

Министерство науки и высшего образования Российской Федерации
Муромский институт (филиал)
федерального государственного бюджетного образовательного учреждения
высшего образования
**«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»
(МИ ВлГУ)**



На правах рукописи

Холкина Наталья Евгеньевна

УДК 681.324

**АЛГОРИТМЫ ОБРАБОТКИ РЕЧЕВЫХ СИГНАЛОВ
ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ В УСЛОВИЯХ ПОМЕХ**

Диссертация на соискание ученой степени

кандидата технических наук по специальности

2.2.15 – Системы, сети и устройства телекоммуникаций

Научный руководитель:

к.т.н., доцент, А. Ю. Проскуряков

д.т.н., профессор, Ю.А. Кропотов

Муром, 2022

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	5
ГЛАВА 1. ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ ПЕРЕДАЧИ РЕЧЕВОЙ ИНФОРМАЦИИ И ПОДХОДЫ К ОБРАБОТКЕ АКУСТИЧЕСКИХ РЕЧЕВЫХ СИГНАЛОВ.	14
1.1. ФОРМАЛИЗОВАННОЕ ОПИСАНИЕ АКУСТИЧЕСКИХ СИГНАЛОВ И АНАЛИЗ ИХ ПАРАМЕТРОВ.	14
1.2. ФОРМИРОВАНИЕ АКУСТИЧЕСКОГО ПОЛЯ СИГНАЛОВ ИНФОРМАЦИОННО- УПРАВЛЯЮЩИХ СИСТЕМ	19
1.3. ХАРАКТЕРИСТИЧЕСКИЕ ПАРАМЕТРЫ АКУСТИЧЕСКИХ СИГНАЛОВ	26
1.4. МОДЕЛИ АКУСТИЧЕСКИХ РЕЧЕВЫХ И МЕТОДЫ ИХ ОЦЕНИВАНИЯ.....	28
1.5. АНАЛИЗ АКУСТИЧЕСКИХ РЕЧЕВЫХ СИГНАЛОВ НА ОСНОВЕ ОЦЕНОК ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ.....	36
1.6. АНАЛИЗ СУЩЕСТВУЮЩИХ СИСТЕМ И РЕШЕНИЙ ГРОМКОГОВОРЯЩЕЙ СВЯЗИ И ОПОВЕЩЕНИЯ.	44
1.7. ВЫВОДЫ И ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЙ.	45
ГЛАВА 2. МЕТОДЫ ОЦЕНИВАНИЯ ПАРАМЕТРОВ МОДЕЛЕЙ АКУСТИЧЕСКИХ СИГНАЛОВ.....	48
2.1. ИССЛЕДОВАНИЕ ВЛИЯНИЯ ОТНОШЕНИЯ АКУСТИЧЕСКИЙ СИГНАЛ/ШУМОВАЯ ПОМЕХА НА СЛОГОВУЮ РАЗБОРЧИВОСТЬ.....	48
2.2. ПРИМЕНЕНИЕ МЕТОДА МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ДЛЯ ИССЛЕДОВАНИЯ ПАРАМЕТРОВ МОДЕЛИ АКУСТИЧЕСКИХ СИГНАЛОВ.....	52
2.3. ПОСТАНОВКА ЗАДАЧИ АППРОКСИМАЦИИ ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ АКУСТИЧЕСКИХ СИГНАЛОВ	56
2.4 ВЫВОДЫ	62
ГЛАВА 3. ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК АКУСТИЧЕСКИХ РЕЧЕВЫХ СИГНАЛОВ	63

3.1. МЕТОДЫ ИССЛЕДОВАНИЯ РАСПРЕДЕЛЕНИЯ ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ АКУСТИЧЕСКИХ СИГНАЛОВ	63
3.2. МОДЕЛЬ ОДНОМЕРНОЙ ФУНКЦИИ ПЛОТНОСТИ ВЕРОЯТНОСТИ ОЦИФРОВАННЫХ АКУСТИЧЕСКИХ РЕЧЕВЫХ СИГНАЛОВ.	66
3.3. ФОРМИРОВАНИЕ МОДЕЛЕЙ ВНЕШНИХ АКУСТИЧЕСКИХ ПОМЕХ В СИСТЕМАХ ГРОМКОГОВОРЯЩЕЙ СВЯЗИ И ОПОВЕЩЕНИЯ.....	73
3.4. ИССЛЕДОВАНИЕ АВТОКОРРЕЛЯЦИОННЫХ ФУНКЦИЙ (АКФ) РЕЧЕВЫХ СИГНАЛОВ И ШУМОВЫХ ПОМЕХ.....	76
3.5. ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК РЕЧЕВЫХ СИГНАЛОВ.....	81
3.6. ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК АКУСТИЧЕСКИХ ШУМОВЫХ ПОМЕХ.....	86
3.7. ВЫВОДЫ	88
ГЛАВА 4 АЛГОРИТМЫ ПОВЫШЕНИЯ ПОМЕХОЗАЩИЩЕННОСТИ АУДИООБМЕНА В ТЕЛЕКОММУНИКАЦИЯХ	90
4.1. ЛИНЕЙНАЯ ФИЛЬТРАЦИИ В ТЕЛЕКОММУНИКАЦИЯХ АУДИООБМЕНА	90
4.2. МЕТОД КОРРЕЛЯЦИОННО-ЭКСТРЕМАЛЬНОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ ЭХОСИГНАЛОВ	94
4.3. АДАПТИВНАЯ КОМПЕНСАЦИЯ ЭХОСИГНАЛОВ В СИСТЕМЕ ТЕЛЕКОММУНИКАЦИЙ ГРОМКОГОВОРЯЩЕЙ СВЯЗИ.	98
4.4. СОЗДАНИЕ СТРУКТУРЫ АЛГОРИТМА ПОДАВЛЕНИЯ АДДИТИВНЫХ АКУСТИЧЕСКИХ ПОМЕХ.....	103
4.5. РАЗРАБОТКА АДАПТИВНОГО МНОГОКАНАЛЬНОГО АЛГОРИТМА ПОДАВЛЕНИЯ АДДИТИВНЫХ И СОСРЕДОТОЧЕННЫХ АКУСТИЧЕСКИХ ПОМЕХ	110
4.6. АЛГОРИТМ АДАПТИВНОГО ПОДАВЛЕНИЯ АКУСТИЧЕСКИХ ПОМЕХ С ФОРМАНТНЫМ РАСПРЕДЕЛЕНИЕМ ПОЛОС РЕЖЕКЦИИ.....	114
4.7. ВЫВОДЫ	119
ЗАКЛЮЧЕНИЕ	122
ЛИТЕРАТУРА.....	124
ПРИЛОЖЕНИЕ А АКТЫ ВНЕДРЕНИЯ	138

ПРИЛОЖЕНИЕ В Свидетельство о внедрении интеллектуальной СОБСТВЕННОСТИ	141
ПРИЛОЖЕНИЕ С Листинг модулей программного обеспечения ТЕЛЕКОММУНИКАЦИОННОЙ СИСТЕМЫ ГРОМКОГОВОРЯЩЕЙ СВЯЗИ И ОПОВЕЩЕНИЯ С ШУМОПОДАВЛЕНИЕМ И ЭХОКОМПЕНСАЦИЕЙ.....	143

ВВЕДЕНИЕ

В диссертационной работе рассматриваются методы улучшения качества передачи речевой информации, исследуются характеристики акустических сигналов в системах телекоммуникаций и громкоговорящей связи. Речь, акустические речевые сигналы являются естественным носителем информации, используемой в целях общения между людьми, анализа окружающей обстановки, управления техническими системами. Посредством акустических сигналов решаются задачи связи и локации, как в воздушной, так и в водной среде. С их помощью осуществляется оценка состояния технических систем, и исследуются природные процессы.

Методам моделирования и исследованиям акустических сигналов посвящена многочисленная литература. При решении задач моделирования используются методы теории сигналов, корреляционного и спектрального анализа.

Продолжаются интенсивные исследования акустических речевых сигналов с применением регрессионного подхода, основанного на методе наименьших квадратов [48, 50]. Обладая достоинством получения решений при незначительном объеме априорных данных, эти методы имеют существенные преимущества, поскольку «опираются на жесткую априорную гипотезу, вытекающую из предположений о переменных, свободных от шума». Эти методы заключают в себе важное математическое содержание, позволяющее считать их «основным математическим инструментом в теории идентификации акустических сигналов в условиях шумов».

В широком смысле акустический сигнал является проявлением акустического поля и зависит от места расположения и ориентации микрофона в пространстве. Установление этой зависимости является в общем случае задачей акустики и требует применения при ее решении особых подходов, отличающихся по цели от перечисленных выше задач. А именно, требует рассмотрения характеристик поля как функций пространственных координат, но, как правило, не рассмотрения методов

обработки сигналов, инвариантных к особенностям акустического окружения. При этом необходимо иметь в виду, что в силу неопределенности граничных условий акустический сигнал, снимаемый с микрофона, не удастся представить с помощью какого-либо выражения, особенно в случае фрактальной структуры поля. Таким образом, акустический речевой сигнал можно представить в виде самостоятельного объекта исследований.

Работа посвящена исследованию, обработке и методам моделирования акустических сигналов в оперативно-командных телекоммуникационных системах обмена информацией и аудиообмена. Обмен информацией является важным средством обеспечения оперативного управления сложными объектами и обеспечивает их надежное функционирование, что предъявляет повышенные требования к достоверности передачи информации и к характеристикам эффективности телекоммуникационных систем [118]. Необходимость исследования и решения ряда проблем технологической связи обусловлена явлениями эха и реверберации, а также шумами и помехами природного и техногенного происхождения.

В работе приведены результаты исследования и обработки акустических сигналов, а также рассмотрены методы их моделирования в объектовых телекоммуникационных системах громкоговорящей связи, так как передача речевой информации является основным и важнейшим средством обеспечения оперативного управления сложными объектами, а качество ее передачи обеспечивает надежное функционирование последних. Это обуславливает предъявление повышенных требований к обеспечению качества передачи информации и к характеристикам эффективности телекоммуникационных систем [118].

На текущий момент такие системы активно исследуются и развиваются, как в теоретическом, так и в прикладном аспектах. Благодаря развитию современной микроэлектроники происходит совершенствование алгоритмов обработки речевой информации, исследуются такие важные для технологической связи характеристики как разборчивость, устойчивость и

надежность. Исследуются характеристики узнаваемости абонентов; вопросы борьбы с акустическим эхом, шумами и помехами; вопросы анализа вероятностных характеристик стандартизированных слогов; вопросы анализа процессов в системах с коммутируемыми фильтрами.

Методам повышения качества передачи речевой информации в системах телекоммуникаций, посвящены работы Фурдуева В.В., Сапожкова М.А., Скучика Е., Фланагана Дж., Покровского П.Б., Кропотова Ю.А., Cohen I., Hansler E. И др.

Однако, известные алгоритмы обработки сигналов, существующие методы увеличения эффективности телекоммуникационных систем обмена речевой информацией требуют проведения более тщательных исследований, разработки новых методов, требуют совершенствования и уточнения моделей информационных сигналов, разработки новых алгоритмов анализа и обработки речевых сигналов [39].

Объект исследования – речевые сигналы и акустические помехи, передаваемые по каналам телекоммуникационных систем.

Предмет исследования – методы, модели и алгоритмы анализа речевых сигналов, повышающие эффективность передачи аудиоинформации в телекоммуникационных системах.

Цель диссертационной работы – разработка методики и алгоритмов повышения помехоустойчивости телекоммуникационных систем передачи речевой информации.

Для достижения сформулированной цели были поставлены и решены следующие **задачи**:

1. Создание методики оценивания функции плотности вероятностей и спектральных характеристик отдельных сегментов речевого сигнала, наблюдаемых, в том числе, на фоне помех, в целях повышения эффективности адаптивного подавления акустических помех.

2. Разработка модели гистограммной оценки плотности вероятностей.

3. Разработка алгоритма выделения основного канала акустической обратной связи, потребность в котором присуща многим акустическим системам, в целях подавления эха и реверберации.

4. Разработка адаптивного алгоритма подавления эха и реверберации, совмещенного с алгоритмом выделения сигнала в основном канале акустической обратной связи, и с алгоритмом подавления акустических помех.

5. Определение области режекции в низкочастотной части спектра для повышения эффективности подавления акустических помех.

6. Разработка алгоритма компенсации и программного обеспечения адаптивного управления областями режекции для подавления сосредоточенных акустических помех.

7. Разработка адаптивного алгоритма подавления аддитивных акустических помех.

Научная новизна полученных результатов заключается в том, что:

1. Разработана методика оценивания слоговой разборчивости в системах телекоммуникаций, отличающаяся применением сеточной функции, обеспечивающая практическое оценивание слоговой разборчивости по отношению сигнал/помеха.

2. Разработана модель гистограммной оценки плотности вероятностей, отличающаяся аппроксимацией речевых сигналов по системе экспоненциальных функций и аппроксимацией акустических шумов по системе гауссовых функций, позволяющая получить погрешность оценивания не более 5%.

3. Разработан алгоритм формирования сигнала управления автоматическим выключением пораженных сосредоточенными помехами каналов, отличающийся применением адаптивной пороговой обработки, что позволяет получить отношение сигнал/помеха более 20 дБ и обеспечить слоговую разборчивость более 93%.

4. Разработан алгоритм подавления эхосигналов и сосредоточенных акустических помех, позволяющий подавить аддитивные и сосредоточенные акустические помехи на 30 дБ для обеспечения помехоустойчивости телекоммуникационных систем.

Методы исследования базируются на методах параметрической оптимизации, локальной аппроксимации и интерполяции, методах теории вероятностей и математической статистики, цифровой обработки сигналов, спектральном анализе.

Теоретическая и практическая значимость работы.

Практическая значимость результатов работы заключается в создании новых алгоритмов подавления аддитивных и сосредоточенных акустических помех, составивших основу комплекса программ по исследованию параметров и спектральных характеристик акустических сигналов, а также оценки эффективности предложенных алгоритмов.

Применение разработанных алгоритмов позволяет повысить эффективность и помехоустойчивость информационных коммуникаций в объектовых телекоммуникационных системах за счет комплексного решения следующих задач: подавления сосредоточенных помех, эхо компенсации и подавления аддитивных помех, а также достижения плановых показателей слоговой разборчивости речи не менее 93%.

Соответствие паспорту специальности. Содержание работы соответствует паспорту специальности ВАК 2.2.15 в части пунктов 2,8,14:

2. Исследование процессов генерации, представления, передачи, хранения и отображения аналоговой, цифровой, видео-, аудио- и мультимедиа информации; разработка рекомендаций по совершенствованию и созданию новых соответствующих алгоритмов и процедур.

8. Исследование и разработка новых сигналов, модемов, кодеков, мультиплексоров и селекторов, обеспечивающих высокую надежность

обмена информацией в условиях воздействия внешних и внутренних помех.

14. Разработка методов исследования, моделирования и проектирования сетей, систем и устройств телекоммуникаций.

Достоверность полученных результатов обеспечивается применением имитационного моделирования и подтверждается их согласованностью с экспериментальными данными, полученными при проектировании телекоммуникационных систем аудиообмена, функционирующих в условиях помех, апробацией результатов исследования на научных конференциях различного уровня.

Результаты диссертационной работы:

Внедрены в программное обеспечение оперативно-командных телекоммуникационных систем громкоговорящей связи, трансляции и оповещения в ходе выполнения ОКР «Модернизация комплекса оперативно-командной громкоговорящей и телефонной связи «КТС-01ЦС»», выполняемой АО «Муромский радиозавод».

Используются в учебном процессе в Муромском институте (филиале) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» при подготовке бакалавров и магистров по направлению 09.03.01, 09.04.01 «Информатика и вычислительная техника».

Применение результатов исследований в учебном процессе и в производстве подтверждено актами внедрения, приведенными в приложении А.

Основные положения, выносимые на защиту:

1. Методика оценивания слоговой разборчивости в системах информационных коммуникаций, которая позволяет осуществить практическое оценивание слоговой разборчивости по отношению сигнал/помеха в телекоммуникационных системах и громкоговорящей связи.

2. Модель гистограммной оценки плотности вероятностей, позволяет получить погрешность оценивания не более 5%.

3. Алгоритм формирования сигнала управления автоматическим выключением пораженных сосредоточенными помехами каналов, который позволяет получить отношение сигнал/помеха более 20 дБ и обеспечить слоговую разборчивость более 93%.

4. Адаптивный многоканальный алгоритм подавления эхосигналов и сосредоточенных акустических помех, который обеспечивает помехоустойчивость телекоммуникационных систем.

Апробация работы. Основные результаты работы доложены на 23-ей Международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2021), опубликованы в материалах и сборниках тезисов докладов на всероссийских и международных конференциях: IV Международной конференции и молодежной школе «Информационные технологии и нанотехнологии» (ИТНТ-2018), XV Международной научно-практической конференции «Инновационные, информационные и коммуникационные технологии» (Москва, 2019), X-XII Всероссийских научных Зворыкинских чтениях «Наука и образование в развитии промышленной, социальной и экономической сфер регионов России» (Муром, 2018-2022); научно-технических семинарах кафедры «Электроника и вычислительная техника» Муромского института в 2017-2022.

Диссертационные исследования выполнялись в рамках госбюджетных научно-исследовательских работ: ГБ НИР МИ ВлГУ № 376/01 «Анализ и синтез электронных устройств, систем и сигналов с применением компьютерных технологий» (2013-2017 гг., 2018-2022 гг.).

Публикации по работе. По теме диссертации опубликовано 20 работ, в том числе 3 статьи в ведущих рецензируемых изданиях из перечня ВАК, 2 статьи проиндексированы в наукометрической базе SCOPUS, 8 статей в центральных изданиях, 6 докладов на научных конференциях; получено 1 свидетельство о государственной регистрации программы для ЭВМ.

Личный вклад. Все результаты, изложенные в работе, получены автором лично или при его непосредственном участии.

