров. Параметры кодека, оказывающие доминирующее влияние на качество кодирования РС, играют роль переменных при решении задачи синтеза системы кодирования и определяют значение целевой функции вида:

$$D(\bar{S}_0, \hat{\bar{S}}_0) \xrightarrow{\bar{\varepsilon}} \min, \bar{\varepsilon} = \{C_x, z^{(C_x)}_i, \alpha^{(C_x)}, P, y_i^{(C_x)}\}, \quad (35)$$

где C_x – конечное число непересекающихся классов сегментов РС; $z^{(C_x)}_i$ – число векторов возбуждения, зависящих с вектором ЛСЧ; $\alpha^{(C_x)}$ – мощность множества состояний кодера для C_x -го класса сегментов РС; P – количество коэффициентов ЛП; $y_i^{(C_x)}$ – объем i -ой КК для C_x -го класса сегментов РС.

Использование зависимостей элементов декомпозиции РС $\{\bar{a}\}$ и $\{\bar{e}\}$ при создании модели, учитывающей эти связи, требует формирования кластеризованных пространств $\{\bar{a}\}$, $\{\bar{e}\}$ и нахождение зависимостей между ними.

Согласно обобщенной математической модели использование таких зависимостей требует динамического перераспределения мощностей подпространств $\{Y_i\}$ представлений кодируемых параметров с возможным изменением отображений F_{2i} . Такая задача также относится к классу задач эффективного кодирования, а в условиях идеального канала связи – к классу задач сжатия РС. На рисунке 3 представлена модифицированная структурная схема кодера, реализующего предложенный метод. Особенностью схемы является введение процедуры классификации сегментов РС на непересекающееся число классов, поскольку реализация гипотезы о зависимости элементов декомпозиции РС требует перехода к наиболее адекватным условиям предлагаемой оптимизации системам классифицированного ВК, позволяющим снизить существующую избыточность фиксированных КК.

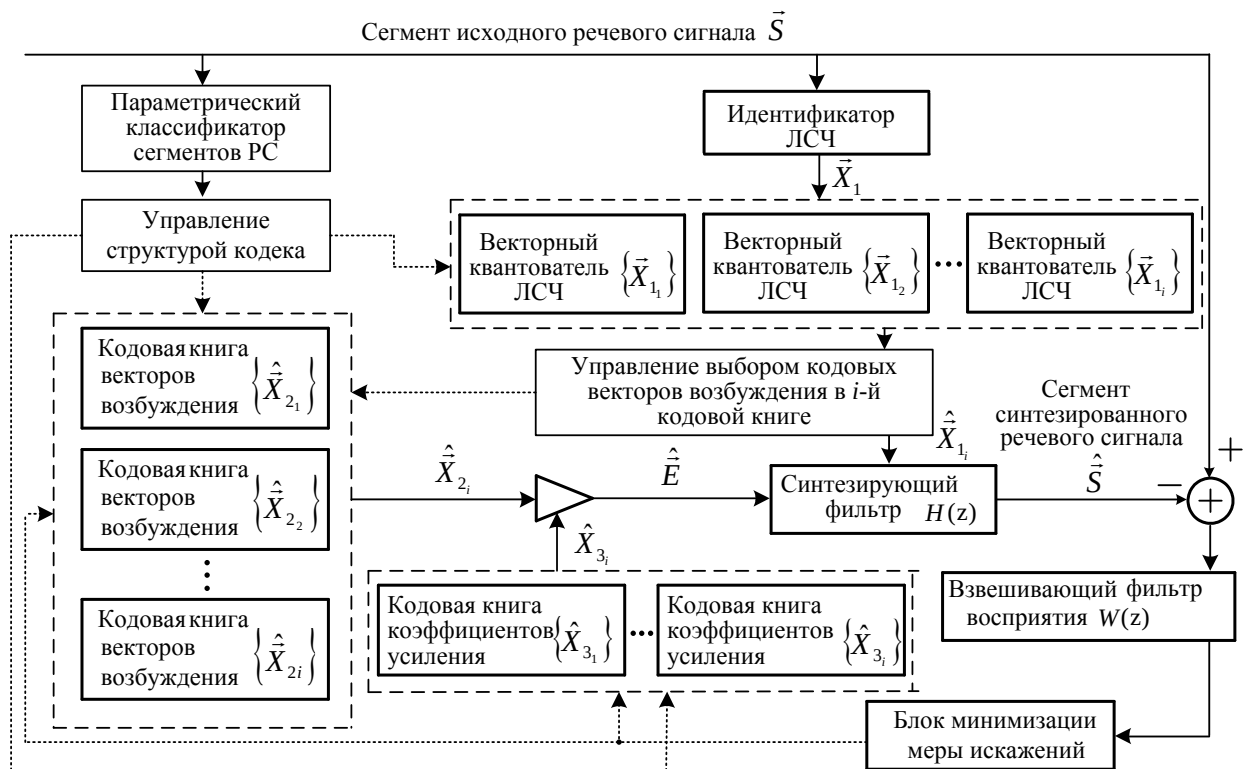


Рисунок 3 – Структурная схема кодера системы обработки речевого сигнала с учетом зависимостей элементов декомпозиции

В работе доказано следующее утверждение: при ограниченном порядке ЛП речи векторные множества элементов декомпозиции РС (вектора параметров синтезирующей модели $\{A\}$ и сигнала возбуждения $\{E\}$) становятся зависимыми и определяются многозначными необратными отображениями между соответствующими векторными множествами. Для его доказательства в диссертации применен метод от противного в рамках положений метода порождающего процесса в предположении, что сигнал остатка ЛП представляет собой дискретный белый гауссовский шум в соответствии с моделью авторегрессионного цифрового спектрального анализа, то есть его реализации не коррелированы между собой и не связаны с векторами параметров формирующей модели.

На основании вышеизложенного сформулировано следующее утверждение: учет зависимости элементов множеств декомпозиции РС, получаемых на основе модели ЛП, между пространствами кластеризации при его моделировании, приводит к уменьшению мощности используемого множества представления одного из векторных пространств при обработке. На основании данного утверждения разработан и в разделе 6 диссертации представлен алгоритм уменьшения избыточности представления

РС в процессе его обработки и низкоскоростного кодирования для передачи по каналам цифровой связи. Основная трудность реализации системы кодирования с учетом зависимостей связана с отсутствием априорной информации о вероятной мощности пространства представлений различных вариантов КК.

Таким образом, использование зависимостей элементов декомпозиции РС при ЛП с ВК его параметров позволяет усовершенствовать системы обработки РС, поскольку большинство подобных устройств осуществляли независимое представление и кодирование данных параметров, что определяло информационную избыточность их кадра передачи. Учет связей элементов декомпозиции позволяет уменьшить число бит, необходимых для представления векторов сигнала возбуждения, что приводит к снижению скорости передачи при сохранении качества синтезированной речи по показателям естественности ее звучания и узнаваемости. Разработанная математическая модель системы кодирования РС показывает пути совершенствования систем обработки РС в условиях соответствия тенденциям развития инфокоммуникационных систем при учете зависимостей элементов декомпозиции РС. В работе сформулировано следующее утверждение: нарушение характера и степени связи элементов множеств декомпозиции РС, определяемой многозначным отображением между соответствующими векторными множествами, указывает на наличие постороннего шума или несанкционированного использования системы кодирования РС. Взаимосвязи множеств $\{\bar{a}\}, \{\bar{e}\}$ в условиях наличия акустических шумовых воздействий начинают разрушаться, что связано с неправильной оценкой и выбором вектора $\bar{a}$ при поиске в КК. При организации доступа к ресурсу сети связи используются различные механизмы подтверждения легитимности абонента, основанные на биометрической аутентификации. Основным недостатком подобных методов является однократная проверка легитимности пользователя с дальнейшим предоставлением ему прав использования ресурсов сети на весь сеанс связи. Для его устранения разработаны предложения по непрерывной аутентификации диктора, причем процедура аутентификации оказывается связанной непосредственно с процессом ведения телефонных переговоров по сети связи [9].

В четвертом разделе развита теория анализа и принципы обработки РС при шумоподавлении, разработана математическая модель системы фильтрации РС на основе применения полиспектрального анализа [8,12], представлены новые технические решения по выделению сегментов паузы и оценки характеристик акустических

воздействий в условиях зашумления обрабатываемого РС [26, 27, 46], разработан метод шумоподавления РС на основе применения полиспектрального анализа [26].