Постановка цели и задач, обсуждение планов исследований и полученных результатов выполнены совместно с научным руководителем.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы из 127 наименований и приложений. Работа включает 157 страниц машинописного текста, содержит 48 рисунков и 9 таблиц.

Краткое содержание диссертационной работы.

Во введении обоснована актуальность темы, определены объект и предмет исследования, сформулированы цель и задачи диссертационной работы, определены ее научная новизна и практическая значимость, представлены положения, выносимые на защиту, приведены сведения об апробации работы и публикациях.

В первой главе рассмотрено современное состояние вопроса разработки алгоритмов и программных структур обработки сигналов в телекоммуникациях аудиообмена, функционирующих в помеховой обстановке. Приведены общие характеристики сигналов информационно-управляющих систем, рассмотрены параметры акустических сигналов, а также вопросы анализа случайных информационных сигналов. Сформулированы основные задачи исследований.

Во второй главе проанализировано влияние, которое оказывают на слоговую разборчивость и полноту восприятия информации абонентами системы такие параметры, как интегральный индекс артикуляции, определяющий формантную разборчивость, а также отношение сигнал/акустический шум, воздействующий на систему передачи сообщений извне. Исследованы параметры модели акустических сигналов.

В третьей главе разработаны модели плотности вероятностей акустических речевых сигналов и акустических шумовых помех,

исследуются корреляционные и спектральные функции параметров акустических сигналов.

В четвертой главе рассмотрены алгоритмы повышения помехозащищенности: разработан метод управления областью режекции, алгоритм формирования сигнала управления автоматическим выключением пораженных сосредоточенными помехами каналов, адаптивный многоканальный алгоритм подавления аддитивных и сосредоточенных помех.

ГЛАВА 1. ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ ПЕРЕДАЧИ РЕЧЕВОЙ ИНФОРМАЦИИ И ПОДХОДЫ К ОБРАБОТКЕ АКУСТИЧЕСКИХ РЕЧЕВЫХ СИГНАЛОВ.

1.1. Формализованное описание акустических сигналов и анализ их параметров.

Отличительной особенностью информационно-управляющих систем является наличие значительного объема информации, поступающей от источников в форме разного рода акустических сигналов. Эта информация может быть заключена в речевых сообщениях, поступающих по каналам технологической связи, в акустических и механических колебаниях, обусловленных движущимися объектами и неисправными механизмами и машинами. Во многих случаях информационные сигналы принимаются на фоне значительных акустических помех, что ставит задачу их подавления. Необходимо также учитывать, что информационно-управляющие системы относятся к классу человеко-машинных систем, в которых речевые сообщения обычно передаются в форме коротких команд.

Таким образом, акустический сигнал – это локальная характеристика акустических волн, наблюдаемых (измеряемых) в заданной точке пространства с помощью соответствующего датчика или прибора. При этом наблюдаемый сигнал зависит от целого ряда факторов, а именно, от местоположения и ориентации датчика (микрофона), от характеристик измерительного прибора, источника акустических волн и канала распространения звуковых колебаний. Каждый из указанных факторов должен быть представлен адекватной ему математической моделью.

Под математической моделью понимается описание объекта некоторой системой алгебраических, дифференциальных или разностных уравнений. Сложность модели принято согласовывать с целью задачи, априорной информацией об объекте и вычислительными возможностями используемых аппаратных средств. Это обуславливается тем, что детализация модели при

определенных условиях может не только не повышать, но даже снижать точность решения задачи в целом. Поэтому предпочтение обычно отдается более простым моделям, если, они обеспечивают необходимую точность, то есть являются адекватными данным, получаемым из эксперимента. Такие упрощенные модели называются имитационными.

Таким образом, построение модели сигналов в системах телекоммуникаций громкоговорящей связи основывается на структуре абонентского устройства передачи речи, который представлен на рисунке 1.1.

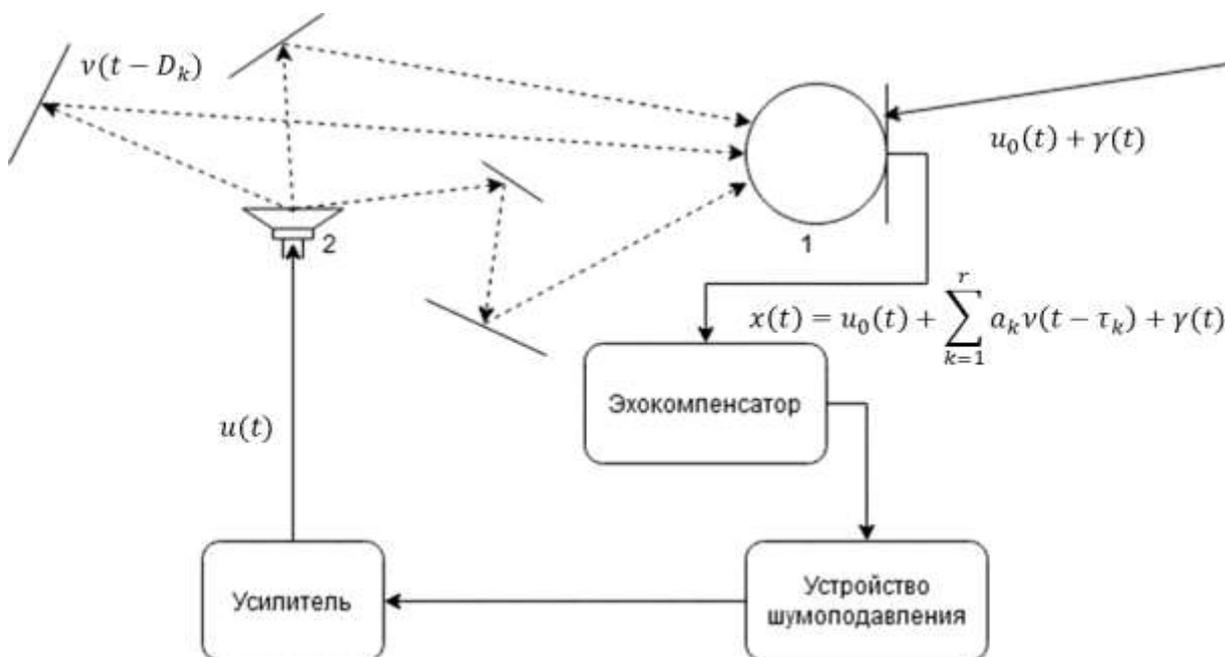


Рисунок 1.1 – Структурная схема абонентского устройства передачи речевой информации в системе телекоммуникаций

Представленная на рисунке 1.1 модель, содержит следующие элементы: 1 – микрофон, 2 – устройство излучения информационного сигнала.

Рассмотрим модель эхосигнала $u(t)$, которую можно представить в виде комбинации многочисленных копий истинного воспроизводимого сигнала $v(t)$, запаздывающих на некоторые временные промежутки [40]. Её можно представить следующим выражением:

$$u(t) = \sum_{k=1}^r \alpha_k v(t - \tau_k),$$

в котором $v(t) = \int_0^t f(\theta)u_0(t - \theta)d\theta \div F(s)U_0(s)$.

В данном выражении, формализующем вышеуказанную модель, τ_k и α_k являются, соответственно, временем задержки запаздывания и параметром значения амплитуды запаздывающих копий сигналов, которые попадают на входы принимающего устройства по пути с номером k , $f(\theta)$ - является импульсной функцией, обеспечивающей достаточное усиление входных сигналов $u_0(t)$, символ \div можно считать знаком, определяющим эквивалентность изображения функций её оригиналу, при этом r определяет количество возможных путей по которым распространяются запаздывающие эхосигналы.

Математическую модель источника акустического сигнала, являющегося следствием функционирования механизмов и машин, можно представить системой обыкновенных дифференциальных уравнений [5, 91, 103, 108, 116], то есть акустическое поле можно описать дифференциальными уравнениями, являющимися следствием законов непрерывности, сохранения энергии, а также уравнениями состояния жидкости или газа. Использование этих уравнений позволяет установить характер распространения звука в неоднородной среде, в частности, распространения звука в атмосфере при изменении температуры и скорости ветра с высотой над поверхностью земли. Форма звуковых лучей в атмосфере при изменении температуры с высотой приведена на рисунке 1.2, а при изменении скорости ветра – на рисунке 1.3. Анализ распространения звука в неоднородной среде, например в атмосфере с зависящей от высоты температуры воздуха и скорости ветра, указывает на возможность образование зон молчания.

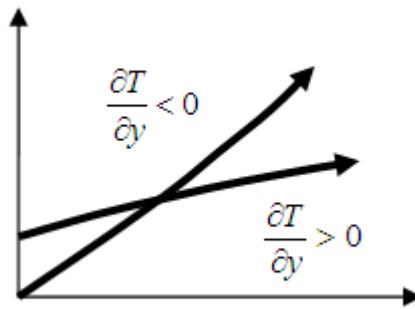


Рисунок 1.2 – Ход лучей звука при изменении с высотой температуры

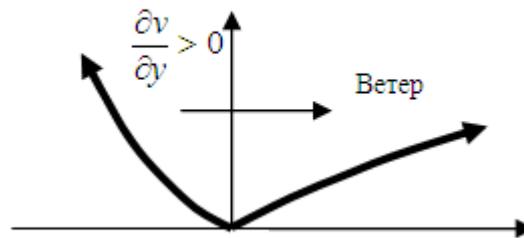


Рисунок 1.3 – Ход лучей звука при росте с высотой скорости ветра

Однако даже в линейном приближении названные уравнения в частных производных слишком сложны для практического использования, что ограничивает их применение, как правило, простыми задачами, за исключением задач теории поля, где они играют основополагающую роль. В области задач обработки сигналов более подходящими являются упрощенные модели, основанные на методах геометрической акустики и акустических линий с распределенными параметрами [23, 113, 114], которые, например, используются в качестве моделей каналов передачи информации аудиосигналами и акустических полостей, обеспечивающих требуемую акустику помещений [7, 14, 23].

Применение методов геометрической акустики приводит к моделям, описываемых дифференциальными уравнениями с запаздыванием или, в общем случае, функционально-дифференциальными уравнениями. Нестационарный характер сигналов может быть учтен, если принять, что коэффициенты указанных уравнений изменяются во времени, то есть, модель относится к классу систем с переменными параметрами [21, 79]. Аналогично

может быть представлен и нестационарный сигнал источника акустического поля.

Устройства связи и оповещения, используемые в информационно-управляющих системах и оповещения, являются, фактически, элементами человеко-машинных систем управления сложными объектами. Такие устройства должны обеспечивать минимальные затраты времени на надежное и оперативное доведение командной информации до соответствующих абонентов, что особенно важно в аварийных ситуациях. Это предполагает реализацию высокой разборчивости речи и, соответственно, ставит задачу подавления акустических помех и эхосигналов.

Принципы формирования поля распространения акустических сигналов следующие: оно существенно зависит от геометрических параметров той области и того пространства, в котором осуществляется озвучивание акустических речевых сообщений системами громкоговорящей связи и оповещения [39, 96, 102]. Если системы громкоговорящей связи применяются на открытых площадках, то формирование акустического поля реализуется с применением законов геометрической акустики, его можно охарактеризовать существенными задержками, имеющими место в каналах распространения эхосигналов от объектов на местности, при этом сигналы в значительной степени рассеяны. В случае работы систем громкоговорящей связи в замкнутых помещениях и пространствах, длины волн колебаний акустических сигналов могут быть существенно меньше габаритных параметров объектов располагающихся в замкнутом пространстве, но вышеуказанные принципы распространения речевых сигналов также можно рассматривать. При этом значимым может стать дополнительный эффект реверберации, которая появляется благодаря множественным отражениям сигналов от многочисленных поверхностей [40]. Схема формирования акустического поля в условиях множественных отражений приведена на рисунке 1.4.

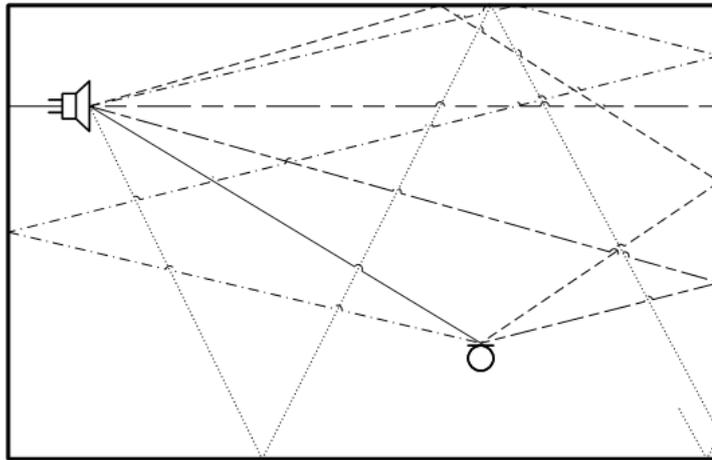


Рисунок 1.4 – Схема формирования акустического поля

1.2. Формирование акустического поля сигналов информационно-управляющих систем

Вопросам формирования и восприятия акустических речевых сигналов, продолжают посвящаться достаточно большое число исследований, проводимых в рамках обработки и восстановления, выяснения механизмов восприятия, машинного распознавания, а также сжатия и кодирования методами линейного и нелинейного предсказания. В значительной мере все эти исследования направлены на уменьшение влияния акустических помех, шумов и эха. При этом решение этой задачи осложнено нестационарным характером сигналов и, в первую очередь аудиосигнала, один из фрагментов которого приведен на рисунке 1.5.

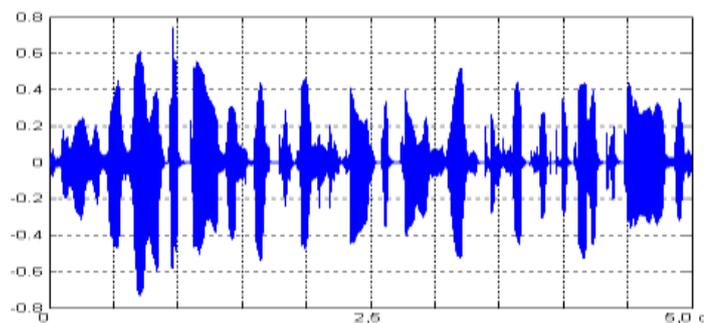


Рисунок 1.5 – Фрагмент речевого аудиосигнала

Речевые сигналы можно отнести к классу нестационарных, поэтому для их подробного анализа можно использовать различные виды

сегментации в том числе фонетическую и лингвистическую, которая проводится в частотном диапазоне 50-200 Гц, с учетом длины интервалов рассмотрения сигнала в несколько десятков или сотен миллисекунд. При этом, чтобы проанализировать характеристики разделения акустических звуков речи, может быть применен график, отражающий зависимость интенсивности речи от времени. Полученные интервалы рассмотрения сигналов могут быть охарактеризованы рядом медленно меняющихся характеристик и параметров фонетической модели речи, таких, как частота основного тона речевого сигнала, частоты формантных составляющих речи, массивы значений коэффициентов линейных функций модели прогноза. К основным параметрам интенсивности речевых сигналов можно отнести среднюю энергию сигнала за кратковременный промежуток времени, среднее значение мощности акустического речевого сигнала. Применение вышеуказанных параметров не обеспечивает нахождение и исследование деталей акустических речевых сигналов даже при использовании предварительного сглаживания сигнала [47].

Формирование речевого аудиосигнала может быть определено многочисленными трудно формализуемыми процессами в артикуляционном аппарате говорящего человека. Они зависят от особенностей работы многочисленных систем и органов диктора, таких как: диафрагма, легкие, голосовые связки, органы ротовой полости. Воздушный поток, нагнетаемый легкими и регулируемый голосовыми связками, возбуждает в случае вокализованных звуков акустические колебания, близкие к гармоническому виду, а в случае фрикативных и взрывных звуков – почти хаотические колебания [82].

Первыми моделями аппарата формирования речи являлись наборы резонансных контуров с частотами, отвечающими значимым формантам речевого аудиосигнала. В связи с развитием техники цифровой обработки сигналов внимание переместилось на методы линейного и нелинейного предсказания, методы динамических и нестационарных систем.

Не менее простыми являются модели слуховой системы, периферический отдел которой состоит из базилярной мембраны, кортиева органа, средней лестницы, рейсснеровой мембраны и слухового нерва. Здесь привлекают к себе внимание физические механизмы, лежащие в основе функционирования базилярной мембраны, основная особенность которой заключена в сочетании высокой частотной избирательности с малым временем переходных процессов. Другая особенность слухового аппарата, которая влияет на маскирующие свойства акустических шумов и помех, заключена в зависимости его чувствительности от частоты сигнала. Подобными особенностями может обладать модель в форме механической линии с распределенными параметрами с подключенным к ней набором слабо связанных резонаторов. Здесь нельзя не учитывать, что многие процессы слухового восприятия происходят на уровне нейронной сети [75].

В обобщенной постановке можно принять, что слуховая система выполняет функции частотно-временного анализа, разнообразные варианты которого, основанные на кратковременном преобразовании Фурье, теории всплесков и атомарных функций, распределениях Вигнера – Вилия и других, используются во многих прикладных и нейрофизиологических исследованиях.

В соответствии с приведенными выше сведениями, основные проблемы обработки сигналов информационно-управляющих систем заключаются в задачах выделения информационных сигналов из разного рода акустических шумов, помех и эха [7, 12], то есть в подавлении последних. Это приводит к повышению параметров устойчивости систем, обладающих акустическими обратными связями и возможности создания адекватных моделей каналов распространения, приемников и формирователей звуковых колебаний. Похожие проблемы возникают в задачах пространственной селекции источников звука, оценивания характеристик и диагностики состояния устройств и каналов передачи информации.

Проблема подавления нежелательных компонентов акустического сигнала по своей постановке включает в себя два аспекта. В рамках первого аспекта задача заключается в исключении нежелательных компонентов из сигнала, предназначенного для передачи по каналу передачи аудиосигналов или человеко-машинного взаимодействия. Особая роль этой задачи связана с использованием техники сжатия речи и речевого ввода информации в систему, стандартные процедуры которых являются, как известно, весьма чувствительными к шумам, помехам и эхосигналам.

Второй аспект связан с повышением разборчивости речи в помещениях и на территориях с высоким уровнем помех и эха, например, аэропортов. Одно из решений этой задачи заключается в создании локальных зон тишины, которое, в принципе, может быть получено посредством адаптивного подавления нежелательных компонентов сигнала. Однако в таких условиях более реальным способом повышения разборчивости является применение головных телефонов со встроенной системой адаптивной компенсации помех и эха.

С задачей подавления акустического эха связана проблема устойчивости, характерная для систем громкоговорящей связи [54, 66, 96]. Имеется специальный подход к решению вышеуказанной проблемы, базирующийся на специальной технике управления голосом диктора, обеспечивающей регулировку параметров затухания сигналов в различных приемо-передающих каналах звуковых трактов абонентов [104]. В основе подобного управления может лежать контроль за интенсивностью и громкостью речи абонентов, участвующих в конференции. Другой способ основывается на алгоритмах адаптивной компенсации эхосигналов.

От телефонии системы громкоговорящей связи отличаются значительно более высокими значениями мощности и длительности эха, что заметно усложняет задачу компенсации в последнем случае, в частности, по причине увеличения порядка адаптивных компенсаторов эха [66]. При этом если в телефонии длительность эха, обусловленного по преимуществу

электрическими характеристиками каналов передачи, имеет величину порядка 30 мс, то эхо в системах громкоговорящей связи, носящее в основном акустический характер, имеет величину в несколько сотен миллисекунд, что требует применения адаптивных фильтров порядка 256 и 4000, соответственно.

В обоих случаях адаптивная компенсация сопряжена с задачей обнаружения встречного сигнала, при наличии которого настройка адаптивного фильтра приостанавливается [1]. Необходимые для решения этой задачи методы сегментации, реализуются более просто в телефонии [22].

Структурная схема системы громкоговорящей связи [67] и соответствующей ей граф сигналов, приведенные на рисунке 1.6, раскрывает состав контуров акустической и электрической обратной связи.

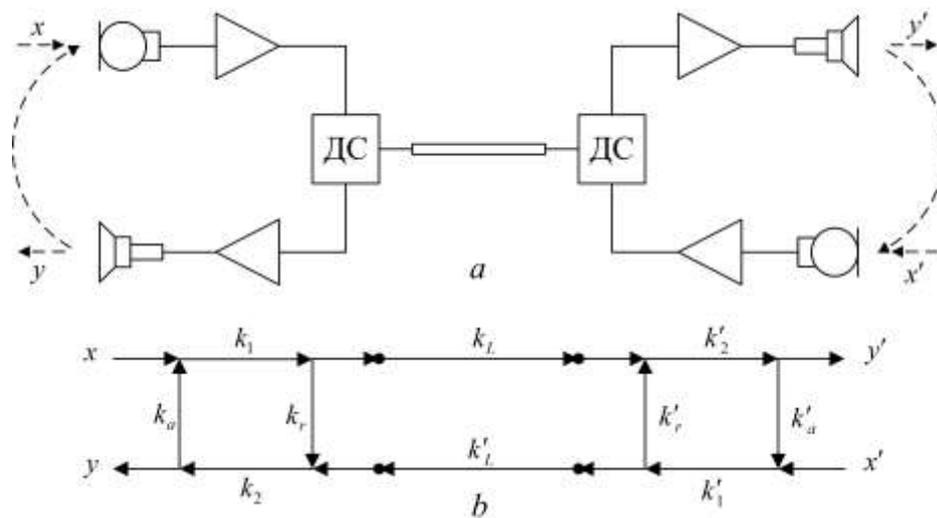


Рисунок 1.6 – Общая структурная схема (а) и граф сигналов (b) системы громкоговорящей связи

Представим матричное уравнение вышеуказанной системы:

$$\begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}, \quad (1.1)$$

в котором

$$t_{11} = \frac{1}{\Delta} k_1 k_2 (k_r + k_L k'_L k'_1 k'_2 k'_a - k_r k'_r k'_1 k'_2 k'_a), \quad t_{12} = \frac{1}{\Delta} k_2 k'_L k'_1,$$

$$t_{21} = \frac{1}{\Delta} k_2' k_L k_1, \quad t_{22} = \frac{1}{\Delta} k_1' k_2' (k_r' + k_L k_L' k_1 k_2 k_a - k_r k_r' k_1 k_2 k_a) \text{ и}$$

$$\Delta = (1 - k_1 k_2 k_a k_r')(1 - k_1' k_2' k_a' k_r') - k_1 k_2 k_a k_L k_1' k_2' k_a' k_L'. \quad (1.2)$$

Приведенные в вышеуказанных соотношениях k_a , k_a' являются коэффициентами передачи, которые характеризуют каналы акустической связи в тракте между устройством воспроизведения (динамиком) и устройством ввода речевой информации (микрофоном), коэффициенты k_r , k_r' демонстрируют отличную от идеальную балансировку параметров дифференциальных систем, k_1 , k_1' - это передаточные функции в каналах передачи, k_2 , k_2' - это передаточные функции в принимающих трактах, k_L , k_L' - это передаточные функции физических линий связи [16, 72, 76]. Те ветви размеченного графа, приведенные на рисунке 1.6. и не помеченные обозначениями обладают коэффициентом передачи, равным 1 [40].

Акустическая обратная связь по электрическому каналу обратной связи достаточно просто устраняется посредством адаптивной компенсации [1, 77], поскольку длительность его импульсного отклика не превышает 30 мс. Акустическая обратная связь присутствует во всех случаях, независимо от наличия или отсутствия обратной электрической связи. Схема системы только с акустической связью приведена на рисунке 1.7.

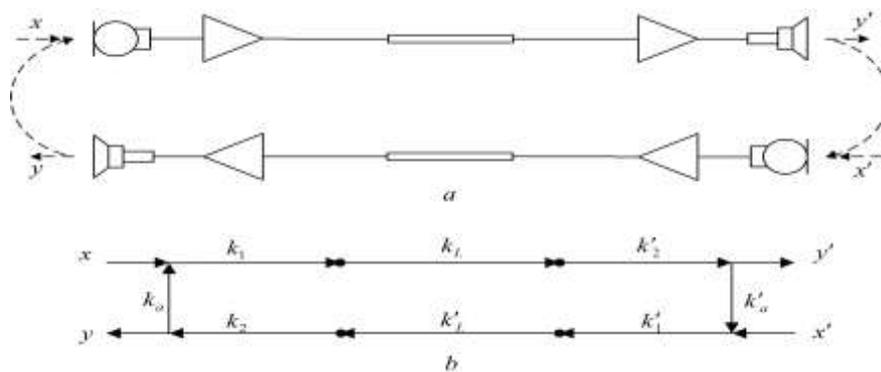


Рисунок 1.7 – Схема (a) и граф сигналов (b) системы с акустической обратной связью

В общем случае акустическое поле как информационной составляющей сигнала, так и помех не является изотропным. Интенсивность колебаний

может сильно зависеть от формы озвучиваемой области и особенностей размещения в ней источников и иного оборудования. Именно по этой причине возможен эффект от применения направленных микрофонов, но при этом только в случае правильного выбора их ориентации в пространстве. Методы адаптивной компенсации или подавления помех позволяют в этом плане достичь лучших результатов. Полезной в этом плане является оценка оптимального, по критерию минимума дисперсии восстанавливаемого сигнала, размера области наблюдений в случайном поле.

Структура абонентского прибора с адаптивной компенсацией помех и эха на передающей стороне по входному информационному сигналу приведена на рисунке 1.8.

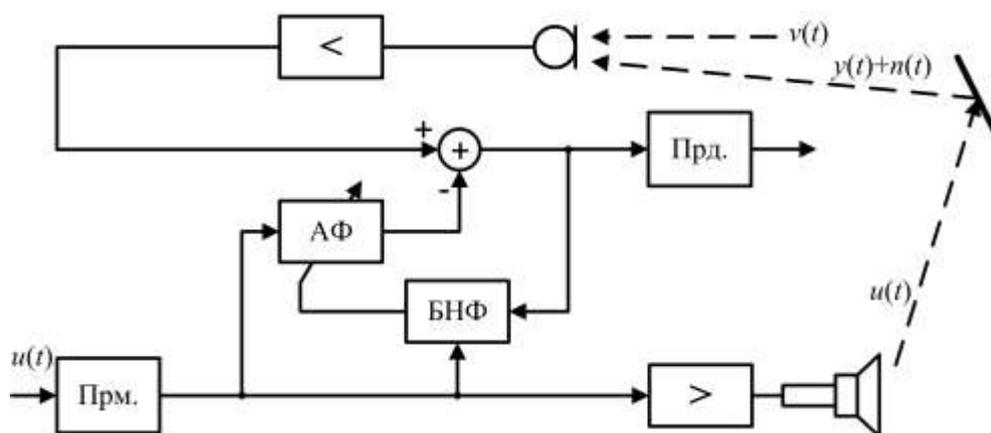


Рисунок 1.8 – Адаптивная компенсация по входному информационному сигналу

На этом рисунке $u(t)$ — входной информационный сигнал и акустический сигнал, формируемый громкоговорителем (без учета вносимых искажений), $y(t)$ — эхосигнал, $v(t)$ — встречный акустический речевой сигнал и $n(t)$ — акустические шумы и помехи. В состав прибора входят приемное (Прм.) и передающее (Прд.) устройства, адаптивный фильтр (АФ) и блок его настройки (БНФ), а также микрофон и громкоговоритель. Порядок адаптивного фильтра, как уже отмечалось, должен иметь в общем случае большую величину, что увеличивает время настройки. При определенных условиях порядок фильтра можно уменьшить, воспользовавшись набором опорных сигналов – копий входного сигнала,

запаздывающих на характерные значения времени распространения эха.

Любое изменение ориентации микрофона пользователя отрицательно влияет и на работу схемы, приведенной на рисунке 1.8, поскольку сопровождается перенастройкой адаптивного фильтра. Возможным решением этой проблемы в обоих случаях является совмещение микрофонов основного и компенсационного канала в одном устройстве.

1.3. Характеристические параметры акустических сигналов

Наиболее полная информация о свойствах случайных акустических сигналов содержится в их одномерных и многомерных функциях распределения плотностей вероятностей [2, 12, 60, 61, 62, 73, 79]. Однако на практике удается использовать, как правило, только одномерные распределения стационарных процессов, а в случае процессов Маркова – еще и переходные плотности вероятностей. При этом в ряде случаев одномерными распределениями описываются и нестационарные процессы. Так образом на больших интервалах описываются, например, акустические аудиосигналы, имеющие по определению нестационарный характер.

В случае рассмотрения одномерных функций распределения плотности вероятности акустических речевых сигналов [40] можно применять сведения из [11].

Для получения вышеуказанных функций применимо, например, гамма – распределение, подробно описанное в [11, 53], в соответствии с выражением:

$$f(x) = \frac{\lambda^\alpha}{2\Gamma(\alpha)} |x|^{\alpha-1} e^{-\lambda|x|}, \quad (1.3)$$

в котором, параметры α и λ принимают положительные значения, $\Gamma(\alpha)$ является гамма-функцией, $\alpha\Gamma(\alpha) = \Gamma(\alpha + 1)$.

Если параметр α , принимает значение, равное 1, то выражение (1.3), определяющее распределение Лапласа, принимает вид, характеризующий функцию распределения плотности вероятностей

$$f(x) = \frac{1}{2} \lambda e^{-\lambda|x|}.$$

Имеются и другие возможные способы и выражения для определения одномерных функций плотности распределения. В [39] показано, что одним из оптимальных подходов к формированию одномерной плотности вероятностей для англоязычной речи является метод, основанный на применении функции следующего вида

$$f(x) = \frac{0,6}{\sqrt{2u_1}} e^{-\frac{\sqrt{2}}{u_1}|x|} + \frac{0,4}{\sqrt{2\pi u_2}} e^{-\frac{1}{2u_2}x^2}, \quad (1.4)$$

а для русского языка – более подходящей является следующая функция [40]

$$f(x) = \frac{0,25}{u_0} \left(e^{-\frac{1}{u_0}|x|} + 10e^{-\frac{10}{u_0}|x|} \right). \quad (1.5)$$

Для формирования функций плотности распределения вероятностей важное значение также имеют характеристики применяемых устройств ввода речевой информации (микрофонов).

Для решения задач анализа акустических речевых сигналов в работе будет применена функция плотности распределения вероятности в виде [11, 53]

$$f(x) = \frac{\eta k_\alpha}{2\Gamma(\alpha)u_0} \left(\frac{k_\alpha}{u_0} |x| \right)^{\alpha-1} e^{-\frac{k_\alpha}{u_0}|x|} + \frac{1-\eta}{\sqrt{2\pi}u_0\beta} e^{-\frac{1}{2u_0^2\beta^2}x^2}. \quad (1.6)$$

Вид выражений (1.3)-(1.6) обеспечивает низкую сложность расчета параметров модели плотности вероятности по полученным экспериментально данным. При этом возможно применение и других подходов к аппроксимации, базирующихся, в том числе на оптимизации функции плотности распределения со следующими ограничениями [49, 55]:

$$f(x) \geq 0 \text{ и } \int_{-\infty}^{\infty} f(x)dx = 1. \quad (1.7)$$

Однако, как показано в исследованиях в главе 2, вышеуказанные виды одномерной плотности вероятностей имеют значительную погрешность. По

экспериментально полученной одномерной плотности вероятностей можно, в частности, определить наличие в каналах нелинейных искажений типа ограничений. В нестационарных условиях это требует сегментации сигнала либо по уровню, либо по моментам изменения его функции распределения. В обоих случаях оценивание функций распределения осуществляется по ограниченному набору выборок сигнала, что представляет собой и самостоятельный интерес, в частности, в плане классификации отдельных звуков.

Полученные таким образом плотности вероятностей акустического речевого сигнала совсем не обязаны описываться приведенными выше выражениями и при аппроксимации предпочтительно использовать универсальные методы аппроксимации. Этим проблема сегментации не исчерпывается, и требуются числовые характеристики, позволяющие оценивать степень расхождения между различными распределениями, некоторые из которых представлены в следующем параграфе.

Сложность аппроксимации и интерпретации многомерных плотностей вероятностей обуславливает применение моментных функций и их преобразований Фурье. Начальные и центральные моментные функции будут при этом называться соответственно корреляционными и ковариационными функциями.

1.4. Модели акустических речевых и методы их оценивания

Задачи обработки сигналов в системах аудиообмена обуславливаются потребностями выделения информации, подавления помех, повышения устойчивости систем связи и компенсации эха. К указанным задачам относятся задачи диагностики объектов по излучаемым шумовым сигналам, задачи повышения эффективности систем связи.

В связи с необходимостью совершенствования методов обработки аудиосигналов можно поставить задачи:

– оценивания и аппроксимации одномерных функций распределения и корреляционных функций по ограниченным наборам данных и, на этой основе, формирования базы априорной информации и выделения интервалов стационарности наблюдаемых сигналов;

- сглаживания и локальной аппроксимации наблюдаемых сигналов и, на этой основе, сегментации нестационарных сигналов;
- обнаружения и оценивания гармонических сигналов, сигналов с дискретным спектром на фоне акустического шума с непрерывным распределением и, на этой основе, диагностики объектов;
- спектрального анализа наблюдаемых сигналов и, на этой основе, разрешения источников излучения и селекции речевого сигнала и помехи;
- формирования модели эхосигналов и оценивания ее параметров;
- идентификации нестационарных сигналов и создания нестационарных моделей под наблюдаемые данные;
- адаптивной компенсации акустических помех и эхосигналов с применением многоканальной и многоскоростной обработки.

Одномерные функции распределения являются достаточно простым средством различения сигналов, которым можно воспользоваться в целях различения речи и шума. Аппроксимация функций распределения речи, может быть сформулирована как задача оптимизации с ограничениями. А именно, как задача минимизации квадрата нормы

$$\|f(\theta) - \hat{f}\|^2 \rightarrow \min ,$$

$$P(x, \theta) \geq 0, \int_{a_0}^{a_n} P(x, \theta) dx = 1. \quad (1.8)$$

Здесь \hat{f} – вектор значений гистограммы, характеризующей распределение выборок x наблюдаемого сигнала. Компонентами вектора $f(\theta)$ являются, соответственно, вероятности, с которыми значения наблюдаемого сигнала попадают в интервалы гистограммы $[a_{k-1}, a_k]$, $k = 1, \dots, n$.

То есть $f_k(\theta) = \int_{a_{k-1}}^{a_k} P(x, \theta) dx$. При этом $P(x, \theta)$ является аппроксимацией искомой функции плотности вероятности, а θ – вектором параметров, по которым осуществляется минимизация. В целях регуляризации на вектор θ параметров могут быть наложены и дополнительные ограничения.

Статистические характеристики вектора параметров, представляющего решение задачи (1.8), существенно зависят от объема используемых выборок. Установление этой зависимости является необходимым основанием для достижения необходимой достоверности полученных результатов аппроксимации.

К другой значимой характеристике речевых сигналов можно отнести АКФ - автокорреляционную функцию, значение которой находится из выражения

$$R_{xx}(t, \tau) = E\{x(t)x(t + \tau)\}.$$

При условии, что процесс стационарный, АКФ находится в зависимости от единственного аргумента и может быть представлена выражением

$$R_{xx}(\tau) = E\{x(t)x(t + \tau)\}.$$

Чтобы получить спектральную плотность мощности рассматриваемого стационарного процесса требуется выполнить преобразование Фурье от полученной АКФ [39]:

$$\Phi_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j\omega\tau} d\tau.$$

Подобным образом можно рассчитать значение взаимной функции корреляции

$$R_{xy}(\tau) = E\{x(t)y(t + \tau)\}$$

а также получить взаимный спектр акустического речевого сигнала [40]

$$\Phi_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j\omega\tau} d\tau.$$

Если рассматривать дискретное представление, то взаимная корреляционная функция может быть представлена выражением [40]

$$\hat{R}_{xy}(k) = \frac{1}{N} \sum_{i=0}^{N-1} x(i) \cdot y(k + i).$$

Подход к оцениванию корреляционных и спектральных функций основан на параметрическом представлении случайных процессов. Этот

подход широко используется при анализе временных рядов и предполагает наличие некоторой дискретной модели, описываемой соответствующим разностным уравнением.