Разработка методов очистки РС от шума является весьма актуальным направлением исследований [12]. В тех случаях, когда шум имеет значительную интенсивность, его наличие может существенно исказить результаты обработки, анализа или распознавания РС. В целом ряде других случаев, например, при обработке сильно зашумленного РС или его восстановления на фоне шумов и помех, задача очистки РС от шума или его восстановления носит самостоятельный характер и является частной целью работы. Анализ условий обработки зашумленного РС показывает, что определение значений линейных спектральных частот выполняется с изначальной ошибкой, что в общем случае приводит к нестабильности процедуры их оценки и коррекции. Основным предположением, которое используется в алгоритмах шумоподавления (с одним входом) на основе спектрального вычитания, является то, что усредненная оценка спектра мощности паузы РС $\langle A_E^2(k, d) \rangle$ равна оценке аддитивного воздействия помехи на сам сигнал $A_{SS}^2(k, d)$ на зашумленном участке:

$$A_U^2(k, d) \approx A_{SS}^2(k, d) + \langle A_E^2(k, d) \rangle. \quad (36)$$

В связи с этим сильно возрастает зависимость системы фильтрации на основе спектрального вычитания (с одним входом) от самого факта правильного обнаружения паузы. В работе предложен и обоснован метод шумоподавления в РС на основе применения полиспектрального анализа. В работе выделены методы шумоподавления на основе спектрального вычитания из-за их относительной простоты и возможности практической реализации на основе современных технологий. С учетом проведенного анализа в рамках диссертации выявлены недостатки этих методов. Для их устранения предложена частная математическая модель системы фильтрации РС на основе полиспектрального анализа (рисунок 4). Данную модель можно описать в виде (37):

$$U(n, d) \xrightarrow{\Pi\{F_1, F_2, F_3\}} S^*(n, d), \quad (37)$$

где $U(n, d)$ – сегмент РС на входе системы фильтрации; Π – оператор системы фильтрации РС; F_1 – функционал выделения сегментов паузы; F_2 – функционал оценки спектральных характеристик АВ; F_3 – функционал шумоподавления на основе прямого вычитания в амплитудных спектрах Фурье; $S^*(n, d) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} A_{SS}^2(k, d) \cdot e^{kn}$ – РС на выходе

системы фильтрации при $n=1, 2, \dots, N$, $k=1, 2, \dots, N$, $d=1, 2, \dots, D$; n – номер отсчета; k – номер спектральной компоненты; d – последовательный номер анализируемого сегмента; N – количество отсчетов на сегменте РС; D – количество сегментов обработки РС; $\dot{A}_{SS}(k, d) = A_{SS}(k, d) \cdot \exp[F_U(k, d)]$ – комплексный спектр Фурье сегмента РС на выходе системы фильтрации; $F_U(k, d)$ – оценка спектра фаз сегмента РС на входе системы фильтрации; $A_{SS}(k, d) = F_3\{A_U(k, d), \langle A_E(k, d) \rangle, I(d)\}$ – оценка спектра амплитуд сегмента РС на выходе системы фильтрации; $\langle A_E(k, d) \rangle = F_2\{A_U^2(k, d), A_U(k, r), J(d)\}$ – усредненная оценка спектра амплитуд АВ для сегмента d , при $r=0, 1, \dots, M$, $A_U^2(k, d)$ – совокупность оценок спектра мощности для D – выборки обрабатываемых сегментов РС, M из которых сегменты паузы и L – сегментов, содержащих полезную составляющую; $A_U(k, r)$ – совокупность оценок спектра амплитуд M – выборки сегментов паузы обрабатываемого РС; $J(d) = F_1\{A_U^2(k, d), B_U^3(p, q, d)\}$, $d \in M \Rightarrow J(d) = 1$, $d \in L \Rightarrow J(d) = 0$ – бинарная функция признака паузы; $B_U^3(p, q, d) = A_U(p, d) \cdot A_U(q, d) \cdot A_U(p+q, d)$ – совокупность оценок биамплитуды для D – выборки обрабатываемых сегментов, синтезируемых согласно прямому методу; $J(d) = 1 \Rightarrow I(d) = 1$, $J(d) = 0 \Rightarrow I(d) = 1/2$ – коэффициент усреднения.

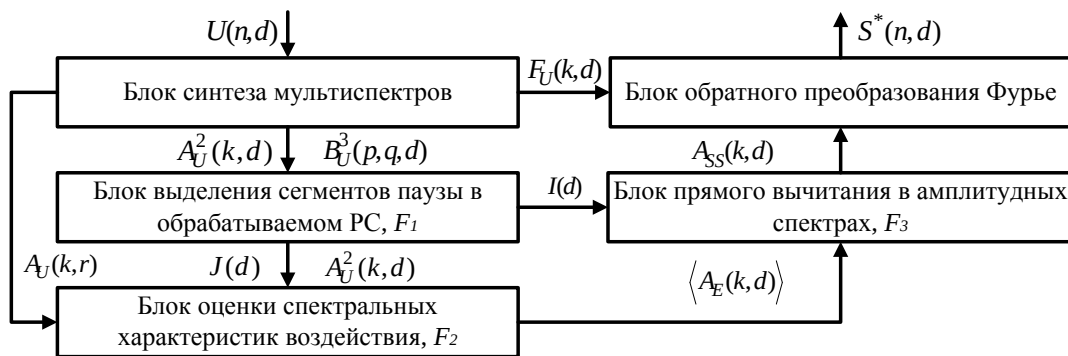


Рисунок 4 – Функциональная схема модели системы фильтрации РС с полиспектральным анализом

Случайный характер РС в УИАВ даже при достаточно точной усредненной оценке спектральных характеристик АВ не позволяет провести шумоподавление на основе спектрального вычитания без возникновения искажений. Математическую интерпретацию линейной разности ОСШ ΔG выход-вход относительно РС в отсутствии зашумления $S(n, d)$ можно представить следующим выражением:

$$\Delta G = G^{out} - G^{in}, \quad (38)$$

где $G^{out} = \frac{1}{D} \cdot \sum_{d=1}^{d=D} 10 \cdot \lg \left(\sum_{n=0}^{n=N-1} \frac{[S(n, d)]^2}{[S^*(n, d) - S(n, d)]^2} \right)$ – ОСШ на выходе системы фильтрации РС;

$G^{in} = \frac{1}{D} \cdot \sum_{d=1}^{d=D} 10 \cdot \lg \left(\sum_{n=0}^{n=N-1} \frac{[S(n, d)]^2}{[U(n, d) - S(n, d)]^2} \right)$ – ОСШ на входе системы фильтрации РС.

Анализ (39) показывает максимум, который теоретически достижим при $S^*(n, d) - S(n, d) = 0 \quad \forall d$. Таким образом, выполнение функционала F_3 зависит от F_2 , который, в свою очередь, зависит от F_1 , вследствие чего главной задачей идеализированной системы фильтрации является сведение усредненной оценки $\langle A_E(k, d) \rangle$ к оценке амплитудного спектра акустического воздействия $A_E(k, d) \quad \forall k, d$:

$$A_E(k, d) - \langle A_E(k, d) \rangle = 0. \quad (39)$$

Выполнение равенства (39) определяет: 1) получение точной оценки спектральных характеристик АВ для каждого обрабатываемого сегмента РС ведет к достижению максимума в рамках предложенной модели системы фильтрации; 2) вырождение процедуры усреднения сегментов паузы с целью формирования остаточного шума, что ведет к возникновению воспринимаемых слуховым аппаратом человека незначительных искажений (для случая УИАВ), связанных с использованием в системе фильтрации исходных фазовых соотношений. При практической реализации выполнение (37) в условиях реальных АВ высокой интенсивности труднодостижимо. Поэтому на выходе системы фильтрации присутствует шум $E^*(n, d) = S^*(n, d) - S(n, d)$, который представлен (40):

$$E^*(n, d) = E_{rez}^*(n, d) + E_{mat}^*(n, d), \quad d \in L; \frac{1}{2} \cdot (E_{rez}^*(n, d) + E_{mat}^*(n, d)), \quad d \in M, \quad (40)$$

где $E_{rez}^*(n, d)$ – остаточный шум на выходе системы фильтрации; $E_{mat}^*(n, d)$ – собственные шумы системы (шумы округлений, квантования и др.).