В задачах оптимальной фильтрации процессы часто моделируются с помощью динамических систем, возбуждаемых случайными сигналами с известными характеристиками. В случае акустических сигналов такой подход также вполне обоснован, особенно если учесть вполне детерминированный механизм их формирования. В этом случае входные воздействия можно отнести скорее к регулярным сигналам с неизвестными, возможно изменяющимися, параметрами, нежели к случайным процессам. Например, акустический шум вращающейся машины может быть обусловлен в основном гармониками частоты ее вращения.

Акустические и механические резонансы природных и искусственных объектов под воздействием ветра обуславливают звуки на соответствующих резонансных частотах, интенсивность которых зависит от скорости и направления ветра. Характеристики ветра, имеющего на небольших интервалах вполне регулярный характер, в целом подвержены, конечно, значительным, возможно скачкообразным изменениям. Все это в плане борьбы с шумами и выделения полезных сигналов приводит к задаче идентификации параметров как самой модели, так и входных воздействий.

На рисунке 1.9, приведен пример структуры алгоритма формирования акустических помех $x(t)$ с применением резонансной системы (блок H) при учете воздействия параметров $u(t)$ и v . На рисунке 1.9, характеристики h и g - это операторы, значения которых зависят от условий и постановки задачи [40].

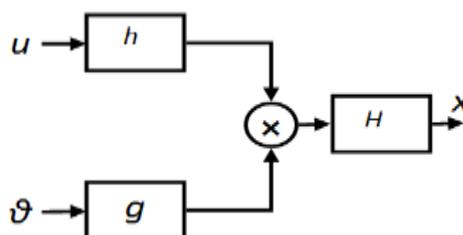


Рисунок 1.9 – Структура модели формирования акустических помех

Для систем телекоммуникаций аудиообмена, громкоговорящей связи и оповещения с адаптивным шумоподавлением и эхокомпенсацией важными для учета параметрами являются характеристики моделей акустических речевых сигналов, аддитивных и сосредоточенных помех. При этом учет вышеуказанных параметров необходим в ходе проектирования алгоритмов и программных структур обработки акустических сигналов, с целью выделения полезного речевого сигнала из искаженного шумовыми помехами и компонентами эха.

В широком смысле подавление помех может основываться как на методах адаптивной фильтрации, так и на методах компенсации. В первом случае помеха устраняется с помощью соответствующим образом спроектированных заграждающих фильтров, а во втором – посредством вычитания ее оценки из наблюдаемого сигнала. Однако в последующем, в целях упрощения, подавление будет чаще рассматриваться в узком смысле – как фильтрация помехи.

В качестве моделей сигналов используются дискретные системы, описываемые разностными уравнениями, либо вида авторегрессии – скользящего среднего [20]

$$x_k = a_1 x_{k-1} + a_2 x_{k-2} + \dots + a_p x_{k-p} + u_k - b_1 u_{k-1} - b_2 u_{k-2} - \dots - b_q u_{k-q}, \quad (1.9)$$

либо вида системы в переменных состояния

$$\begin{aligned} y(k+1) &= Ay(k) + Bu(k), \\ x(k) &= Cy(k) + Du(k), \end{aligned} \quad (1.10)$$

где $y \in R^n$ – вектор состояния, $u \in R^m$ – вектор входных воздействий и $x \in R^h$ – вектор выходов. Соответственно матрицы A , B , C и D имеют размеры: $n \times n$, $n \times m$, $h \times n$ и $h \times m$. При этом оценивание спектров сводится к оцениванию параметров модели (1.9) или (1.10).

Уравнение (1.9) является, как это следует из его названия, комбинацией двух моделей, модели авторегрессии, если все коэффициенты $b_k = 0$, и модели скользящего среднего – если $a_k = 0$.

Для системы (1.10) изображения вектора переменных состояния и вектора выхода модели записываются в виде

$$\begin{aligned} Y(z) &= (zI - A)^{-1} BU(z), \\ X(z) &= (C(zI - A)^{-1} B + D)U(z). \end{aligned} \quad (1.11)$$

Соответственно, спектры выходов описываются выражением

$$X(\omega) = \left| C(e^{j\omega}I - A)^{-1} B + D \right| \cdot |U(e^{j\omega})|. \quad (1.12)$$

Обычно модель, описываемая выражениями (1.10) – (1.12), зависит только от состояния системы, при этом $x(k) = Cy(k)$, а матрица $D = 0$.

Задача идентификации модели (1.10) может заключаться не только в оценивании матриц A , B и C , но и в оценивании состояния системы. При этом иногда используется прием, основанный на замене модели в переменных состояния эквивалентной моделью авторегрессии.

Задача идентификации во многих случаях – это задача минимизации некоторого функционала потерь, характеризующего отклонение результата аппроксимации $\hat{x}(k)$ от наблюдаемых данных $\bar{x}(k)$, $k = 1, \dots, N$:

$$\rho(\hat{\mathbf{x}}, \bar{\mathbf{x}}) \rightarrow \min.$$

Здесь $\hat{\mathbf{x}} = (\hat{x}(1) \ \hat{x}(2) \ \dots \ \hat{x}(N))^T$ и $\bar{\mathbf{x}} = (\bar{x}(1) \ \bar{x}(2) \ \dots \ \bar{x}(N))^T$ – векторы результатов аппроксимации и наблюдаемых данных. Уравнение (1.9) или (1.10) выступает при этом в качестве ограничения задачи. Возможны и другие ограничения, обусловленные, например, условиями устойчивости модели.

Указанный функционал потерь при решении некорректных задач дополняется регуляризирующим функционалом $\Omega(\hat{x}(t))$, который областью своего определения может иметь множество функций как дискретного, так и непрерывного времени [52, 83]. При этом задача идентификации принимает вид

$$\rho(\hat{\mathbf{x}}, \bar{\mathbf{x}}) + \alpha \Omega(\hat{x}(t)) \rightarrow \min.$$

Во многих случаях в качестве функционала потерь используется норма,

характеризующая расстояние между наблюдаемыми данными и значениями функции, полученной в результате идентификации:

$$\rho(\hat{\mathbf{x}}, \bar{\mathbf{x}}) = \frac{1}{2} \|\hat{\mathbf{x}} - \bar{\mathbf{x}}\|^2.$$

В случае гильбертова пространства норма представляется скалярным произведением, а функционал потерь

$$\rho(\hat{\mathbf{x}}, \bar{\mathbf{x}}) = \frac{1}{2} \langle \hat{\mathbf{x}} - \bar{\mathbf{x}}, \hat{\mathbf{x}} - \bar{\mathbf{x}} \rangle$$

является квадратичной функцией оценки сигнала по наблюдаемым данным.

Если ограничения также имеют вид квадратичных или линейных функций, то имеет место задача квадратичного программирования.

Вектор \hat{x} может быть образован выборками функции $\hat{x}(t)$ непрерывного времени, представленной в виде линейной или нелинейной регрессии. А именно, в виде функции $\hat{x}(t) = g(t, \mathbf{a})$, зависящей от вектора \mathbf{a} параметров оптимизации. В случае линейной регрессии эта функция принимает вид $\hat{x}(t) = \boldsymbol{\varphi}^T(t) \mathbf{a}$, где $\boldsymbol{\varphi}(t) = (\varphi_1(t) \ \varphi_2(t) \ \dots \ \varphi_n(t))^T$ представляет собой вектор по системе линейно независимых функций $\varphi_k(t)$, $k = 1, \dots, n$ [109, 110, 111].

Вектор \hat{x} можно при этом записать в виде

$$\hat{x} = \Phi \mathbf{a},$$

где строками матрицы Φ являются значения транспонированных векторных функций $\boldsymbol{\varphi}(k)$ в точках наблюдения:

$$\Phi = \begin{pmatrix} \boldsymbol{\varphi}^T(1) \\ \boldsymbol{\varphi}^T(2) \\ \vdots \\ \boldsymbol{\varphi}^T(N) \end{pmatrix}.$$

В случае нелинейной регрессии вектор

$$\hat{x} = \begin{pmatrix} g(1, \mathbf{a}) \\ g(2, \mathbf{a}) \\ \vdots \\ g(N, \mathbf{a}) \end{pmatrix} \equiv \mathbf{g}(\mathbf{a}).$$

Тогда задачу минимизации функции потерь в случае нелинейной регрессии можно записать в виде

$$\frac{1}{2} \|\mathbf{g}(\mathbf{a}) - \bar{\mathbf{x}}\|^2 \rightarrow \min, \quad (1.13)$$

а в случае линейной регрессии – в виде

$$\frac{1}{2} \|\Phi \mathbf{a} - \bar{\mathbf{x}}\|^2 \rightarrow \min. \quad (1.14)$$

Как известно [63], математическое ожидание функции (1.13) или (1.14) достигает своего наименьшего значения, если вектор $\mathbf{g}(\mathbf{a})$ или $\Phi \mathbf{a}$ совпадает с условным математическим ожиданием вектора \mathbf{x} , $E\{\mathbf{x}|\bar{\mathbf{x}}\}$, рассматриваемым как функция наблюдаемых данных $\bar{\mathbf{x}}$. При этом величина потерь совпадает с условной дисперсией, если только отбросить коэффициент 1/2.

Полученный в результате минимизации функции потерь вектор коэффициентов регрессии $\hat{\mathbf{a}}$ определяет оптимальное решение $\hat{x}^*(t) = \hat{\phi}^T(t)\hat{\mathbf{a}}$ как функцию непрерывного времени. Это позволяет при необходимости наложить на ее поведение между выборками данных дополнительные ограничения, с помощью, например, регуляризирующего функционала.

Существуют и другие способы оценки параметров модели, например, способ, основанный на согласовании корреляционных функций. В частности, такого рода методы используются при сжатии речи и вычислении спектров. Используются методы, основанные на собственных значениях и сингулярных разложениях ковариационных матриц, имеющих тёплицеву структуру.

Достижимая точность приближения наблюдаемых данных функцией регрессии зависит в значительной степени от размеров области ее определения. С увеличением размеров области, в пределах которой наблюдаемые данные не стремятся к нулю, точность снижается. Устранить этот недостаток можно, воспользовавшись методом локальной аппроксимации. Приближение наблюдаемых данных в этом случае обеспечивается последовательностью функций регрессии, каждая из которых задана на своем конечном интервале. Одновременно это позволяет

аппроксимировать нестационарные сигналы и системы.

Вопрос сопряжения отдельных функций регрессии можно решить, если дополнить ограничения, представленные в задаче минимизации, условиями согласования значений этих функций и, возможно, значений их производных в узлах сопряжения.

Таким образом, проблема оптимальной обработки сигналов в телекоммуникационных системах передачи речи обусловливается недостаточностью достоверной априорной информации о статистических характеристиках наблюдаемых сигналов. Задача оценивания параметров моделей решается различными методами в зависимости от ограничений на статистические характеристики наблюдаемых сигналов. Наиболее перспективными методами для данных условий можно считать метод минимизации функции потерь, метод максимального правдоподобия, метод согласования корреляционных функций.

1.5. Анализ акустических речевых сигналов на основе оценок плотности вероятностей

1.5.1. Непараметрические оценки расстояний между распределениями

Сравнение распределений предполагает введение соответствующей числовой меры, характеризующей статистическое различие, по меньшей мере, двух случайных величин или результатов двух независимых серий наблюдений над одной величиной [55]. Такая мера, в частности, необходима для интерпретации и сравнительного анализа результатов аппроксимации плотности вероятностей, полученных с помощью различных алгоритмов и по отличающимся данным. Подобная задача возникает и в других областях математической статистики, при анализе временных рядов и обнаружении изменений свойств систем [6].

В качестве указанной числовой меры, характеризующей различие распределений, возможно использовать непараметрические оценки – функционалы от точных и эмпирических функций распределения. Наиболее

известными являются статистики Смирнова, Колмогорова и Кульбака. Статистика, полученная по методике Смирнова, позволяющая задать «расстояние» между истинным распределением P и полученным опытным путем эмпирическим распределением P_n^* , определена выражением

$$\omega_n^2 \equiv d(P, P_n^*) = n \int (F(x) - F_n^*(x))^2 dx. \quad (1.16)$$

в котором, параметр n – является объемом выборки, $F(x)$, $F_n^*(x)$ – соответственно точные и эмпирические функции плотности распределения вероятностей, которые представляются выражением вида

$$F_n^*(x) = \frac{1}{n} \sum_{l=1}^n I_{(-\infty, x]}(x_l) \quad (1.17)$$

в котором $I_B(x) = \begin{cases} 1, & x \in B \\ 0, & x \notin B \end{cases}$.

Другим вариантом вычисления «расстояния» между истинной и эмпирической функциями распределения плотности вероятностей акустических речевых сигналов является подход, основанный на методе Колмогорова, статистика по которому вычисляется по выражению

$$d_K(P, P_n^*) = \sqrt{n} \sup_x |F(x) - F_n^*(x)| \quad (1.18)$$

С учетом рассмотренных выше статистик (1.16) или (1.18) и функций распределения плотности вероятностей возможно проведение анализа вычислительной сложности и связанной с ним точности представления данных при обработке для обеспечения требуемого отношения сигнал / помеха. Параметром выбора решения, определяющего соответствие эмпирических данных выбранному распределению может служить неравенство вида

$$n \int (F(x) - F_n^*(x))^2 dx \leq c_\varepsilon, \quad (1.19)$$

либо вида

$$\sqrt{n} \sup_x |F(x) - F_n^*(x)| \leq c_\varepsilon. \quad (1.20)$$

Выбор уровня пороговой обработки c_ε может быть основан на

теоретической оценке условия

$$\Pr(d(P, P_n^*) > c) = \varepsilon \quad (1.21)$$

В выражении (1.21) $\Pr(A)$ определяет вероятность возникновения события A . При этом нахождение параметров истинного распределения рассматриваемой статистики $d(P, P_n^*)$, зачастую, сопряжено со значительными вычислительными затратами, и требует применения подходов, основанных на принципах асимптотических, либо непараметрических методов оценивания. Малые размеры выборки не позволяют при применении асимптотических методов достичь точных результатов.

Подобным образом осуществляется сравнение различных эмпирических статистик, таких как статистика Смирнова и Колмогорова [6]. По данным статистикам «расстояния» между функциями распределения $F_n^*(x)$ а также $G_n^*(x)$ могут быть получены в соответствии с формулами вида

$$d(P_{1,n}^*, P_{2,n}^*) = n \int (F_n^*(x) - G_n^*(x))^2 dx \quad (1.22)$$

$$d_K(P_{1,n}^*, P_{2,n}^*) = \sqrt{n} \sup_x |F_n^*(x) - G_n^*(x)|. \quad (1.23)$$

Вышеуказанные равенства позволяют осуществить поиск характерных изменений в акустических речевых сигналах и проектировать соответствующие фильтры.

Список возможных к применению статистик можно расширить статистиками, определяющими «расстояния» Леви и Прохорова. Вышеуказанным «расстоянием» Леви между функциями плотности распределения вероятностей $F(x)$ и $G(x)$ называется величина, заданная выражением

$$d_L(F, G) = \inf \{ \varepsilon \mid F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon, \forall x \} \quad (1.24)$$

Для расчета «расстояния» Прохорова можно применять выражение

$$d_{II}(F, G) = \inf \{ \varepsilon \mid F\{A\} \leq G\{A^\delta\} + \varepsilon, \forall A \subset B \}, \quad (1.25)$$

в котором $A^\delta = \left\{ x \in \Omega \mid \inf_{y \in A} d(x, y) \leq \delta \right\}$ определяет некоторую окрестность множества A [40].

Отличительной особенностью вышеуказанных «расстояний» Прохорова и Леви может служить тот факт, что их можно считать некоторыми метриками функций распределения плотности вероятностей.

В качестве дополнения, можно отметить метрику Липшица, которую принято считать ограниченной

$$d_{BL}(F, G) = \sup_{\varphi} \left| \int_X \varphi(x) dF(x) - \int_X \varphi(x) dG(x) \right|, \quad (1.26)$$

в которой точная верхняя грань формируется по всем функциям $\varphi(x)$, которые удовлетворяют ограничениям Липшица $|\varphi(x) - \varphi(y)| \leq d(x, y)$ [39].

Преимуществом метрики Липшица служит потенциальная возможность сокращения чувствительности производимой оценки к имеющим место нежелательным «выбросам» и «всплескам» наблюдаемых данных. Данные особенности метода Липшица обусловили появление метода робастной статистики [49, 97, 99], обеспечивающего формирование ряда критериев, в которых обобщаются подходы Смирнова и Колмогорова.

Формирование возможности нахождения особенностей и изменения свойств акустических речевых сигналов с применением одной из рассмотренных статистик, базируется на анализе и сравнении эмпирических функций плотности распределений вида

$$F_{n,m}^*(x) = \frac{1}{n} \sum_{l=(m-1)n+1}^{mn} I_{(-\infty, x]}(x_l), \quad G_{n,m}^*(x) = \frac{1}{n} \sum_{l=(m-1)n+q+1}^{mn+q} I_{(-\infty, x]}(x_l), \quad m=1, 2, \dots \quad (1.27)$$

В (1.27) m – номер окна данных, n – размер окна и q – величина смещения окна данных статистики G_n^* относительно F_n^* .

1.5.2. Проблема некорректности оценок плотности вероятностей

Потребность в преодолении некорректного характера задачи возникает не только при восстановлении плотности вероятности случайного сигнала [4,

17, 25], но и при анализе временных рядов, идентификации и оптимизации систем [38, 70, 98]. Принципы решения некорректных задач имеют независимый от области применения характер [16, 35, 52, 83, 84]. Здесь задача рассматривается в применении к оцениванию плотности вероятностей, что предполагает, что случайная величина характеризуется непрерывным распределением [17].