Проверка модели на чувствительность и адекватность позволяет пренебречь шумами вычислений на выходе системы фильтрации и уточнить (37):

$$\begin{aligned} U(n, d) &\xrightarrow{\Pi\{F_1, F_2, F_3\}} S^*(n, d), \\ A_E(k, d) - \langle A_E(k, d) \rangle &\rightarrow \min, \Leftrightarrow E_{rez}^*(n, d) \rightarrow \min \Leftrightarrow \Delta G \rightarrow \max, \\ \frac{Q^{out} \geq Q^o \quad \forall Q^{in} \leq Q^o}{Q^{out} \geq Q^{in} \quad \forall Q^{in} > Q^o}, &\tau \leq \tau^o, \end{aligned} \quad (41)$$

где Q^{out}, Q^{in}, Q^o – субъективная оценка качества РС на выходе, входе и требуемое значение, τ, τ^o – алгоритмическая задержка и ее требуемое значение.

В пятом разделе работы раскрыт предложенный новый метод адаптивной цифровой фильтрации РС в условиях акустического зашумления [8, 26, 46, 50], проведено обоснование оценки критерия качества РС на выходе системы обработки, показана оценка повышения качества РС при шумоподавлении согласно предлагаемым техническим решениям.

На основании разработанной математической модели, которая отличается дополнительным введением оценки биспектральной плотности, разработан метод адаптивной цифровой фильтрации РС с целью шумоподавления. Он *отличается* от известных тем, что учитывает бикорреляционные свойства сигнала на основе анализа полиспектров с применением полученных параметров в процедурах выделения сегментов паузы и оценки спектральных характеристик акустического воздействия. Разработанный метод не требует априорных сведений об ОСШ и характере акустического воздействия. В диссертации показана возможность коррекции и правильной оценки положения ЛСЧ за счет реализации технических решений по шумоподавлению. На вход системы фильтрации поступает последовательность отсчетов РС, который можно описать тремя возможными состояниями: отсутствие зашумления; наличие только акустического воздействия; наличие аддитивной смеси РС и акустического воздействия. Блок-схема алгоритма адаптивной цифровой фильтрации РС в УИАВ, реализующего данный метод, представлена на рисунке 5, в нем осуществляется:

1) Ввод данных $U(n, d)$, d – последовательный номер сегмента, n – номер отсчета в сегменте, сброс счетчика количества итераций i ;

2) Условие проверки $d \leq D_m$ наличия необходимого числа сегментов для проведения процедуры оценки спектральных характеристик акустического воздействия;

3) Выполнение условия – отсутствие усредненной оценки спектральных характеристик акустического воздействия, порогового значения по каждому коэффициенту суммарной бикорреляции (КСБ) $\lambda_j^{tr}(d)$, все сегменты определяются как активные участки РС, D_m – необходимое количество сегментов;

4) При невыполнении условия $d \leq D_m$ – определение характера АВ:

– оценка спектра мощности анализируемых сегментов $A_U^2(k, d)$ на интервале:

$$\bullet \quad A_U(k, d) = \sum_{n=0}^{N-1} U(n, d) \cdot e^{-kn}, \quad (42) \quad A_U^2(k, d) = \frac{1}{2} \cdot \left(A_U(k, d) \right)^2; \quad (43)$$

– оценка коэффициентов относительной мощности $R_j(d)$:

$$R_j(d) = \frac{1}{F_j} \cdot \sum_{\Omega F_j} A_U^2(k, d) \quad (44); \quad F_j \xleftrightarrow{\alpha} \frac{F_d}{2} \quad (45); \quad \Delta R_j \leq r^o \Rightarrow H(d) = 1, \quad (46)$$

где F_j – область дискретных частот, F_d – нормированная частота дискретизации; $H(d)$ – бинарная функция признака широкополосного АВ;

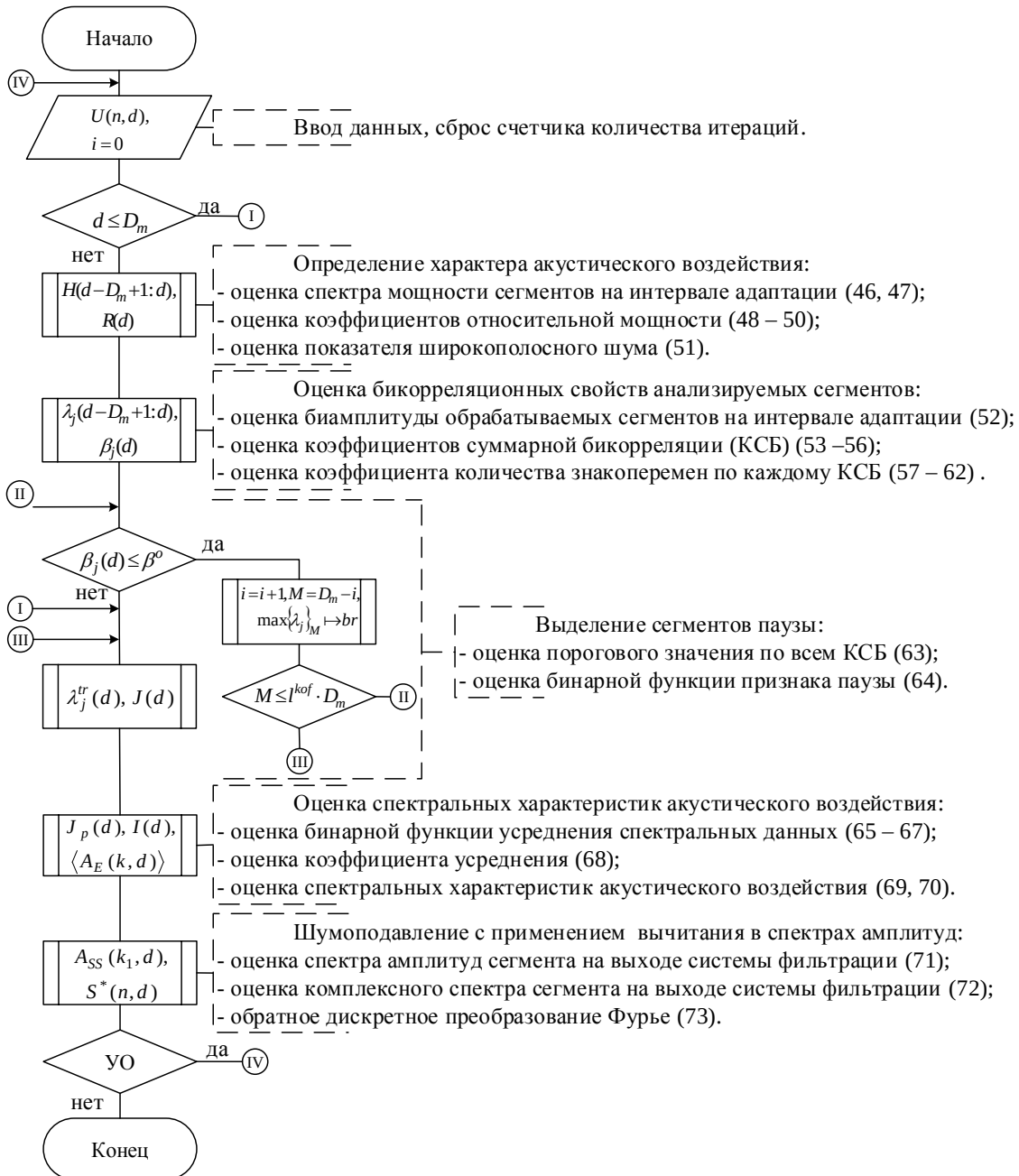


Рисунок 5 – Блок-схема алгоритма адаптивной цифровой фильтрации РС

- оценка показателя широкополосного шума $R(d)$ при $m = d - D_m + 1, d - D_m + 2, \dots, d$:

$$R(d) = \sum_{m=d-D_m+1}^{m=d} H(m). \quad (47)$$

5) Оценка бикорреляционных свойств анализируемых сегментов:

- оценка биамплитуды обрабатываемых сегментов $B_U^3(p, q, d)$ на интервале адаптации:

$$B_U^3(p, q, d) = A_U(p, d) \cdot A_U(q, d) \cdot A_U(p + q, d); \quad (48)$$

- оценка КСБ $\lambda_j(d)$:

$$\lambda_1(d) = \|\Delta F_1, \Delta F_1\| = \sum_{p=f_0}^{p=f_1} \sum_{q=f_0}^{q=f_1} B_U^3(p, q, d), \quad (49) \quad \lambda_2(d) = \|\Delta F_1, \Delta F_2\| = \sum_{p=f_0}^{p=f_1} \sum_{q=f_1}^{q=f_2} B_U^3(p, q, d), \quad (50)$$

$$\lambda_3(d) = \|\Delta F_1, \Delta F_3\| = \sum_{p=f_0}^{p=f_1} \sum_{q=f_2}^{q=f_3} B_U^3(p, q, d), \quad (51) \quad \lambda_4(d) = \|\Delta F_1, \Delta F_4\| = \sum_{p=f_0}^{p=f_1} \sum_{q=f_3}^{q=f_4} B_U^3(p, q, d), \quad (52)$$