В рамках одного из методов оценку плотности вероятностей можно найти из интегрального уравнения, по которому могут быть установлены взаимосвязи между функциями распределения и плотностью вероятностей акустических речевых сигналов. В случае использования одномерных функций плотности вероятностей, соотношение может принять вид

$$F(x) = \int_{-\infty}^x f(v)dv = \int_{-\infty}^{\infty} I(x-v)f(v)dv. \quad (1.28)$$

В (1.28) $I(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ - это функция Хевисайда, F - определяет

значения функции распределения, а f , в свою очередь является плотностью вероятностей. Имеется возможность замены истинной функции распределения аппроксимирующей эмпирической функцией распределения $F_n(x)$.

$$F_n^*(x) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(x_i). \quad (1.29)$$

Отсюда следует интегральное уравнение относительно искомой плотности вероятностей:

$$\int_{-\infty}^{\infty} I(x-v)f(v)dv = F_n(x). \quad (1.30)$$

Однако известно, что задача решения подобных уравнений является некорректной, то есть малые отклонения $F_n(x)$ от точной функции распределения могут приводить к большим ошибкам [16, 35]. В рассматриваемом случае ошибки возникают по двум причинам [107]. Во-первых, из-за ограниченного набора данных и, во-вторых, из-за особенностей

выбранного алгоритма их аппроксимации.

Задача решения уравнения (1.30), которое в операторной форме имеет вид $A \circ \hat{f}_n = F_n$, относится к классу обратных задач. Используя обратный оператор A^{-1} , получаем

$$\hat{f}_n = A^{-1} \circ F_n \quad (1.31)$$

При этом задача считается корректной, если выполняются условия:

- 1) решение \hat{f}_n может существовать для любой функции $F_n \in \Omega_F$,
- 2) решение \hat{f}_n является единственно возможным для любой функции $F_n \in \Omega_F$,
- 3) решение $\hat{f}_n \in \Omega_f$ обеспечивает непрерывную зависимость от функции $F_n \in \Omega_F$, при этом, для любых сходящихся последовательностей функций $F_n \xrightarrow{\Omega_F} F \in \Omega_F$, полученная последовательность формируемых решений $\hat{f}_n \xrightarrow{\Omega_f} f \in \Omega_f$, в которых f служит решением задачи $A \circ f = F$.

Задача фактически заключается в нахождении производной функции $F_n(x)$. Отсюда следует, что одной из возможных причин некорректности является наличие высокочастотных ошибок в эмпирической функции распределения [35, 107].

Подход к получению решения некорректных задач в соответствии с методом Тихонова основывается на уменьшении множества потенциально возможных решений Ω_f до вложенного множества $\Omega'_f \subset \Omega_f$, при котором действие оператора A обеспечивает формирование подмножества $\Omega'_F = A \circ \Omega'_f \subset \Omega_F$, при котором на нем имеет место оператор A^{-1} обратный по отношению к оператору A .

Одним из наиболее распространенных подходов к реализации вышеуказанного метода, является принцип регуляризации, при котором дополнительно вводится функционал регуляризации $G(f)$. Данный функционал обязательно соответствует ряду условий [40]:

1. Найденные решения полученного операторного уравнения должны быть включены в область определения рассматриваемого функционала;

2. Полученный функционал $G(f)$, не может принимать положительных значений или быть равен нулю;

3. Множества, включенные в набор $A_G = \{f \mid G(f) \leq c\}$ являются компактными.

При этом чтобы применить метод регуляризации, требуется вычислить некоторую последовательность решений f_α , которая обеспечивает возможность минимизации функционала

$$Q(f, F_n) = \rho^2(A \circ f, F_n) + \alpha G(f), \quad (1.32)$$

с учетом разнообразных значений $\alpha > 0$, которые стремятся к нулевому значению ($\alpha \rightarrow 0$). В выражении (1.32) слагаемое $\rho(A \circ f, F_n)$ представляет собой меру «расстояния» между рассматриваемыми функциями $A \circ f$ и F_n .

С учетом описанных условий, полученное решение может быть представлено пределом следующего вида [40]

$$\hat{f} = \lim_{\alpha \rightarrow 0} f_\alpha. \quad (1.33)$$

В развернутой форме функционал (1.23) для задачи (1.21) имеет вид

$$Q(f, F_n) = \rho^2 \left(\int_{-\infty}^{\infty} I(x-v) f(v) dv, F_n(x) \right) + \alpha G(f(x)) \quad (1.34)$$

В качестве меры расстояния между распределениями можно, например, взять функцию

$$\rho(\hat{F}(x), F_n(x)) = \sup_x |\hat{F}(x) - F_n(x)| \quad (1.35)$$

или квадратичную функцию

$$\rho^2(\hat{F}(x), F_n(x)) = \int_{-\infty}^{\infty} w(x) \left(\int_{-\infty}^{\infty} I(x-v) f(v) dv - F_n(x) \right)^2 dx \quad (1.36)$$

где $w(x)$ – неотрицательная весовая функция.

Регуляризирующий функционал можно задать выражением вида

$$G(f) = \int f^2(x)dx \quad (1.37)$$

Чтобы оценить плотность вероятностей речевых сигналов, можно применить не только выражение (1.30), но и некоторые его модификации, которые могут быть получены с использованием линейного оператора B , являющегося невырожденным. В этом случае равенство (1.30) может быть записано в виде:

$$B \circ \int_{-\infty}^{\infty} I(x-v)f(v)dv = B \circ F_n(x). \quad (1.38)$$

Применение оператора B предоставляет дополнительные возможности по регуляризации задачи. При этом становится понятным, что при условии подстановки в правую часть выражения (1.38) точной плотности $F(x)$, определенной выражением $F_n(x) = F(x)$, то корни уравнений (1.30) а также (1.38) совпадут, в независимости от формы оператора B , при условии его невырожденности.

Очевидно, что с переходом к уравнению (1.38) функционал (1.34) задачи (1.30) преобразуется к виду

$$Q(p, F_n) = \rho^2 \left(B \circ \int_{-\infty}^{\infty} I(x-v)f(v)dv, B \circ F_n(x) \right) + \alpha G(f(x)) \quad (1.39)$$

где $\rho(B \circ \hat{F}(x), B \circ F_n(x)) = \sup_x |B \circ \hat{F}(x) - B \circ F_n(x)|$

или $\rho^2(B \circ \hat{F}(x), B \circ F_n(x)) = \int_{-\infty}^{\infty} w(x) \left(B \circ \int_{-\infty}^{\infty} I(x-v)f(v)dv - B \circ F_n(x) \right)^2 dx$.

1.6. Анализ существующих систем и решений громкоговорящей связи и оповещения.

Разработкой систем громкоговорящей связи и оповещения для объектов промышленности занимается много производителей. Известны решения от BOSCH¹, ЛайтБау², Коммуникации³, Максиком⁴. В рамках разрабатываемых систем необходимо решать задачи повышения разборчивости речи в системах трансляции. Этим разработкам посвящен ряд исследований, например

«Устройства и способы для обработки аудио сигнала с целью повышения разборчивости речи, используя функцию выделения нужных характеристик» по патенту RU 2507608 от 3 августа 2009 г. (Фраунхофер-Гезелльшафт цур Фердерунг дер ангевандтен Форшунг Е.Ф. (DE), МПК G10L 21/02 (2013.01), опубликовано 20.02.2014, Бюл. №5). Устройства обработки аудиосигнала для получения контрольной информации для фильтра повышения разборчивости речи.

«Улучшение разборчивости речи в мобильном коммуникационном устройстве путем управления работой вибратора в зависимости от фонового шума» по патенту RU 2411595 от 1 августа 2006 г. (Конинклейке филипс электроникс Н.В. (NL), МПК G10L 21/02 (2006.01) H04M 1/60 (2006.01), опубликовано 10.02.2011, Бюл. №4), за счет анализа фонового шума и формирования управляющего сигнала на вибратор, для улучшения эффективности передачи.

«Способ повышения разборчивости речи» по патенту RU 2676022 от 13 июля 2016 г. Общество с ограниченной ответственностью "Речевая аппаратура "Унитон" (RU), МПК G10L 21/02 (2013.01), G10L 21/0364

¹ Проектирование систем оповещения сайт] URL = <http://bosch-audio-system.ru/> (дата обращения: 16.10.2020). - Текст: электронный

² Громкоговорящая связь | Инжиниринговая компания LightBau | Москва [сайт] URL = <https://www.lightbau.com/> (дата обращения: 16.10.2020). - Текст: электронный.

³ Communications.su - Системная интеграция связи для промышленности [сайт] URL = <http://www.communications.su/ru/> (дата обращения: 16.10.2020). - Текст: электронный

⁴ Официальный сайт производителя АТС в России: Мультиком [сайт] URL = <https://www.multicom.ru/> (дата обращения: 16.10.2020). - Текст: электронный.

(2013.01), СПК G10L 21/02 (2006.01), G10L 21/0364 (2006.01), опубликовано 25.12.2018, Бюл. №36), за счет использования полосовых фильтров, настроенных по спектральным зонам, несущим основные фонетические признаки звуков речи.

«Способ измерения разборчивости речи» по патенту RU 2620569 от 17 мая 2016 г. Иванов Николай Александрович (RU), Стародубцев Петр Юрьевич (RU), МПК G10L 15/00 (2013.01), H04R 29/00 (2006.01), опубликовано 26.мая 2017, Бюл. №15), за счет применения пространственно-распределенного преобразователя акустического сигнала.

«Способ повышения разборчивости и информативности звуковых сигналов в шумовой обстановке» по патенту RU 2589298 от 29.12.2014 (Бредихин Александр Юрьевич (RU), МПК G10L 19/02 (2013.01), G10L 21/0232 (2013.01), G10L 21/028 (2013.01), опубликовано 10.07.2016, Бюл. №19), за счет деления на фреймы и адаптивной компенсации шумовых составляющих.

«Корреляционный адаптивный инвариантный эхокомпенсатор» по патенту RU 168793 от 16 июня 2016 г. (Федеральное государственное бюджетное образовательное учреждение высшего образования "Сибирский государственный университет телекоммуникаций и информатики" (СибГУТИ) (RU), МПК H04B 3/20 (2006.01), опубликовано 21.02.2017, Бюл. №6), за счет вычисления величины корреляции суммы принимаемых сигналов, помехи и экосигналов с передаваемыми сигналами.

1.7. Выводы и постановка задачи исследований.

На многофункциональных объектах, возникает проблема достоверного и оперативного доведения информации до абонентов системы с применением систем громкоговорящей связи и оповещения. Для этого необходимо повышать эффективность алгоритмов и методов обработки и передачи информации в телекоммуникационных системах обмена акустическими речевыми сигналами, функционирующими при воздействии интенсивных

акустических помех и экосигналов. Требование повышения качества при обработке и передаче речевой информации [106] обуславливается недостаточностью достоверной априорной информации о статистических характеристиках наблюдаемых сигналов и соответственно недостаточной точностью моделирования [26]. Задачи оценивания параметров моделей решаются различными моделями в зависимости от ограничений на статистические характеристики наблюдаемых сигналов [6, 80]. Наиболее перспективными методами для данных условий можно считать, это метод минимизации функции потерь, метод максимального правдоподобия, метод локальной аппроксимации.

Чтобы проектировать современные, эффективные подходы к формированию алгоритмов обработки акустических речевых сигналов [33, 54], требуется первоначально получить данные об огибающей спектральных функций, нахождение которых обусловлено необходимостью подбора оптимальных частотных полос, а также формы амплитудно-частотной характеристики в передающем канале системы обмена речевыми сообщениями [40]. Это приводит к необходимости формирования усовершенствованных алгоритмов ДПФ, обладающих значительно меньшей вычислительной сложностью, а также формирования более эффективных алгоритмов обработки речевых сигналов [39].

Таким образом, создание новых моделей, новых программных структур анализа речи является своевременной и значимой задачей, в связи с чем основной целью диссертационной работы служит повышение эффективности телекоммуникационных систем обмена речевыми сообщениями за счет разработки новых алгоритмов обработки сигналов, обеспечивающих увеличение точности при анализе речевых сообщений и существенное повышение эффективности при передаче информации речевыми сообщениями на многофункциональных объектах.

Для достижения поставленной цели необходимо решение следующих задач.

1. Разработка статических моделей представления речевых сигналов с минимизированной погрешностью аппроксимации.

2. Разработка методов оценивания задержки акустических сигналов.

3. Формирование эффективных алгоритмов обработки речевых сигналов, а также соответствующего программного обеспечения, существенно повышающих ОСШ (отношения сигнал/шум) при их передаче, которое влечет за собой обеспечение установленной нормативными документами достаточной слоговой разборчивости при обмене речевой информацией в телекоммуникационных системах связи и оповещения.

ГЛАВА 2. МЕТОДЫ ОЦЕНИВАНИЯ ПАРАМЕТРОВ МОДЕЛЕЙ АКУСТИЧЕСКИХ СИГНАЛОВ.

2.1. Исследование влияния отношения акустический сигнал/шумовая помеха на слоговую разборчивость

Качественная передача речевой информации на многофункциональных объектах является значимой проблемой, решение которой обеспечит достоверное доведение оперативно-командной информации до абонентов системы. Поэтому повышение эффективности функционирования таких систем, как системы оповещения, системы громкоговорящей связи, системы обмена речевыми аудиосообщениями является актуальной проблемой и требует проведения дальнейших исследований. Наиболее значимым параметром, который определяет эффективность функционирования систем передачи информации речевыми сообщениями, служит слоговая разборчивость, обозначаемая S %. Слоговая разборчивость является мерой оценки качества акустических речевых сигналов по шкале средней субъективной оценки восприятия данных, называемой шкалой MOS (Mean Opinion Score) [123, 125]. В стандарте [120, 121] определено, что система передачи речевых сообщений может считаться эффективной, при условии, что вся передаваемая речевая информация будет восприниматься принимающими абонентами полностью и без каких-либо существенных затруднений. В обратном случае эффективность функционирования объекта массового обслуживания или многофункционального объекта системы существенно снижается. Стандарт определяет минимальный уровень слоговой разборчивости значением $S \geq 93$ % для обеспечения полного восприятия передаваемых данных [120, 121, 122], что эквивалентно оценке в 3,9 баллов по шкале MOS [124, 125, 126].

Исследование зависимости уровня слоговой разборчивости от разнообразных факторов, оказывающих влияние на функционирование телекоммуникационных систем передачи речевых сообщений, проводилось в ряде научных работ [10, 11, 29, 66, 71, 73, 74, 78, 81]. Однако,

в рассмотренных вышеуказанных источниках, результаты проведенного анализа и исследований по зависимости слоговой разборчивости речи от ОСШ (отношение акустический сигнал/шумовая помеха) при передаче речевых сообщений в системах связи представлены в не достаточно полной степени, что обосновывает значимость проводимых исследований. В работе использовались разнообразные подходы и методы, такие как метод оценивания формантной разборчивости R (в англоязычной литературе называемой интегральным индексом артикуляции), метод оценивания значения слоговой разборчивости в соответствии с параметром разборчивости формант речи, метод вычисления коэффициентов восприятия формантных составляющих при разных значениях отношений сигнал/шум [104].

Первоначально были исследованы вопросы восприятия формантных составляющих p_i . На графике (рисунок 2.1) представлена функция зависимости вычисленных коэффициентов формантных составляющих речевых сигналов от значения параметра интенсивности формант Q_i , которые вычислялись в каждой частотной полосе i на заданной среднегеометрической частоте $f_{cp,i}$.

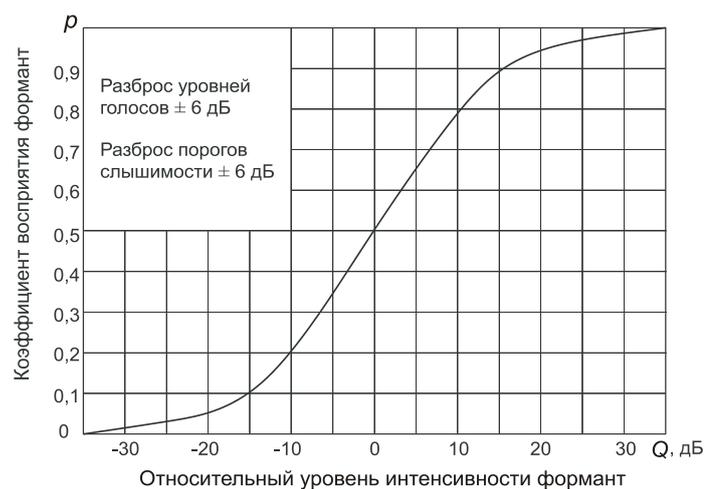


Рисунок 2.1 – График функциональной зависимости p_i от значений Q_i

При этом параметр формантной интенсивности Q_i , определялся в соответствии с выражением

$$Q_i = q_i - \Delta A_i, \quad (2.1)$$

в котором q_i – ОСШ в i -той формантной частотной полосе.

В том случае, если количество октавных полос $N=5$, то становится возможным вычисление параметров пределов $f_{н.i}$ и $f_{в.i}$ по частотным полосам, а также определение $f_{ср.i}$ которые отражают среднегеометрические значения частоты в i -той частотной полосе, по которым вычисляются значения ΔA_i , представляющие формантные параметры.

Формантные параметры ΔA_i , позволяют с помощью (2.1) найти значения относительных уровней интенсивности формант Q_i , находящихся в зависимости от ОСШ q_i .

По графику, представленному на рисунке 2.1, возможно определение коэффициентов восприятия формант p_i в соответствии со значениями Q_i для всех частотных полос i , при изменении значений ОСШ.

Чтобы вычислить формантную разборчивость R_i , требуется найти весовые коэффициенты k_i , отображающие вероятность нахождения формант речевых сообщений в частотных полосах i в соответствии с выражением

$$k_i = k(f_{в.i}) - k(f_{н.i}),$$

в котором $k(f_{в.i})$ и $k(f_{н.i})$ – весовые коэффициенты верхней ($f_{в.i}$) и нижней ($f_{н.i}$) граничных частот в i -той полосе частотного спектра акустических речевых сигналов.