где $\lambda_1(d) - \lambda_4(d)$ – коэффициенты суммарной бикорреляции для сегмента d , $\|\Delta F_1, \Delta F_4\|$ – взаимосвязь формантных областей сегмента РС, $f_0 - f_4$ – границы частотных областей в биспектральной плотности. – оценка коэффициента количества знакоперемен по каждому КСБ $\beta_j(d)$:

$$\lambda_j(m) < \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) \geq \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 1, \lambda_j(m) \geq \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) < \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 1, \quad (53)$$

$$\lambda_j(m) < \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) < \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 0, \lambda_j(m) > \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) > \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 0,$$

$$\lambda_j^a(d) \Big|_{M=D_m-i} = \frac{1}{M} \cdot \sum_{\Omega M} \lambda_j(m), \quad (54) \quad \beta_j(d) \Big|_{M=D_m} = \frac{1}{M} \cdot \sum_{\Omega M} \mu_j^\pm(m), \quad (55)$$

где $\lambda_j^a(d)$ – среднее значение j -го КСБ для сегмента d , M – количество анализируемых сегментов, $\mu_j^\pm(m)$ – сегментный признак смены полярности по j -му КСБ относительно $\lambda_j^a(d)$.

- оценка коэффициента количества знакоперемен по каждому КСБ $\beta_j(d)$:

$$\lambda_j(m) < \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) \geq \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 1, \lambda_j(m) \geq \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) < \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 1, \quad (56)$$

$$\lambda_j(m) < \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) < \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 0, \lambda_j(m) > \lambda_j^a(d) \Big|_M \ \& \ \lambda_j(m+1) > \lambda_j^a(d) \Big|_M \Rightarrow \mu_j^\pm(m) = 0,$$

$$\lambda_j^a(d) \Big|_{M=D_m-i} = \frac{1}{M} \cdot \sum_{\Omega M} \lambda_j(m), \quad (57) \quad \beta_j(d) \Big|_{M=D_m} = \frac{1}{M} \cdot \sum_{\Omega M} \mu_j^\pm(m), \quad (58)$$

где $\lambda_j^a(d)$ – среднее значение j -го КСБ для сегмента d , M – количество анализируемых сегментов, $\mu_j^\pm(m)$ – сегментный признак смены полярности по j -му КСБ относительно $\lambda_j^a(d)$.

6) Выделение сегментов паузы (использование итеративного алгоритма "с выкалыванием" максимального значения для получения массива сегментов, для которых $\beta_j(d)$ превышает пороговое значение):

- оценка порогового значения по всем КСБ: $\lambda_j^{tr}(d) = \max\{\lambda_j\}_M$

- оценка бинарной функции признака паузы $J(m)$ при $m = d - D_m + 1, d - D_m + 2, \dots, d$:

$$\lambda_j(m) \leq \lambda_j^{tr}(d) \Rightarrow J(m) = 1, \lambda_j(m) > \lambda_j^{tr}(d) \Rightarrow J(m) = 0. \quad (59)$$

7) Оценка спектральных характеристик акустического воздействия:

- оценка бинарной функции усреднения спектральных данных $J_p(m)$ (использована для повышения точности усредненной оценки спектральных характеристик на основе анализа эффективной полосы частот АВ):

$$\Delta P_a|_M = \frac{1}{M} \cdot \sum_{m=1}^{m=M} \Delta P(m), \quad (60)$$

$$\Delta P^{tr} = \max[\Delta P]_M, \quad (61)$$

$$J(m)=0 \Rightarrow J_p(m)=0, J(m)=1 \ \& \ \Delta P(m) \leq \Delta P^{tr} \Rightarrow J_p(m)=1, J(m)=1 \ \& \ \Delta P(m) > \Delta P^{tr} \Rightarrow J_p(m)=0, \quad (62)$$

где $\Delta P_a|_M$ – среднее значение эффективной полосы частот для M – выделенных сегментов паузы, ΔP^{tr} – пороговое значение эффективной полосы частот АВ;

- оценка коэффициента усреднения $I(d)$:

$$J(d)=0 \Rightarrow I(d)=\frac{1}{2}, J(d)=1 \Rightarrow I(d)=1; \quad (63)$$

- оценка спектральных характеристик АВ $\langle A_E(k, d) \rangle$ при $m=d-D_m+1, d-D_m+2, \dots, d$:

$$A_U^a(k, m) = A_U(k, m) \cdot J_p(m), \quad (64) \quad \langle A_E(k, d) \rangle = \frac{1}{\sum_{m=d-D_m+1}^{m=d} J_p(m)} \cdot \sum_{m=d-D_m+1}^{m=d} A_U^a(k, m), \quad (65)$$

где $A_U^a(k, m)$ – массив спектральных данных.

8) Шумоподавление с применением вычитания в спектрах амплитуд:

- оценка спектра амплитуд сегмента на выходе системы фильтрации $A_{SS}(k, d)$:

$$A_{SS}(k, d) = I(d) \cdot |A_U(k, d) - \langle A_E(k, d) \rangle|; \quad (66)$$

- оценка комплексного спектра сегмента на выходе системы фильтрации $\dot{A}_{SS}(k, d)$:

$$\dot{A}_{SS}(k, d) = A_{SS}(k, d) \cdot \exp[F_U(k, d)], \quad (67)$$

где $F_U(k, d)$ – оценка спектра фаз сигнала на входе системы фильтрации;

- получение отсчетов сегмента РС на выходе системы фильтрации $S^*(n, d)$ путем обратного дискретного преобразования Фурье:

$$S^*(n, d) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} \dot{A}_{SS}(k, d) \cdot e^{kn}. \quad (68)$$

В диссертационной работе выполнена оценка результативности алгоритма адаптивной цифровой фильтрации (рисунок 6), показано, что использование предлагаемых технических решений по шумоподавлению позволяет сохранить положение ЛСЧ, сохраняя при этом выявленные взаимосвязи элементов декомпозиции РС в процессе ЛП при их кластерном представлении.