Параметры $k(f_{в.i})$, $k(f_{н.i})$ находятся по графику, который изображен на рисунке 2.2.

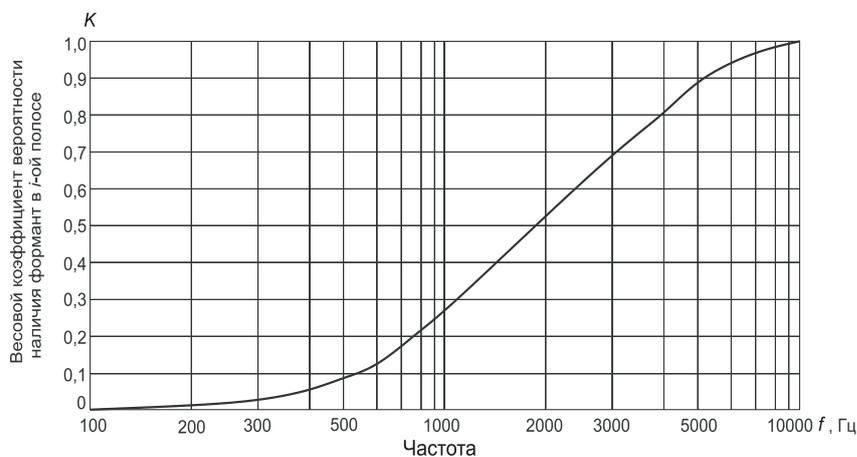


Рисунок 2.2 – График зависимости весового коэффициента вероятности наличия формант в различных частотных полосах

По полученным значениям весовых коэффициентов вероятности нахождения формант речевых сообщений в частотных полосах i , определяются значения параметров разборчивости формант R_i в зависимости от ОСШ q_i , после чего параметры R_i были сведены в таблицу 2.1.

Таблица 2.1.

Значения разборчивости формант R_i , при различных значениях отношения с/ш

$R_i = p_i \cdot k_i$	$R_1 = p_1 \cdot k_1$	$R_2 = p_2 \cdot k_2$	$R_3 = p_3 \cdot k_3$	$R_4 = p_4 \cdot k_4$	$R_5 = p_5 \cdot k_5$
$q_i = 0$ дБ	$0.07 \cdot 0.03 = 0.002$	$0.180 \cdot 0.14 = 0.025$	$0.22 \cdot 0.18 = 0.039$	$0.29 \cdot 0.37 = 0.107$	$0.34 \cdot 0.29 = 0.098$
$q_i = 3$ дБ	$0.11 \cdot 0.03 = 0.003$	$0.021 \cdot 0.14 = 0.029$	$0.31 \cdot 0.18 = 0.055$	$0.38 \cdot 0.37 = 0.140$	$0.41 \cdot 0.29 = 0.118$
$q_i = 6$ дБ	$0.20 \cdot 0.03 = 0.006$	$0.300 \cdot 0.14 = 0.042$	$0.40 \cdot 0.18 = 0.072$	$0.48 \cdot 0.37 = 0.177$	$0.51 \cdot 0.29 = 0.147$
$q_i = 10$ дБ	$0.24 \cdot 0.03 = 0.007$	$0.410 \cdot 0.14 = 0.057$	$0.51 \cdot 0.18 = 0.091$	$0.53 \cdot 0.37 = 0.196$	$0.62 \cdot 0.29 = 0.179$
$q_i = 20$ дБ	$0.47 \cdot 0.03 = 0.014$	$0.600 \cdot 0.14 = 0.084$	$0.79 \cdot 0.18 = 0.091$	$0.81 \cdot 0.37 = 0.299$	$0.88 \cdot 0.29 = 0.255$
$q_i = 30$ дБ	$0.81 \cdot 0.03 = 0.024$	$0.900 \cdot 0.14 = 0.126$	$0.94 \cdot 0.18 = 0.169$	$0.96 \cdot 0.37 = 0.355$	$0.98 \cdot 0.29 = 0.284$

В результате проведённых исследований значений разборчивости формант R_i , представленных в таблице 1, возможно нахождение зависимости уровня формантной разборчивости R от ОСШ, позволяющее получить зависимость слоговой разборчивости при передаче речевых сообщений от ОСШ. Полученные результаты соответствия слоговой разборчивости от ОСШ приведены в таблице 2.2.

Таблица 2.2.

Слоговая и формантная разборчивость для различных значений ОСШ

Формантная разборчивость, R	0.273	0.348	0.4455	0.5323	0.7952	0.9589
Слоговая разборчивость, S	25%	35%	53%	65%	93%	98.5%
ОСШ, q_i	0 дБ	3 дБ	6 дБ	10 дБ	20 дБ	30 дБ

Графическое изображение функциональной зависимости S от ОСШ приведено на рисунке 2.3.

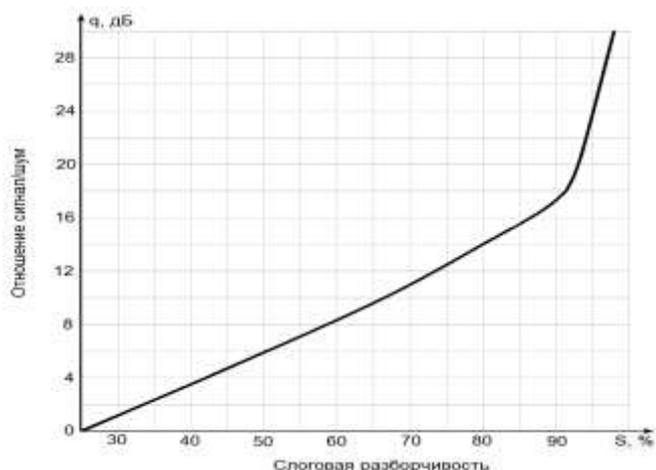


Рисунок 2.3 – Графическое соотношение между S и ОСШ (q)

По графику (рисунок 2.3) можно определить, что для обеспечения требуемой, для полного восприятия информации абонентами системы телекоммуникаций аудиообмена, слоговой разборчивости значением $S \geq 93\%$, необходимо достижение ОСШ уровнем не менее 20 дБ. В результате исследований получена значимая, для практического применения, зависимость слоговой разборчивости от ОСШ [104].

Полученная сеточная функция на рисунке 2.3, определяет второе основное положение, выносимое на защиту.

2.2. Применение метода максимального правдоподобия для исследования параметров модели акустических сигналов

Поставим задачу определения параметров в наблюдаемом сигнале $x(t)$, являющимся аддитивной суммой оцениваемого сигнала $s(t)$ и акустической помехи $\eta(t)$, которая также считается комплексной функцией. В дискретной форме этот сигнал имеет вид

$$x(k) = s(k) + \eta(k) = \sum_{n=1}^p a_n e^{j\left(n \frac{\omega_1}{f_0} k + \varphi_n\right)} + \eta(k). \quad (2.2)$$

Здесь f_0 – частота дискретизации.

Задачу определения параметров функции (2.2) можно решить также методом максимального правдоподобия [3, 7, 8, 9, 15, 86], применение

которого осложнено недостаточной надежностью априорной информации о распределениях помех [26, 51, 59]. Метод максимального правдоподобия, как известно, в случае независимых одинаково распределенных гауссовых величин эквивалентен методу наименьших квадратов. В этом случае параметры функции можно, в принципе, найти методами нелинейного программирования [55, 63, 65].

$$\{\hat{\omega}_1, \hat{a}, \hat{\varphi}\} = \arg \min_{\{\omega_1, a, \varphi\}} \|s - x\|^2.$$

Если ввести векторы выборок выделяемого $\mathbf{s} = (s(1) \ s(2) \ \dots \ s(N))^T$ и наблюдаемого $\mathbf{x} = (x(1) \ x(2) \ \dots \ x(N))^T$ сигнала, векторы $\mathbf{a} = (a_1, \dots, a_p)^T$ и $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)^T$, то норму можно записать в виде

$$\|s - x\|^2 = \langle s - x, s - x \rangle = \sum_{k=1}^N \left| \sum_{n=1}^p a_n e^{j \left(\frac{\omega_1}{f_0} k + \varphi_n \right)} - x(k) \right|^2. \quad (2.3)$$

Если, к тому же, помеха является коррелированной, с корреляционной функцией $R(t)$ и корреляционной матрицей

$$\mathbf{R}(k, n) = \begin{pmatrix} R(1, 1) & R(1, 2) & \dots & R(1, K) \\ R(2, 1) & R(2, 2) & \dots & R(2, K) \\ \vdots & \vdots & \dots & \vdots \\ R(K, 1) & R(K, 2) & \dots & R(K, K) \end{pmatrix}, \quad (2.4)$$

то квадрат нормы в задаче минимизации принимает вид

$$\|s - x\|^2 = \langle s - x, s - x \rangle = (\mathbf{s} - \mathbf{x})^T \mathbf{R}^{-1} (\mathbf{s} - \mathbf{x})^*.$$

Если принять, что оцениваемый сигнал имеет вид многочлена $\mathbf{a}^T \boldsymbol{\varphi}(t)$ по системе линейно независимых функций, то это приводит к параметрической задаче максимизации функции правдоподобия $P(y_a^b | \mathbf{a})$, решение которой по наблюдаемым на интервале $[a, b]$ данным y_a^b записывается в виде:

$$\hat{\mathbf{a}} = \arg \sup_{\mathbf{a}} P(y_a^b | \mathbf{a}). \quad (2.5)$$

Метод максимального правдоподобия может быть использован при определенных условиях, если наблюдаемые данные представляют собой последовательность независимых случайных величин с плотностью вероятности $p(y_k|\mathbf{a})$. При этом функция правдоподобия имеет вид $P(y_a^b|\mathbf{a}) = \prod_{k=1}^n p(y_k|\mathbf{a})$. Аналогично, если известны одномерная плотность вероятности $p(y_k|\mathbf{a})$ и условная плотность вероятности $p(y_{k+1}|y_k, \mathbf{a})$, то функция правдоподобия описывается выражением [95]

$$P(y_a^b|\mathbf{a}) = p(y_1|\mathbf{a}) \prod_{k=1}^{n-1} p(y_{k+1}|y_k, \mathbf{a}).$$

В связи с рассмотренными выше методами параметрической оптимизации и локальной аппроксимации, следует отметить, что по своему содержанию они практически эквивалентны проекционным методам решения операторных уравнений [64], проекционным и интерполяционным методам анализа и расчета систем [24, 25, 54, 58].

2.2.1. Параметрические методы оценивания плотности вероятностей акустических сигналов

Методы как параметрического, так и непараметрического оценивания плотности вероятностей [57, 92, 93, 115] заключаются в минимизации ее отклонения от некоторой эталонной функции. Если обозначить это отклонение функционалом

$$J(\hat{f}, f, \theta) = \int_{-\infty}^{\infty} Q(\hat{f}(x, \theta), f(x)) \mu(dx) \equiv \int_{-\infty}^{\infty} Q(x, \theta) \mu(dx), \quad (2.6)$$

то задача принимает вид

$$J(\hat{f}, f, \theta) \rightarrow \min_{\theta} \text{ или } \theta^* = \arg \min_{\theta} J(\hat{f}, f, \theta).$$

В параметрическом случае – это задача безусловной минимизации. В непараметрическом же случае на решение должны быть наложены ограничения вида (2.7)

$$\hat{f}(x, \theta^*) \geq 0 \text{ и } \int_{-\infty}^{\infty} \hat{f}(x, \theta^*) dx = 1. \quad (2.7)$$

Это приводит к задаче условной минимизации.

Если за расстояние между распределениями принять расстояние Смирнова, то выражение (2.6) принимает вид (2.8)

$$J(\hat{f}, f, \theta) = \int_{-\infty}^{\infty} (\hat{F}(x, \theta) - F(x))^2 \mu(dx), \quad (2.8)$$

в котором $\hat{F}(x, \theta)$ и $F(x)$ – функции распределений, соответствующие \hat{f} и f .

Эмпирический функционал по значениям функций распределения в N точках представляется при этом суммой

$$J(\hat{f}, f, \theta) = N \sum_{l=1}^N (\hat{F}(x_l, \theta) - F(x_l))^2. \quad (2.9)$$

При параметрическом оценивании широкое применение находят методы максимального правдоподобия, байесовские и минимаксные методы [15, 17, 54, 117], а при непараметрическом, помимо метода наименьших квадратов, используются методы моментов, разложения по системам ортогональных функций и аппроксимации Паде [20, 24, 48, 50].

Оценка максимального правдоподобия является решением задачи

$$\hat{\theta} = \arg \max_{\theta} \int \ln f_{\theta}(x) dF(x) \cong \arg \max_{\theta} \frac{1}{N} \sum_{l=1}^N \ln f_{\theta}(x_l). \quad (2.10)$$

Поскольку решение этой задачи в общем случае не является единственным, то не единственной является и полученная таким способом оценка. Оценку максимального правдоподобия можно при этом отнести к оценкам минимального риска, получаемым из уравнения вида (2.6)

$$R(\theta) = \int Q(x, \theta) dF(x) \cong \frac{1}{N} \sum_{l=1}^n Q(x_l, \theta) = R_n(\theta), \quad (2.11)$$

если принять что функция

$$Q(x, \theta) = -\ln f_{\theta}(x).$$

Выражение (2.10) можно представить как оценку подстановки, основанной на эмпирической плотности вероятностей. Оно имеет

определенную связь с функцией расстояния Кульбака

$$d(f, f_{\hat{\theta}}) = -\int \ln \left(\frac{f_{\hat{\theta}}(x)}{f(x)} \right) f(x) dx.$$

2.3. Постановка задачи аппроксимации плотности вероятностей акустических сигналов

Плотность вероятностей случайной величины может быть выражена в виде обобщенного многочлена

$$P(x, a, \theta) = \sum_{k=1}^m a_k \varphi_k(x, \theta) = a^T \varphi(x, \theta), \quad (2.12)$$

по системе независимых функций $\varphi_k(x, \theta)$ эквивалентно проецированию пространства F функций плотности $f(x) \in F$ в пространство P многочленов $P(x, a, \theta) \in P$. В общем случае этот многочлен зависит от неизвестных параметров – компонентов векторов a и θ .

Задача при этом заключается в нахождении таких значений параметров a и θ , при которых многочлен (2.12) может быть принят в качестве оценки плотности вероятностей

$$\hat{f}(x) = P(x, \hat{a}, \hat{\theta}).$$

Функция распределения и плотность вероятностей случайной величины, как известно, связаны соотношением

$$F(x) = \int_c^d I(x-v) f(v) dv.$$

Тогда, если заменить неизвестную плотность многочленом $P(x, a, \theta)$, а функцию распределения – эмпирической функцией $F_n^*(x)$, то вычисляемые значения представляют собой решение задачи минимизации [117]

$$\{\hat{a}, \hat{\theta}\} = \arg \min_{a, \theta} \rho^2(\tilde{F}(x, a, \theta), F_n^*(x)) + \alpha_j Q(P(x, a, \theta)), \quad (2.13)$$

с ограничениями

$$P(x, \hat{a}, \hat{\theta}) \geq 0 \text{ и } \int_c^d P(x, \hat{a}, \hat{\theta}) dx.$$

здесь $\tilde{F}(x, a, \theta) = \int_c^d I(x-v)P(v, a, \theta)dv$, $Q(\cdot)$ – регуляризирующий функционал, $\rho(\tilde{F}(x, a, \theta), F_n^*(x))$ – расстояние в пространстве функций $\tilde{F}(x, a, \theta)$ и $F_n^*(x)$, а α_j – параметр регуляризации. При этом расстояние в задачах оптимизации обычно задается нормой Евклида

$$\rho(\tilde{F}(x, a, \theta), F_n^*(x)) = \|\tilde{F}(x, a, \theta) - F_n^*(x)\| = \sqrt{\int_c^d (\tilde{F}(x, a, \theta) - F_n^*(x))^2 dx}$$

Тогда задача (2.13) может быть записана в виде

$$\{\hat{a}, \hat{\theta}\} = \arg \min_{a, \theta} \|\tilde{F}(x, a, \theta) - F_n^*(x)\|^2 + \alpha_j Q(P(x, a, \theta)). \quad (2.14)$$

Регуляризирующий функционал в общем случае можно представить выражением

$$Q(x, a, \theta) = \sum_{l=0}^r \int_c^d (P^{(l)}(v, a, \theta))^2 dv. \quad (2.15)$$

Аналогично записывается задача минимизации по гистограмме

$$\{\hat{a}, \hat{\theta}\} = \arg \min_{a, \theta} \|P(x_l, a, \theta) - f_l\|^2 + \alpha_j Q(P(x, a, \theta)). \quad (2.16)$$

Однако в отличие от (2.14) норма здесь задается формулой

$$\|P(x_l, a, \theta) - f_l\| = \sqrt{\sum_{l=1}^n (P(x_l, a, \theta) - f_l)^2}.$$

Регуляризирующий функционал определяется выражением (2.15).