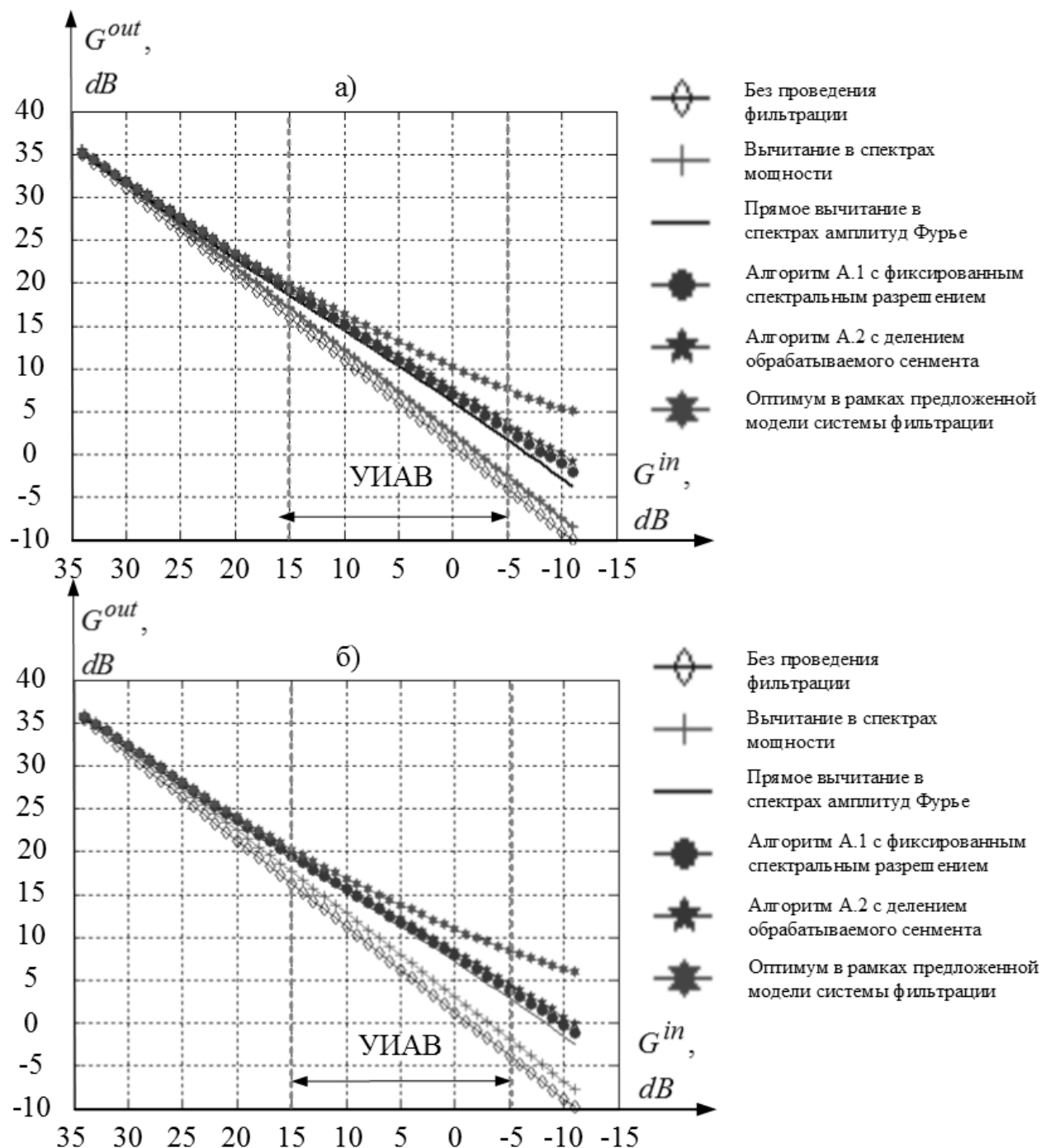


Рисунок 6 – Результативность алгоритмов фильтрации речевого сигнала:

а) зависимость ОСШ на выходе системы фильтрации от ОСШ на входе при выполнении различных алгоритмов для случая зашумления широкополосным акустическим воздействием (на примере шума автомагистрали);

б) зависимость ОСШ на выходе системы фильтрации от ОСШ на входе при выполнении различных алгоритмов для случая узкополосного воздействия (на примере шума двигателя грузового автомобиля).

Также в работе показаны составляющие данных методов обработки РС и представлены предложения по технической реализации предлагаемых решений в современных системах инфокоммуникаций. Отражено использование и комбинация ап-

паратных и программных средств реализации технических решений обработки РС в классе алгоритмов цифровой обработки сигналов.

Полученные результаты указывают на повышение точности выполнения оценки значений параметров элементов декомпозиции РС в процессе его обработки на основе метода ЛП с ВК его параметров при применении данных технических решений по шумоподавлению в РС. Правильная идентификация векторов КК параметров передаточной функции голосового тракта и соответствующего сигнала возбуждения дополнительно приведет к подавлению остаточного акустического шумового воздействия, так как значения данных векторов обучены на фрагментах активного РС при отсутствии побочных шумовых составляющих различной природы. Положительный эффект от предобработки РС с целью шумоподавления проиллюстрирован на (рисунке 7), его анализ показывает, что после проведения шумоподавления в очищенном сигнале происходит восстановление состава локальных мод огибающих спектра амплитуд (ЛМОСА), что приведет на сегменте анализа к точному определению линейных спектральных частот РС.

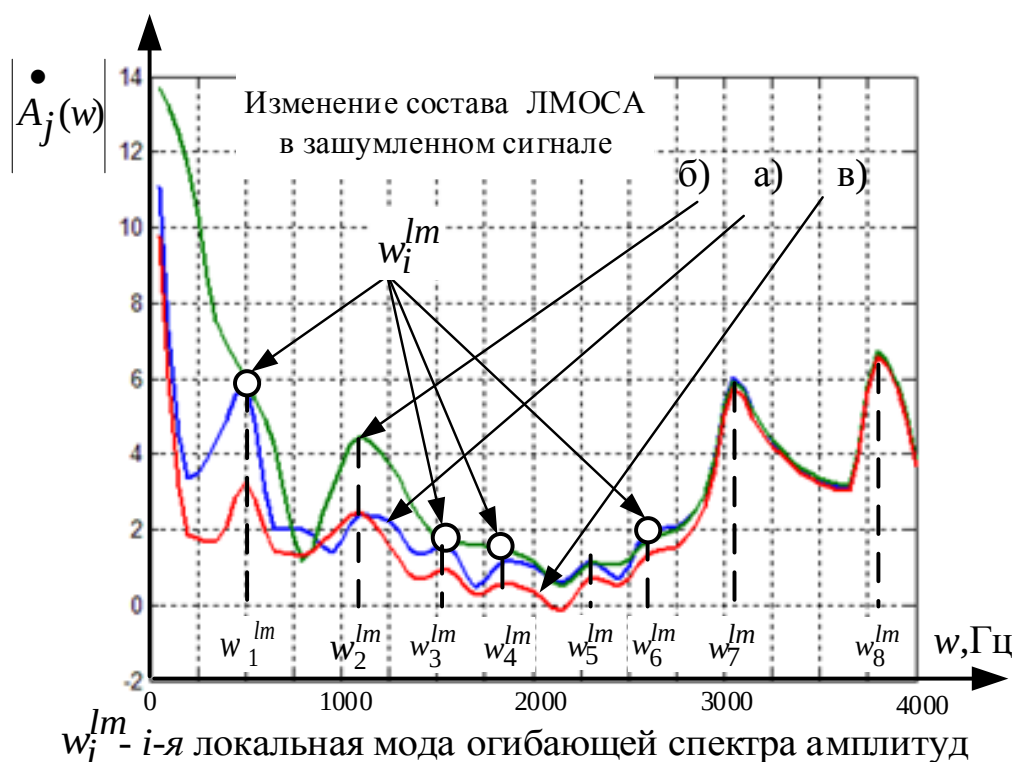


Рисунок 7 – Восстановление исходного состава ЛМОСА в очищенном РС:
а) огибающая спектра амплитуд чистого РС; б) огибающая спектра амплитуд зашумленного РС; в) огибающая спектра амплитуд очищенного РС.

В шестом разделе работы рассмотрены подходы к определению характеристик РС на сегменте анализа и способы классификации сегментов [3, 27, 37, 50, 52], представлены технические решения по выбору длительности однородного сегмента анализа при его обработке на основе ЛП [1, 14, 18, 22, 29, 40, 41, 42, 44, 51, 53] и нахождения зависимостей параметров декомпозиции РС при ЛП [7, 11, 24, 28, 39, 43, 45, 47, 48, 49, 54], выявлены особенности реализации и применения систем обработки РС в средствах и комплексах инфокоммуникаций [20, 25, 35, 36, 38].

Показано, что максимально возможный сегмент одновременно анализируемых данных составляет 80 мс, что связано с требованиями по задержке РС при передаче, определяемыми рекомендацией G.114 Международного союза электросвязи. Если на протяжении 80 мс не произошло перехода между областями, то увеличение длительности сегмента прекращается. Экспериментально установлено, что принадлежность значения логарифма отношения "эффективной" области, соответствует переходному процессу, и в случае нахождения анализируемой выборки в «эффективной» области по достижении 80 мс процесс заканчивается принятием основной гипотезы. Показателем принадлежности сегмента анализа, получаемого на каждом этапе эксперимента, к определённой области является логарифм отношения функций правдоподобия $Z[X]$, вычисляемый по соотношению (11). Использование такого подхода к формированию сегментов обработки РС позволяет наиболее полно использовать статистическую взаимосвязь между случайными значениями РС для выделения однородных сегментов анализа, что дает возможность не осуществлять на длительном сегменте анализа РС перестройки параметров формирующей (передаточной) функции голосового тракта. Каждая ЛСЧ определяет положение соответствующей форманты, что дает возможность осуществлять их межсегментную интерполяцию на однородных сегментах РС с «гладкими» изменениями спектральной плотности мощности. В низкоскоростных системах обработки РС, повышения качества синтезируемой речи можно добиться, используя в качестве параметров для передачи информацию о состоянии КЛП ЛСК. В таблице 1 представлены значения ЛСЧ в радианах на однородных сегментах РС. Алгоритм, реализующий выделение однородных сегментов РС представлен в шестом разделе диссертации. Проведенные экспериментальные исследования выявили возможности применения предлагаемого способа, которые показал уменьшение объема анализируемых данных РС при низкоскоростном кодировании в среднем на 20% при описании формирующей (передаточной

функции) голосового тракта на однородных сегментах при незначительном ухудшении качественных показателей синтезированного РС. В низкоскоростных методах кодирования РС около 25% информационного ресурса кадра передачи используется для передачи данных о параметрах формирующей функции, около 50% в РС занимают паузы, а 70% активной речи занимают вокализованные участки.

Таким образом, выделение однородных участков РС на длительности нескольких классических сегментов анализа дает возможность до 20% уменьшать среднюю скорость передачи данных в КС, за счет сохранения информации о параметрах формирующей модели. Данные результаты являются следствием внедрения динамического изменения границ сегмента анализа и расчета значений параметров, описывающих передаточную

Таблица 1 – Значения ЛСЧ (в радианах) на однородных сегментах речевого сигнала

N	1:196	1:246	1:296	1:346
1.	0,252	0,252	0,251	0,251
2.	0,37	0,371	0,371	0,372
3.	0,474	0,474	0,474	0,476
4.	0,618	0,62	0,619	0,618
5.	1	0,998	0,989	0,983
6.	1,664	1,623	1,629	1,607
7.	1,98	1,978	1,977	1,974
8.	2,108	2,108	2,102	2,1
9.	2,544	2,538	2,544	2,537
10.	2,82	2,827	2,822	2,823

функцию голосового тракта на этих сегментах. В диссертации показаны дальнейшие пути развития систем обработки РС. При реализации предлагаемых решений выявлено, что использование зависимостей параметров голосового тракта и сигналов возбуждения при разработке низкоскоростных систем кодирования речи позволяет сократить среднюю скорость передачи при сохранении разборчивости синтезированной речи. Применение такого подхода возможно в системах обработки речевых данных на основе процедуры ЛП при кластерном представлении параметров декомпозиции РС. Сигнал возбуждения формирующего фильтра в идеальном случае рассматривается как результат аппроксимации ошибки ЛП, где в качестве критерия используется минимум дисперсии ошибки оценивания. Последовательность значений ошибки ЛП, используемая для возбуждения, взаимосвязана с оценкой линейного предсказания до тех пор, пока не будет представлять собой белый шумовой процесс. Определение параметров возбуждающего сигнала в системе анализа через синтез с ЛП основано на исследовании сигнала ошибки $e(n)$, при этом сигнал возбуждения формирующего фильтра голосового тракта является аппроксимацией сигнала ошибки. На этапе функционирования применяются заранее обученные КК ВК с конечным числом классов разделений пространств параметров декомпозиции сегментов

РС. Обучение КК ВК параметров формирующей модели РС и сигналов возбуждения позволяет определить зависимость между данными параметрами декомпозиции. Пример установления взаимосвязи для элементов декомпозиции РС при их кластерном представлении на основе ЛП и их дальнейшем ВК представлен в таблице 2.

Таблица 2 – Соответствие векторов элементов декомпозиции РС

Номера векторов кодовой книги параметров голосового тракта	Номера векторов кодовой книги сигналов возбуждения										
1	18	25	34	38	54	69	125	203	250	—	...
2	36	42	50	54	56	67	92	—	—	—	...
3	71	76	80	84	89	93	144	189	226	245	...
4	5	18	104	200	212	—	—	—	—	—	...
5	9	32	40	—	—	—	—	—	—	—	...
...	...	...	...	...	...	...	...	...	...	...	...

При установлении взаимосвязей определить точное соответствие вектора параметров голосового тракта и вектора сигнала возбуждения точно для всех векторов не представляется возможным, что связано с особенностями обучения соответствующих КК. Исследования сформированных КК на тестовом материале, показало, что максимальное число векторов возбуждения для одного вектора ЛСЧ может достигать 16-20, для КК состоящей из 256 векторов сигналов возбуждения.. Таким образом, сокращается количество бит для представления вектора в подпространстве КК сигналов возбуждения, что дает возможность до 20% уменьшать среднюю скорость передачи в КС при сохранении, качественных показателей синтезированного РС, а при использовании индивидуально ориентированных кодовых книг существует возможность осуществлять постоянный контроль легитимности доступа абонента к сети связи. Учет зависимости элементов декомпозиции РС при ЛП на основе CELP-подобных алгоритмов в процессе его обработки и низкоскоростного кодирования в условиях наличия акустических шумов позволяет по-новому рассмотреть задачу создания индивидуально ориентированных систем абонентского речевого кодирования.

В заключении сформулированы основные результаты и выводы по работе.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В диссертационной работе сформулирована научно-техническая проблема, решение которой дало возможность развить теорию анализа РС, на основании которой предложить модели и методы совершенствования систем обработки РС, учитывающие особенности природы его возникновения и принятые технологии обработки в системах инфокоммуникаций.

Представленные технические решения подтверждают выдвинутую гипотезу исследования о том, что использование положений теории линейной фильтрации и результатов полиспектрального анализа РС в условиях наличия акустических шумовых воздействий, а также учет зависимости элементов его декомпозиции и особенностей речеобразования при его моделировании в системах связи, дают возможность улучшить характеристики функционирования систем его обработки.

Основные результаты выполненной работы.

1. Показана актуальность модернизации существующих систем обработки РС, выявлены объективные предпосылки снижения средней скорости передачи в КС без ухудшения качества синтезированной речи.

2. Представлено комплексное техническое решение по формированию методологии понижения скорости кодирования РС при его передаче с сохранением качественных показателей синтезированной речи в условиях наличия акустических шумовых воздействий;

3. Развита теоретическая концепция обработки РС на основе его моделирования в виде свертки параметров, описывающих передаточную функцию ограниченного порядка голосового тракта и сигнала возбуждения, которые отличаются от известных методов учетом зависимости элементов декомпозиции РС, динамическим изменением длительности сегмента анализа при фиксированных значениях параметров синтезирующего фильтра голосового тракта.

4. Разработаны:

- модель кодирования РС со структурно-параметрической адаптацией, основанная на его представлении в виде свертки параметров, описывающих передаточную функцию ограниченного порядка голосового тракта и сигнала возбуждения, отличающаяся от известных учетом особенностей формирования РС и технологий его

кодирования, и позволяющая разработать методологию снижения средней скорости передачи данных в КС.

- модель обработки РС, основанная на учете зависимости элементов декомпозиции РС при его низкоскоростном кодировании на основе метода линейного предсказания, позволяющая осуществить снижение мощности пространства представления сигналов возбуждения.

- модель системы фильтрации РС, отличающаяся применением полиспектрального анализа и позволяющая разработать методы шумоподавления в РС в широком диапазоне значений ОСШ.

5. Предложены:

- метод выделения сегментов анализа РС различной длительности на однородных участках РС, позволяющий осуществлять его синтез при фиксированных значениях параметров передаточной функции голосового тракта;

- метод обработки РС, основанный на учете зависимости элементов декомпозиции РС при ЛП, позволяющий уменьшать среднюю скорость передачи в канале связи при сохранении качественных показателей синтезированной речи.

- метод адаптивной цифровой фильтрации РС в условиях акустического зашумления, отличающийся применением технологий полиспектрального анализа и позволяющий осуществлять качественное шумоподавление при низких отношениях сигнал-шум до 0÷-5дБ.

6. Разработаны и представлены алгоритмы анализа и обработки РС, позволяющие повысить характеристики функционирования телекоммуникационных устройств в системах инфокоммуникаций в условиях наличия акустических шумов, отличающиеся от известных учетом особенностей формирования и обработки РС, применением результатов полиспектрального анализа с целью шумоподавления, а также учетом зависимости элементов декомпозиции РС при линейном предсказании, устраняющие противоречия между особенностями речеобразования и существующими в современных системах связи технологиями обработки.

Востребованность предложенных в диссертации элементов теории, моделей, методов и алгоритмов анализа и обработки РС подтверждается результатами решения исследовательских задач, направленных на практическое применение предлага-

емых технологий обработки РС в интересах инфокоммуникационной среды Российской Федерации. Полученные в диссертационной работе результаты соответствуют паспорту специальности 05.12.13 – Системы, сети и устройства телекоммуникаций.