Функция $\tilde{F}(x, a, \theta)$ является, очевидно, линейной по компонентам вектора a и нелинейной – по компонентам вектора θ . При этом, если ввести функцию [30]

$$\psi_l(x, \theta) = \int_c^d I(x-v)\varphi_l(v, \theta)dv, \quad l = 1, \dots, m, \quad (2.17)$$

и вектор $\psi(x) = (\psi_1(x), \dots, \psi_m(x))^T$, функция $\tilde{F}(x, a, \theta)$ принимает вид

$$\tilde{F}(x, a, \theta) = \sum_{k=1}^m a_k \psi_k(x, \theta) = a^T \psi(x, \theta).$$

Отсюда в частности следует, что задачу минимизации (2.14) можно

записать в виде

$$\{\hat{a}, \hat{\theta}\} = \arg \min_{a, \theta} \left\| a^T \psi(x, \theta) - F_n^*(x) \right\|^2 + \alpha_j Q(P(x, a, \theta)). \quad (2.18)$$

2.3.1. Линейно-регрессионная задача аппроксимации плотности вероятностей обобщенным многочленом

Если параметры θ функций базисной системы известны или заданы, то (2.18) является задачей квадратичного программирования и задача оценивания плотности вероятностей может быть представлена формулой [3]

$$\hat{a} = \arg \min_a \left\| a^T \psi(x) - F_n^*(x) \right\|^2 + \alpha_j Q(P(x, a)), \quad (2.19)$$

$$P(x, \hat{a}) \geq 0 \text{ и } \int_c^d P(x, \hat{a}) dx = 1.$$

Регуляризирующий функционал может принять вид [117]

$$Q(x, a) = \int_c^d (P(v, a))^2 dv = 1,$$

При этом (2.17) принимают вид

$$\psi_l(x) = \int_c^d I(x-v) \varphi_l(v) dv = 1, \quad l = 1, \dots, m.$$

Соответственно многочлен

$$P(x, a) = \sum_{k=1}^m a_k \varphi_k(x) = a^T \varphi(x).$$

В соответствии с определением нормы Евклида градиент функционала задачи (2.19)

$$J(a) = \left\| a^T \psi(x) - F_n^*(x) \right\|^2 + \alpha_j Q(P(x, a))$$

можно записать в виде

$$\nabla J(a) = 2 \sum_{l=1}^n (a^T \psi(x_l) - F_n^*(x_l)) \psi(x_l) + 2\alpha_j \int_c^d (a^T \varphi(x)) \varphi(x) dx.$$

Если ввести функцию Лагранжа

$$L(a) = J(a) + \lambda \left(\int_c^d a^T \varphi(x) dx - 1 \right),$$

то решение задачи (2.19) с учетом ограничения равенства находится из системы уравнений

$$\nabla L(a) = \nabla J(a) + \lambda \nabla \int_c^d a^T \varphi(x) dx = 0,$$

$$\int_c^d a^T \varphi(x) dx - 1 = 0,$$

то есть из системы

$$\sum_{l=1}^n (a^T \psi(x_l) - F_n^*(x_l)) \psi(x_l) + \alpha_j \int_c^d (a^T \varphi(x)) \varphi(x) dx + \frac{\lambda}{2} \int_c^d \varphi(x) dx = 0,$$

$$a^T \int_c^d \varphi(x) dx - 1 = 0.$$

Эту систему можно преобразовать к виду

$$\left(\sum_{l=1}^n \psi(x_l) \psi^T(x_l) + \alpha_j \int_c^d \varphi(x) \varphi^T(x) dx \right) a = \sum_{l=1}^n F_n^*(x_l) \psi(x_l) - \frac{\lambda}{2} \int_c^d \varphi(x) dx,$$

$$\left(\int_c^d \varphi(x) dx \right) a = 1.$$

Тогда если ввести матрицу

$$\Psi(\alpha_j) = \left(\sum_{l=1}^n \psi(x_l) \psi^T(x_l) + \alpha_j \int_c^d \varphi(x) \varphi^T(x) dx \right) \quad (2.20)$$

и вектор

$$d_F(\lambda) = \sum_{l=1}^n F_n^*(x_l) \psi(x_l) - \frac{\lambda}{2} \int_c^d \varphi(x) dx, \quad (2.21)$$

то решение системы можно записать в виде

$$\hat{a}(\lambda) = \Psi^{-1}(\alpha_j) d(\lambda), \quad (2.22)$$

где множитель Лагранжа находится из уравнения [18]

$$\left(\int_c^d \varphi^T(x) dx \right) \Psi^{-1}(\alpha_j) \hat{a}(\lambda) = 1. \quad (2.23)$$

Выражения (2.19) – (2.23) дают, таким образом, решение задачи оценивания плотности вероятностей в явном виде при различных значениях

параметра регуляризации α_j . В ряде случаев таким способом удается получить неотрицательный многочлен, что позволяет считать его искомой оценкой плотности вероятностей.

В случае если оценка на некоторых интервалах оказывается отрицательной можно попытаться ввести дополнительные ограничения типа равенства или попытаться изменить параметр регуляризации. В этом плане еще большие возможности предоставляют методы обобщенной локальной аппроксимации.

Гистограммная оценка плотности вероятностей в рассматриваемом случае находится аналогичным способом. Выражения (2.19) – (2.23) сохраняют при этом свою силу, за исключением того, что функции $\psi_l(x)$ в них заменяются на функции $\phi_l(x)$ базисной системы, а эмпирическая функция распределения $F_n^*(x)$ – на эмпирические значения f_l плотности.

2.3.2. Многочлены по базисным системам функций

При оценивании плотностей вероятностей используются тригонометрические и алгебраические полиномы [28, 31, 38, 40, 49]

$$P(x, a) = \sum_{l=1}^m a_l x^{l-1},$$

$$P(x, a) = \sum_{l=1}^{m/2} (a_{2l-1} \cos(l-1)x + a_{2l} \sin lx),$$

многочлены по системам ортогональных функций (функции Чебышёва, Лагерра и др.) и типовым функциям плотностей, например, нормального и двустороннего экспоненциального распределения

$$P(x, a, \theta) = \sum_{k=1}^m a_l \frac{1}{\sigma_l \sqrt{2\pi}} e^{-\frac{(x-m_l)^2}{2\sigma_l^2}},$$

$$P(x, a, \theta) = \sum_{k=1}^m a_l \frac{\lambda}{2} e^{-\lambda_l |x-m_l|}.$$

Коэффициенты алгебраических и тригонометрических многочленов

находятся рассмотренным выше методом решения линейной задачи: методом наименьших квадратов. Аналогично находятся и коэффициенты многочленов по базисным функциям плотностей вероятностей, если их параметры заранее известны или заданы [100]. В этом случае для обеспечения условий, налагаемых на оценку плотности, достаточно, чтобы выполнялись ограничения

$$a_l \geq 0, \sum_{l=1}^m a_l = 1. \quad (2.24)$$

Подобным способом оцениваются, например, смеси двух распределений, представленных функциями плотности $\varphi_1(x)$ и $\varphi_2(x)$, так что искомая плотность вероятностей имеет вид

$$f(x) = \alpha \varphi_1(x) + (1 - \alpha) \varphi_2(x), \alpha \in [0,1]. \quad (2.25)$$

Однако условия (2.24) в общем случае не являются необходимыми и точное следование им, возможно упрощая задачу, не обязано обеспечить наиболее точную аппроксимацию эмпирических данных. Здесь есть некоторая неопределенность, заключающаяся в трактовке многочлена $P(x, a, \theta)$.

Можно, например, принять, что распределение случайной величины действительно является смесью распределений, представленных соответствующими базисными плотностями. Тогда задача построения многочлена – это задача параметрического оценивания и условия (2.24) становятся не только достаточными, но и необходимыми.

Если же распределение случайной величины не рассматривается как смесь базисных плотностей, то задача оценивания относится к классу непараметрических задач и многочлен должен удовлетворять ограничениям задачи (2.23), из которых условия (2.24) следовать не обязаны. В данном случае имеет место задача аппроксимации.

2.4 Выводы

1. Разработка практических решений алгоритмов минимизации приведенных функционалов обеспечивает выбор порядка полиномов и базисных функций, при аппроксимации плотности распределения.

2. При практическом применении алгоритмов минимизации приведенных функционалов имеется возможность использования систем степенных и ортогональных многочленов. При задействовании методов сплайновой аппроксимации также возможно получить результаты, отвечающие требуемым параметрам. Построение имитационных моделей с помощью прикладных средств выявляет, что вышеуказанные методы имеют достаточно высокую эффективность. При некоторых условиях ограничения-неравенства были применены в автоматическом режиме и задача условной минимизации в таком случае не требовала решения. Для оценки смеси распределения применим метод барьерной функции.

ГЛАВА 3. ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК АКУСТИЧЕСКИХ РЕЧЕВЫХ СИГНАЛОВ

3.1. Методы исследования распределения плотности вероятностей акустических сигналов

Для получения оптимальных результатов при оценивании распределений речевых сигналов и сигналов акустических помех, требуется наличие экспериментальных результатов измерений. В таком случае можно применить параметрические и непараметрические, прямые и косвенные методы анализа [39].

Основным направлением исследования является сопоставление исходного объема данных одному из известных распределений [66]. Вышеуказанные методы предполагают, что плотность вероятностей имеет вид: $f(x, \theta) \equiv f_{\theta}(x)$, где $x \in R^n$ и $\theta \in R^n$ соответствуют строкам неизвестных параметров и случайным переменным. В задачах оценки считается, что вектор параметров θ неизвестен. При этом, функция $f(x, \theta)$ может быть известна. В таком случае методы восстановления являются параметрическими [9, 24].

При осуществлении аппроксимации $f(x, \theta)$ должна соответствовать ограничениям

$$f(x, \theta) \geq 0 \text{ и } \int_{-\infty}^{\infty} f(x, \theta) dx = 1. \quad (3.1)$$

Четко разделить параметрические и непараметрические методы возможно не во всех случаях. Так, вычисление приближения данных смесью известных распределений [66], которые определяются функциями плотности $\psi_k(x, \theta_k)$, $f(x, \theta) = \sum_k a_k \psi_k(x, \theta_k)$, $\sum_k a_k = 1$, имеет смысл определять как непараметрические задачи. Но при явном определении коэффициентов $a_k \geq 0$, рассматриваемую задачу явно необходимо определить как параметрическую [39].

В класс непараметрических задач входят задачи нелинейной и линейной регрессии, а также задачи, которые можно свести к применению метода наименьших квадратов. Подобные методы решения вышеуказанных задач также называются проекционными.

Нужно учитывать, что определение непараметрических методов, приведенное выше, применяется только в математической статистике [7, 48]. В области теории систем, оптимизации, аппроксимации и приближения они, определяются как параметрические [3, 49], исходя из смысла задач, который состоит в нахождении конечного числа неизвестных параметров [39].

3.1.1. Методы моментов и подстановки

Задача получения оценки неизвестного распределения с помощью метода моментов вычисляется из условий равенства некоторого числа теоретических моментов, имеющих зависимость от неизвестного вектора параметров θ [39],

$$m_{g_k}(\theta) = E_{\theta}\{g_k(x)\} = \int_{-\infty}^{\infty} g_k(x) dF_{\theta}(x), \quad k = 1, \dots, m,$$

где $g(x) \in R^m$ вектор линейно независимых функций случайной величины x , и их эмпирических средних значений, определяемых по формуле

$$\bar{m}_{g_k}(\theta) = \frac{1}{N} \sum_{l=1}^N g_k(x_l), \quad k = 1, \dots, m.$$

Система уравнений относительно неизвестного вектора параметров θ записывается при этом в виде

$$m_{g_k}(\theta) = \frac{1}{N} \sum_{l=1}^N g_k(x_l), \quad k = 1, \dots, m,$$

или в матричном виде

$$m_g(\theta) = \frac{1}{N} \sum_{l=1}^N g(x_l). \quad (3.2)$$

Если функции $m_{g_k}(\theta)$ разрешимы относительно неизвестного вектора θ параметров, то оценка неизвестного вектора может быть записана в виде

$$\hat{\theta} = m_g^{-1} \left(\frac{1}{N} \sum_{l=1}^N g(x_l) \right).$$

Решение системы (3.2) может и не существовать. В этом случае оценку параметров можно получить посредством минимизации квадратичной функции невязок

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{m} \sum_{k=1}^m (m_{g_k}(\theta) - \bar{m}_{g_k})^2.$$

Нахождение явных выражений для функций $m_{g_k}(\theta)$ не всегда является простой задачей. Ее решение может быть получено с помощью соответствующей аппроксимации теоретических моментов. Решение в этом случае можно, в принципе, получить и посредством численного решения соответствующего (3.2) векторного интегрального уравнения

$$\int_{-\infty}^{\infty} g(x) dF_{\theta}(x) = \frac{1}{N} \sum_{l=1}^N g(x_l).$$

В рассмотренной постановке метод моментов фактически эквивалентен методу подстановки, в соответствии с которым неизвестный вектор θ параметров распределения представляется в форме некоторого функционала G от распределения P

$$\theta = G(P).$$

При этом оценка находится из выражения

$$\hat{\theta} = G(P_n),$$

где P_n обозначает эмпирическое распределение наблюдаемого вектора x .

Если зависимость θ от распределения P имеет неявный вид $\theta = G(\theta, P)$, то оценку можно найти посредством минимизации $|G(\theta, P_n)|$,

$$\hat{\theta} = \arg \min_{\theta} |G(\theta, P_n)|. \quad (3.3)$$

Эти методы применимы при решении, как параметрических, так и непараметрических задач. При этом если в первом случае аналитическое выражение плотности вероятностей считается известным, то во втором –

принимается, что плотность вероятностей приближается функций из функционального пространства, задаваемого вектором параметров θ . То есть, во втором случае находится проекция плотности на множество функций $\mathfrak{F} = \{P(x, \theta) \mid \theta \in \Theta\}$.

Очевидно, что для обеспечения необходимой точности это пространство должно быть достаточно полным. С другой стороны полученная таким способом функция не обязана являться плотностью, что может потребовать сужения пространства функций $\mathfrak{F} = \{P(x, \theta) \mid \theta \in \Theta\}$ до пространства, удовлетворяющего ограничениям (3.1).

3.2. Модель одномерной функции плотности вероятности оцифрованных акустических речевых сигналов.

3.2.1. Алгоритм гистограммной оценки плотности вероятностей речевого сигнала.

Функция плотности вероятностей является важнейшей статистической характеристикой акустического речевого сигнала. Исследованы и разработаны алгоритмы построения гистограмм по цифровым дискретам на заданном временном интервале [31].

$$P(x_i) = \frac{1}{N} \sum_{j=-(2^n-1)}^{2^n-1} \sum_{i=1}^N \delta_{i,j}, \quad (3.4)$$

где $P(x_i)$ – оценка вероятности выборки X_i ,

x_i – i -тая выборка, представленная номером уровня квантования,
 $x_i \in [-(2^n - 1), 2^n - 1]$, $|x_{i,\max}| = 2^n - 1$,

N – число выборок в рассматриваемом акустическом речевом сигнале,

i – номер выборки в пределах интервала ($i = 1, 2, \dots, N$),

j – номер интервала гистограммы для анализа числа попаданий в него

x_i - тых выборок, $j \in [-(2^n - 1), 2^n - 1]$,

n – количество разрядов двоичного числа, представляющего выборку

x_i и представляющего номер интервала гистограммы,

$$\delta_{i,j} - \text{символ Кронекера, имеющий вид } \delta_{i,j} = \begin{cases} 1, & \text{если } x_i = j \\ 0, & \text{если } x_i \neq j \end{cases}$$

Блок-схема рассматриваемого алгоритма (3.4) представлена на рис 3.1.

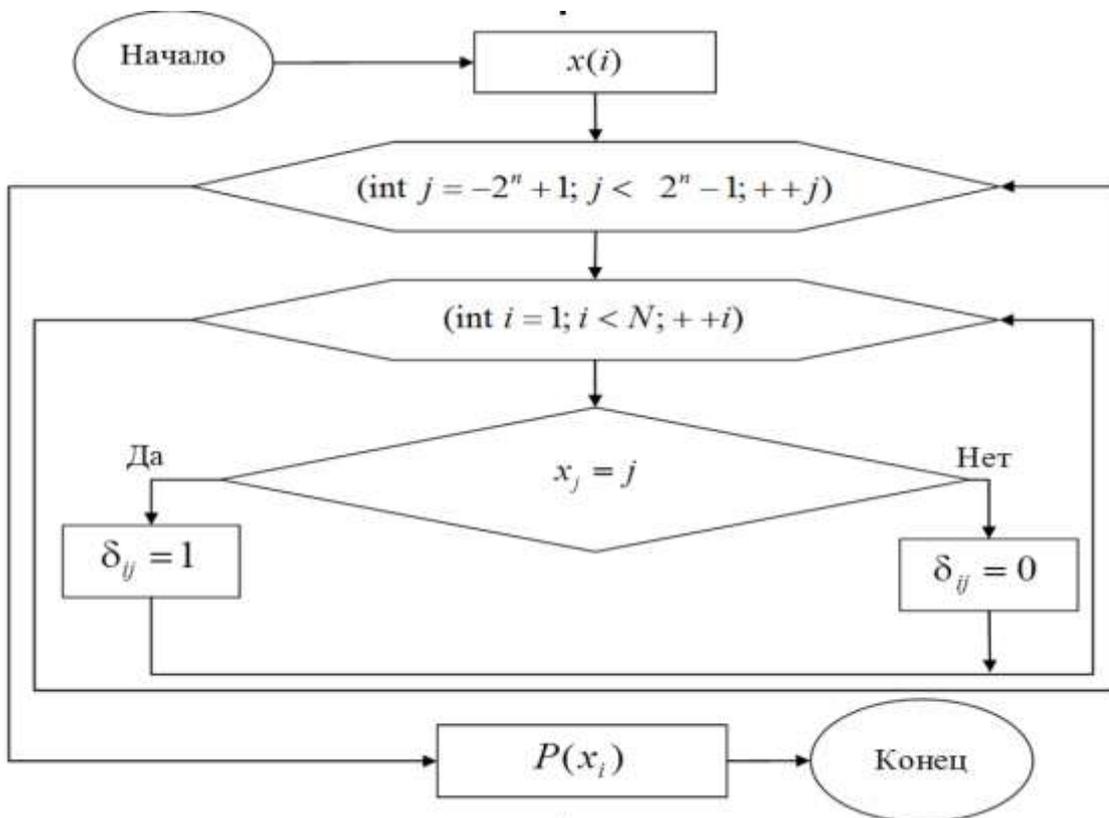
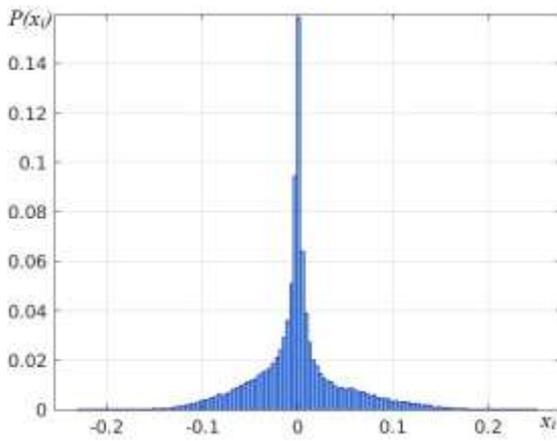


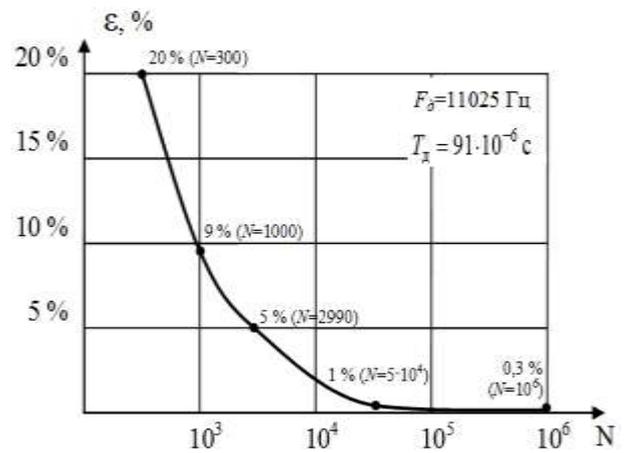
Рисунок 3.1 – Блок схема алгоритма формирования гистограмм по дискретным отсчетам n -разрядными двоичными кодами

Как видно из блок-схемы алгоритма программы, представленной на рисунке 3.1, можно получить гистограммную оценку закона распределения речевого сигнала по оцифрованным дискретам. Так как форма гистограммы носит изрезанный характер, при первоначальном объеме выборки менее тысячи отсчетов, для её сглаживания необходимо расширить диапазон разбиений. Чтобы повысить степень сглаживания гистограммы необходимо использовать объем выборки более 10^4 отсчетов [31, 44]. Так как речевой сигнал имеет нестационарный характер – форма гистограммы зависит от исследуемого фрагмента.