Значение диссертации для экономического развития и обороноспособности страны определяется тем, что в ее рамках разработаны и предложены методы улучшения качественных показателей функционирования устройств обработки РС, применение которых определяет пути совершенствования инфокоммуникационных систем государства. В диссертационной работе раскрыты перспективные направления развития устройств обработки РС в инфокоммуникационных системах, применение которых даёт возможность перейти на новый этап использования подобных технологий обработки РС. На основании полученных результатов работы можно сделать вывод о достижении поставленной цели диссертационного исследования.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в рецензируемых научных изданиях:

1. Афанасьев, А.А. Выделение сегментов квазистационарности при анализе речевого сигнала в вокодерах с линейным предсказанием/ А.А. Афанасьев, В.Г. Трубицын // Фундаментальные и прикладные проблемы техники и технологии. – 2010. – № 2(280). – С. 97–103.
2. Афанасьев, А.А. Синтез сигналов возбуждения в низкоскоростных вокодерах с линейным предсказанием/А.А. Афанасьев//Телекоммуникации–2010.– №11.–С.8–13.
3. Афанасьев, А.А. Психоакустический классификатор кадров при низкоскоростном кодировании широкополосного речевого сигнала /А.А. Афанасьев, О.О.Басов, М.В. Илюшин // Фундаментальные и прикладные проблемы техники и технологии. – 2010. – № 3(281). – С. 73–78.
4. Афанасьев, А.А. Использование психоакустической модели слуха при разработке вокодеров с линейным предсказанием/ А.А. Афанасьев, М.В. Илюшин // Цифровая обработка сигналов.– 2011.– № 1. – С. 49–52.
5. Афанасьев, А.А. Уменьшение объема данных при широкополосном кодировании речевого сигнала/А.А. Афанасьев, В.Г. Трубицын// Цифровая обработка сигналов. – 2011. – № 1. – С. 53–56.
6. Афанасьев, А.А. Метод снижения скорости передачи в вокодерах с линейным предсказанием на основе применения нейронных сетей при переменной длине сегмента анализа / А.А. Афанасьев, А.П. Рыжков // Информационные системы и технологии. – 2012. – № 6 (74). – С. 20 – 28.

7. Афанасьев, А.А. Использование взаимозависимостей параметров линейного предсказания при реализации процедур обработки речевых данных/ А.А. Афанасьев, А.П. Рыжков //Телекоммуникации.– 2012.– № 13. – С. 32–36.

8. Афанасьев, А.А. Модель системы предобработки речевого сигнала на основе полиспектрального анализа / Афанасьев А.А., Титов О.Н. // Вестник Рязанского государственного радиотехнического университета.– 2014.– Вып.50. – С. 9–16.

9. Афанасьев, А.А. Непрерывная аутентификация диктора при ведении телефонных переговоров по низкоскоростным цифровым каналам / А.А. Афанасьев // Вопросы кибербезопасности.– 2015. – № 3(16).– С. 60–68.

10. Афанасьев, А.А. Перспективы развития систем низкоскоростного кодирования речи при учете взаимозависимостей элементов декомпозиции речевого сигнала/ А.А. Афанасьев // Информационные системы и технологии.–2016.–№2(94)– С.14-21.

11. Афанасьев, А.А. Система обработки речевого сигнала на основе модернизации метода прямого предсказания / Афанасьев, А.А. // Проектирование и технология электронных средств. – 2017.– №3.– С.

12. Афанасьев, А.А. Модель системы предобработки речевого сигнала на основе накопления базы биспектров сигналов речи / А.А. Афанасьев, Д.Д. Сергеев, О.Н. Титов // Информационные системы и технологии. – 2018.– №2. – С.56-67.

Патенты на изобретения и полезные модели:

13. Патент № 2343564 РФ, МПК G10L 19/06. Способ адаптивного кодирования речевых сигналов на основе системы с переменной структурой / А. А. Афанасьев, Г.В. Богачев и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2006143249/09; заявл. 06.12.2006, опубл. 10.01.2009. Бюл. № 1.

14. Патент на полезную модель № 78977 РФ, МПК G10L 11/00. Анализатор основного тона и сигнала тон-шум / А.А. Афанасьев, О.О. Басов и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2008126595/22; заявл. 30.06.2008, опубл. 10.12.2008. Бюл. № 34.

15. Патент №2374653 РФ, МПК G01R23/16.Способ цифрового спектрального анализа сигналов / А.А. Афанасьев, М.Ю. Конышев и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России.–№2008127698/28;заявл.07.07.2008, опубл. 27.11.2009. Бюл.№ 33.

16. Патент № 2400832 РФ, МПК G10L 19/08. Способ формирования сигнала возбуждения в низкоскоростных вокодерах с линейным предсказанием / А.А. Афанасьев, М.Ю. Конышев и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2008146294/09; заявл. 24.11.2008, опубл. 27.05.2010. Бюл. № 15.

17. Патент № 2394284 РФ, МПК G10L 19/08. Способ сжатия и восстановления речевых сигналов для систем кодирования с переменной скоростью передачи / А.А. Афанасьев, С.Р. Габдулгазиев и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2009110842/09; заявл. 24.03.2009, опубл. 10.07.2010. Бюл. № 19.

18. Патент № 2400831 РФ, МПК G10L 19/00. Способ выделения сегментов квазистационарности при анализе речевого сигнала в вокодерах с линейным предсказанием / А.А. Афанасьев, С.Р. Габдулгазиев и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2009121174/09; заявл. 03.06.2009, опубл. 27.09.2010. Бюл. № 27.

19. Патент № 2432624 РФ, МПК G10L 19/00. Способ уменьшения объема данных при широкополосном кодировании речевого сигнала / А.А. Афанасьев, В.Г. Трубицын и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2010115915/09, заявл. 21.04.2010, опубл. 27.10.2011. Бюл. № 30.

20. Патент № 2435214 РФ, МПК G06F 17/30, H03M 7/00. Способ быстрого поиска в кодовой книге при векторном квантовании / Афанасьев А.А., Габдулгазиев С.Р.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2010103261/08, заявл. 01.02.2010, опубл. 27.11.2011. Бюл. № 33.

21. Патент № 2445719 РФ, МПК G10L 19/08. Способ улучшения восприятия синтезированной речи при реализации процедуры анализа через синтез в вокодерах с линейным предсказанием. / А.А. Афанасьев, М.В. Илюшини др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2010115914/08, заявл. 21.04.2010, опубл. 20.03.2012. Бюл. № 8.

22. Патент № 2445718 РФ, МПК G10L 19/00. Способ выделения сегментов обработки речи на основе анализа корреляционных зависимостей в речевом сигнале / А.А. Афанасьев, В.Г. Трубицын и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2010136618/08, заявл. 31.08.2010, опубл. от 20.03.2012. Бюл. № 8.

23. Патент № 2459373, МПК H04L 29/06, H04L 12/56. Способ определения длины кадра передачи кодеков речевых сигналов на основе линейного предсказания в сетях с пакетной коммутацией на основе IP-протокола / А.А. Афанасьев, О.Н. Титов; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2010151566/08, заявл. 15.12.2010, опубл. 20.08.2012. Бюл. № 23.

24. Патент № 2495504, МПК G10L 19/08. Способ снижения скорости передачи низкоскоростных вокодеров с линейным предсказанием / А.А. Афанасьев, А.П.

Рыжков; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2012126522/08, заявл. 25.06.2012, опубл. 10.10.2013. Бюл. №28.

25. Патент № 2504027, МПК G10L 19/09, G10L 19/038 Способ создания кодовой книги и поиска в ней при векторном квантовании данных / А.А. Афанасьев, А.П. Рыжков и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – №2012128241/08, заявл. 03.07.2012, опубл. 10.01.2014. Бюл. №1.

26. Патент №2580796, МПК G10L 19/00 , G06F 17/14 , H04B 15/00. Способ (варианты) фильтрации зашумленного речевого сигнала в условиях сложной помеховой обстановки / А.А. Афанасьев, О.Н. Титов и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2015107227/08, заявл. 02.03.2015, опубл. 10.04.2016. Бюл. №10.

27. Патент № 2606566, МПК G10L 15/08, G10L 15/00. Способ и устройство классификации сегментов зашумленной речи с использованием полиспектрального анализа / А.А. Афанасьев, О.Н.Титов и др.; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2014154081, заявл. 29.12.2014, опубл. 10.01.2017. Бюл. №1.

28. Патент № 2631968, МПК G10L 19/08 Способ низкоскоростного кодирования и декодирования речевого сигнала / А.А. Афанасьев, Е.С. Коробовский; патентообладатель гос. казенное образовательное учреждение высш. проф. образования Академия ФСО России. – № 2015127496, заявл.08.07.2015, опубл. 29.09.2017. Бюл. №28.

Доклады на конференциях:

29. Афанасьев, А.А. Обработка речевых данных на основе использования переменной длины сегмента анализа / А.А. Афанасьев // Сборник статей 2-й региональной научно-практической конференции "Многоядерные процессоры, параллельное программирование, системы обработки сигналов".–Барнаул:АГУ.–2012 – С.121–125.

30. Афанасьев, А.А. Адаптивный кодек речевых сигналов на основе систем с переменной структурой / А.А. Афанасьев // Цифровая обработка сигналов и ее применение: материалы 4-й международной конференции.– Москва: ИПУ РАН, 2002. – С.124–126.

31. Афанасьев, А.А. Модель системы с переменной структурой для кодирования речи / А.А. Афанасьев // Информатизация и информационная безопасность правоохранительных органов: материалы 11 Международной научной конференции. – Москва: 2002. – С. 362–365.

32. Афанасьев, А.А. Система адаптивного кодирования речевого сигнала / А.А. Афанасьев, А.В. Пищальников // Цифровая обработка сигналов и ее применение:

- материалы 5-й международной конференции и выставки. – Москва: ИПУ РАН, 2003. – С. 326–329.
33. Афанасьев, А.А. Особенности системы с переменной структурой для кодирования речевых сигналов / А.А. Афанасьев, Г.В. Богачев, О.О. Басов // Цифровая обработка сигналов и ее применение: материалы 6-й Международной конференции. – Москва: ИПУ РАН, 2004. – С.76–79.
34. Афанасьев, А.А. Анализ возможности перехода к системам с переменной структурой при эффективном кодировании речевого сигнала / А.А. Афанасьев, Г.В. Богачев, М.В. Илюшин // Цифровая обработка сигналов и ее применение: материалы 7-й международной конференции. – Москва: ИПУ РАН, 2005. – С.116–119.
35. Афанасьев, А.А. Применение программы технических расчетов MATLAB для решения задач цифровой обработки сигналов/ А.А. Афанасьев, А.А. Шустров// Материалы 10 Всероссийской научно-техн. конф. студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2005.– С.44-45.
36. Афанасьев, А.А. Построение систем речепреобразования для комплексов телекоммуникаций / А.А. Афанасьев, А.В. Петунин, А.С. Чуйкина // Материалы 11 Всероссийской научно-техн. конф. студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2006. – С. 62–63.
37. Афанасьев, А.А. Разработка классификатора речевых сегментов / А.А. Афанасьев, А.В. Петунин, А.С. Чуйкина // Материалы 11 Всероссийской научно-техн. конф. студентов, молодых ученых и специалистов.–Рязань, РГРТУ, 2006. – С. 63–64.
38. Афанасьев, А.А. Разработка и отладка систем на основе цифровых модулей Texas Instruments с использованием MATLAB / А.А. Афанасьев, Д.А. Жиров // Проектирование научных и инженерных приложений в среде MATLAB: труды всероссийской научной конференции.–С-Пб: 2007. – С.1379–1384.
39. Афанасьев, А.А. Синтез сигналов возбуждения в вокодере с линейным предсказанием / А.А. Афанасьев, Д.А. Жиров, И.В. Ульянов // Новые информационные технологии в научных исследованиях – материалы 13 всероссийской научно-технической конференции студентов.– Рязань: РГРТУ, 2008. – Часть 1. – С. 151–153.
40. Афанасьев, А.А. Аппроксимация значений линейных спектральных частот при низкоскоростном кодировании речи /А.А. Афанасьев //Новые информационные технологии в научных исследованиях и в образовании: материалы 15-й Всероссийской научно-технической конференции. – Рязань: РГРТУ, 2010. – С. 88–89.
41. Афанасьев, А.А. Моделирование речевого сигнала на сегментах одинаковой природы образования / А.А. Афанасьев, О.Н. Титов, И.В. Ульянов // Новые информационные технологии в научных исследованиях и в образовании: материалы 16 всероссийской научно-технической конференции студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2011.– С.77–79.

42. Афанасьев, А.А. Расчет линейных спектральных частот при кодировании речи на сегментах одинаковой природы образования / А.А. Афанасьев // Инновационные технологии: теория, инструменты, практика/ Материалы 3-й международной интернет-конференции (Inno-Tech 2011). – Пермь, 2012. – 5 с. [Электронный ресурс]: [сайт].– Режим доступа: <http://www.altstu.ru/media/f/102.pdf>. – Дата обращения: 5.06.2018.

43. Афанасьев, А.А. Формирование сигнала возбуждения в низкоскоростных вокодерах с линейным предсказанием / А.А. Афанасьев, М.В. Илюшин // Цифровая обработка сигналов и ее применение DSPA-2010: материалы 12-й Международной конференции и выставки//Москва: РНТОРЭС,2010.–С.214-217.

44. Афанасьев, А.А. Выделение параметров модели речевого сигнала на сегментах одинаковой природы образования / А. А. Афанасьев, Б.Р. Иванов, М.В. Илюшин // Цифровая обработка сигналов и ее применение DSPA-2011: Материалы 13-й Международной конференции и выставки.–Т.1,Москва:РНТОРЭС,2011.– С. 192-194.

45. Афанасьев, А.А. Выделение зависимостей элементов декомпозиции речевого сигнала при его обработке на основе процедур предсказания / А.А. Афанасьев, А.П. Рыжков // INTERMATIC–2012: материалы 11 Международной научно-технической конференции. –Москва: МГТУ МИРЭА– ИРЭ РАН,2012. – Часть 6. – С. 25–28.

46. Афанасьев, А.А. Использование биспектральных оценок для очистки речевого сигнала от шума / А.А. Афанасьев, О.Н. Титов // Современные проблемы радиоэлектроники: материалы докладов 16-й всероссийской с международным участием научно-технической конференции. – Красноярск, 2013. – С. 403–406.

47. Афанасьев, А.А. Уменьшение избыточности представления речевого сигнала при его кодировании на основе метода линейного предсказания/ А. А. Афанасьев // Цифровая обработка сигналов и ее применение DSPA-2013: материалы докладов 15-й международной конференции. Том1. – Москва: ИПУ РАН, 2013. – С. 202–205.

48. Афанасьев, А.А. Локализация подпространства представления сигналов возбуждения синтезирующего фильтра при низкоскоростном кодировании речи / А.А. Афанасьев // Цифровая обработка сигналов и ее применение DSPA-2014: материалы докладов 16-й международной конференции и выставки. Москва: ИПУ РАН, 2014. – Том 1. – С. 191–194.

49. Афанасьев, А.А. Потенциальные возможности понижения скорости передачи в системах кодирования речи / А.А. Афанасьев // Будущее технической науки: сборник материалов XIII международной молодежной научно-технической конференции. – Нижний Новгород, 2014. – С.46–47.

50. Афанасьев, А.А. Частотно-временная сегментация обрабатываемого речевого сигнала на основе полиспектрального анализа /А.А. Афанасьев, О.Н. Титов // Ин-

формационные системы и технологии: сборник материалов Международной научно-технической конференции. – Нижний Новгород: НГТУ, 2014. – С. 99–101.

51. Афанасьев, А.А. Вычисление параметров передаточной функции голосового тракта на однородных сегментах речевого сигнала при его низкоскоростном кодировании / А.А. Афанасьев, Р.С. Власов // Системы связи и радионавигации: сборник тезисов III всероссийской научно-технической конференции. – г. Красноярск: АО «НПП «Радиосвязь», 2016. – С.145- 148.

52. Афанасьев, А.А. Классификация сегментов зашумленного речевого сигнала с использованием акустических информационных критериев / А.А. Афанасьев, Титов О.Н.// Современные тенденции и инновации в науке и производстве: сборник материалов 2-й международной научно-практической конференции, г. Междуреченск: КузГТУ, 2014 – С. 245-247.

53. Афанасьев, А.А., Власов, Р.С. Вычисление параметров синтезирующей системы голосового тракта на однородных сегментах анализа при линейном предсказании речевого сигнала / А.А. Афанасьев, Р.С. Власов // Перспективные технологии в средствах передачи информации: материалы 12-ой международной научно-технической конференции. – Владимир: ВЛГУ, 2017. – Том 1. – С. 151-153.

54. Афанасьев, А.А. Снижение размерности кодовой книги сигналов возбуждения в гибридных кодеках с линейным предсказанием на основе учёта зависимости элементов декомпозиции речевого сигнала / А.А. Афанасьев // Перспективные технологии в средствах передачи информации: материалы 12-ой международной научно-технической конференции. – Владимир: ВЛГУ, 2017, том 1. – С.153-156.

55. Afanasjev, A.A. A Working out of Speech Compression Adaptive Algorithms Based on the Systems with Variable Structure/ A.A. Afanasjev // Proceedings of International Workshop "Speech and computer" SPECOM'2001. – Moscow: MSLU, 2001. – P.136 – 138.

56. Афанасьев, А.А. Цифровая обработка сигналов / А.А. Афанасьев, А.А. Рыболовлев, А.П. Рыжков // Учебное пособие. – Москва: Горячая линия - Телеком, 2017. – 356 с.

Подписано в печать _____.2018 г.
Формат 60 × 84/16. Печать офсетная.
Тираж 100 экз. Заказ № ____
ФГКВОУ ВО Академия ФСО России
302015 Орел, ул. Приборостроительная, 35