Приведенная на рисунке 3.2 гистограмма соответствует периодам речевой активности.



а.



б.

Рисунок 3.2 – (а) Гистограмма вероятностей дискретов речевого сигнала на интервале 10^6 ; (б) Зависимость погрешности гистограммной оценки от числа выборок N

Погрешность гистограммной оценки сигнала может быть выражена, как норма отклонения этой оценки от априорно известной плотности вероятностей. Норму при этом можно задать любым известным способом [24].

Для вычисления абсолютной погрешности оценки может применяться следующее выражение [28]

$$\Delta = \frac{1}{2^{n+1}} \sum_{j=-2^n+1}^{2^n-1} \left| P_{x_j}^N - P_{x_j} \right|, \quad (3.5)$$

где $P_{x_j}^N$ – гистограммная оценка вероятности случайной величины x_j при N выборках, P_{x_j} – действительная вероятность принадлежности x_j к интервалу $(n_{k-1}, n_k]$ разбиения гистограммы. Если исследуемый отрезок речевого сигнала стационарный, то P_{x_j} можно получить в ходе проведения экспериментальных исследований с использованием объема данных $K > 10^6$, N – число дискретов на исследуемом интервале, $N \leq K$.

Согласно сказанному выше, относительная погрешность ε вычисляется по формуле [28, 34]

$$\varepsilon = \frac{1}{2^{n+1}} \sum_{j=-2^n+1}^{2^n-1} \frac{1}{P_{x_j}} \left| P_{x_j}^N - P_{x_j} \right|. \quad (3.6)$$

Анализ полученных результатов показал, что если объем исследуемых выборок речевых сигналов $N \geq 0,5 \cdot 10^5$ и период их дискретизации $T_d = 91 \cdot 10^{-6} \text{ с}$, то возможно достижение погрешности менее 1%. При увеличении объема выборки до миллиона дискретов погрешность будет составлять менее 0,3%. Для получения погрешности 5% достаточно исследовать сигнал на интервале 0,3 с. В том случае, если продолжительность анализируемого отрезка сигнала будет составлять 10^{-2} с и $3 \cdot 10^{-1} \text{ с}$, его погрешность будет около 19% и 8 %.

3.2.2. Аппроксимация гистограммы плотности вероятностей акустического речевого сигнала.

Аппроксимация результатов, полученных в предыдущем разделе, может производиться с использованием нескольких математических моделей. При этом можно применять тригонометрические и алгебраические полиномы, а так же полиномы по системам гауссовых или экспоненциальных функций. При использовании тригонометрических и алгебраических полиномов модель может быть представлена формулами [2]

$$\rho(x) = \sum_{k=0}^M a_k x^{M-k}, \quad (3.7)$$

$$\rho(x) = \sum_{k=0}^M (a_k \cos(k\alpha x) + b_k \sin(k\alpha x)), \quad (3.8)$$

Решение задач аппроксимации обеспечивается путем вычисления коэффициентов полинома. При известном значении параметра α вышеуказанную задачу можно отнести к классу регрессионных задач, использующих метод наименьших квадратов [18, 19]. Для получения оценки используется функция невязки в вида

$$Q(\theta) = \sum_{k=0}^M (\bar{f}(x_k, \theta) - P(x_k))^2,$$

где $\bar{f}(x_k, \theta)$ усредненная на k -ом интервале гистограммы оценка функции

плотности вероятности $f(x, \theta)$, а θ – вектор параметров p_k , a_k и b_k , определяющих линейную зависимость в полиномах.

При выборке ограниченного объема рассмотренную выше функцию невязки можно представить в следующем виде

$$Q(\theta) = \sum_{k=0}^M (f(x_k, \theta) - P(x_k))^2.$$

При аппроксимации гистограмм полиномами по системам экспоненциальных и гауссовых функций приведенные выше формулы также могут быть применимы. Однако в данном случае указанные многочлены должны быть представлены в следующем виде:

$$\rho(x) = \sum_{k=1}^M A_k e^{-\frac{(x-C_k)^2}{B_k}},$$

$$\rho(x) = \sum_{k=1}^M A_k e^{-\frac{|x-C_k|}{B_k}}. \quad (3.9)$$

При использовании систем гауссовых и экспоненциальных функций решение задачи минимизации невязки будет уже попадать в класс нелинейных регрессионных задач [18, 19], при условии, что B_k , C_k не будут определены ранее.

Анализ найденной функции плотности вероятностей показал, что она одномодальная и симметричная, и исследуемую функцию можно представить как математическое ожидание для рассматриваемого случайного процесса. Его оценка может быть проведена с применением выражения [45]

$$C_k = \overline{x(nT)} = \frac{1}{N} \sum_{n=0}^{N-1} x(nT).$$

Когда принимается, что для рассматриваемой выборки отсчетов речевого сигнала $\overline{x(nT)} \approx 0$ либо при условии вычитания оценки $\overline{x(nT)}$, анализируемый процесс станет возможным отнести к процессу, имеющему математическое ожидание, равное 0, при этом выражение (3.9) запишется в виде

$$\rho(x) = \sum_{k=1}^M A_k e^{-\frac{|x|}{B_k}}, \quad (3.10)$$

в котором параметры A_k , B_k определяются за несколько итераций.

При наличии гладкой восстанавливаемой функции плотности обеспечение заданной погрешности аппроксимации происходит с помощью многочлена (3.10). Главной задачей в таком случае является подбор такого вида полинома, чтобы он имел минимально возможный порядок.

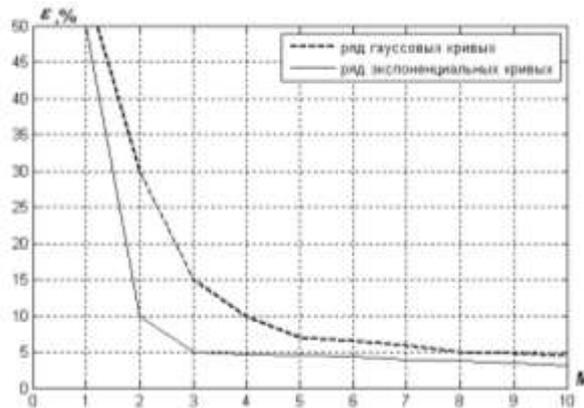


Рисунок 3.3 – Зависимости погрешности аппроксимации от вида и порядка многочлена

По рисунку 3.3 видно, что более оптимальным является подход к аппроксимации на основе полинома, построенного с применением экспоненциальных функций, который при значении порядка, равного 3, позволяет получить погрешность представления менее 5%. Для обеспечения восстановления функции плотности вероятностей, может быть применим метод локальной аппроксимации многочленами, полученными в соответствии с (3.9).

Ряд дополнительных ограничений, например, типа равенства, могут быть наложены в узлах сопряжения. Чтобы найти коэффициенты полиномов для всех интервалов анализа, можно применить подход с использованием минимизации невязки, в соответствии с выражением

$$Q_l(\theta) = \sum_{k=0}^M w(x_k - v_l) (\bar{f}(x_k, \theta) - P(x_k))^2$$

в котором $w(x)$ – является весовой функцией, которая задает «окно данных», и v_l – параметр, определяющий на сколько сдвигается «окно данных» при анализе на интервале с номером l , причем $l = 1, \dots, L$.

Для выполнения аппроксимации применима последовательность $\theta_l = \arg \min_{\theta} Q_l(\theta)$. Полином (3.10) позволяет проводить аппроксимацию негладких функций на всем интервале их рассмотрения.

Рисунок 3.4 демонстрирует результаты аппроксимации функции плотности вероятностей тестового речевого сигнала набором экспоненциальных функций, имеющих порядок, равный 3.

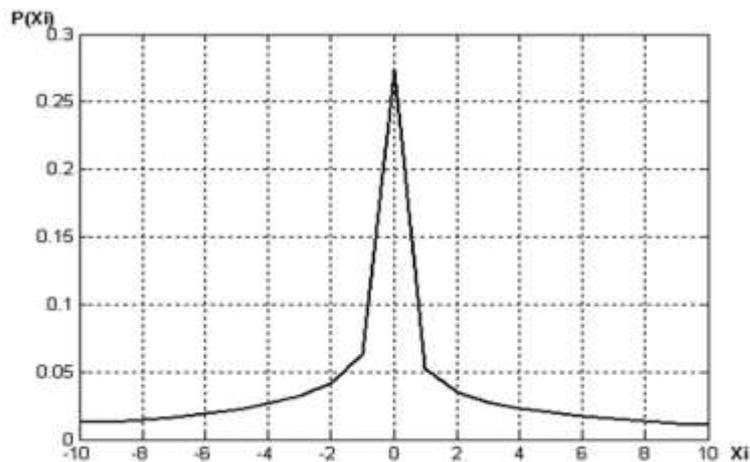


Рисунок 3.4 – Полученная аппроксимация полиномом третьего порядка с применением экспоненциальных функций.

Указанные на рисунке 3 результаты были получены при следующих значениях параметров: порядок полинома равен 3, значения коэффициентов следующие $A_1 = 0.0338$, $A_2 = 0.0032$, $A_3 = 0.2371$, $B_1 = 9.6478$, $B_2 = 7.73975$, $B_3 = 0.4770$, погрешность $\varepsilon \leq 5\%$.

В соответствии с (3.10) многочлен принимает вид

$$\rho(x) = A_1 e^{-\frac{|x|}{B_1}} + A_2 e^{-\frac{|x|}{B_2}} + A_3 e^{-\frac{|x|}{B_3}}, \quad (3.11)$$

где коэффициенты A_i и B_i определены методом последовательных итераций, их значения представлены на рисунке 3.4.

Таким образом, разработан метод восстановления распределения плотности вероятности аудиосигналов с погрешностью не более 5%, который

определил первое основное положение, выносимое на защиту, в соответствии с которым создана модель плотности вероятности речевых сигналов в виде многочлена 3-го порядка, который был сформирован с применением экспоненциальных функций, представленных (3.11) .

3.3. Формирование моделей внешних акустических помех в системах громкоговорящей связи и оповещения.

3.3.1. Исследование одномерных функций плотности вероятностей акустических шумовых помех

Системы обмена и передачи акустической информацией в оперативно-командной обстановке зачастую функционируют в условиях активного воздействия шумовых помех и акустических внешних шумов. В связи с этим снижается эффективность передачи информации и становится необходимым исследовать внешние помехи и их свойства для борьбы с ними.

Одномерная функция плотности вероятностей может быть рассмотрена в качестве наиболее значимой статистической характеристики акустических шумов и помех. За основу была взята следующая статистика вида помех, включающая в себя шумы машинного отделения судна, морские и ветровые шумы [112, 119]. С помощью алгоритма построения гистограмм, представленного выше в главе 3, проводилась гистограммная оценка плотности акустических помех.

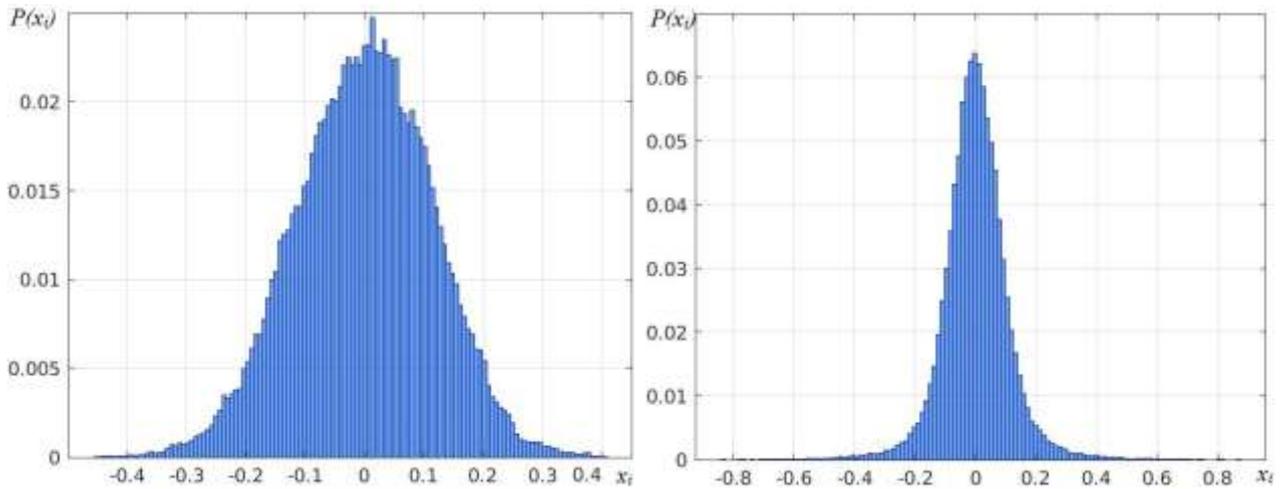
Гистограммы $P_{ак.п.}(x_i)$ для разных источников акустических помех показаны на рисунке 3.5 «а», «б».

Для уменьшения погрешности представления были рассмотрены и исследованы аппроксимации в базисе гауссовых функций и в виде ряда в базисе экспоненциальных функций.

$$P_{ак.п.}(x_i) = \sum_{i=1}^M A_i e^{-\frac{|x_i|}{B_i}},$$

и в базисе гауссовых функций в виде

$$P_{ак.н.}(x_i) = \sum_{i=1}^M A_i e^{-\frac{x_i^2}{B_i}}.$$



а. Ветровой шум

б. Морской шум

Рисунок 3.5 – Гистограммы распределения вероятности дискретов шумовых акустических помех ветра и моря.

На рисунке 3.5 в виде графика приведены гистограммы табличных функции распределения вероятности дискретов акустических помех и шумов. Рисунок 3.6 демонстрирует полученные для шумовых помех аппроксимации с применением подхода на основе использования гауссовых кривых и ряда экспонент. Параметры аппроксимации были определены с применением итерационного алгоритма [28]. Порядок полинома $M = 3$, обеспечил погрешность, не превышающую 5%, при аппроксимации с использованием гауссовых кривых, и не превышающую 9% в ходе аппроксимации рядом экспоненциальных функций.

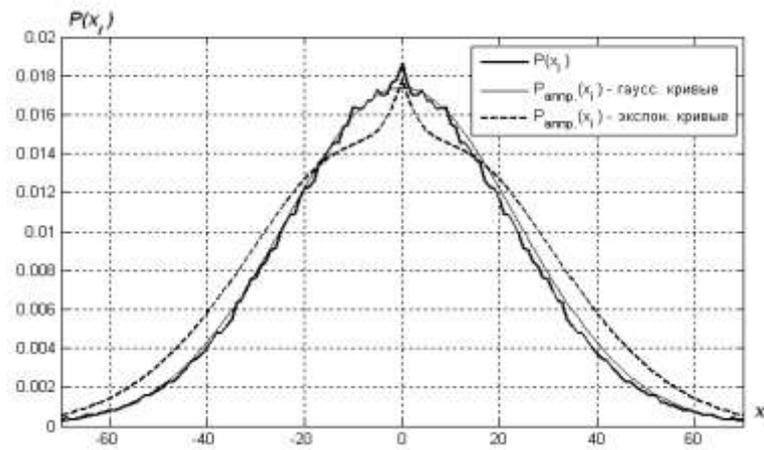


Рисунок 3.6 – Графики функций аппроксимации $P(x_i)$, при порядке $M=3$, построенные для разных базисов

Для достижения показателя погрешности менее пяти процентов при моделировании возможно использовать математическую модель закона распределения вероятности дискретов внешних помех, представленную в виде [32]

$$P_{ак.н.}(x_i) = A_1 e^{-\frac{x_i^2}{B_1}} + A_2 e^{-\frac{x_i^2}{B_2}} + A_3 e^{-\frac{x_i^2}{B_3}},$$

Где $A_1 = 0.0268$; $A_2 = -0.0207$; $A_3 = 0.0122$, $B_1 = 1110.281$; $B_2 = 1141.1400$; $B_3 = 1209.0742$; При этом достигается погрешность моделирования менее 5%.

Таким образом, проведенные исследования позволяют рассчитать дисперсию и математическое ожидание, применяя модель одномерной функции плотности вероятности звуковых помех с допустимой погрешностью. Результаты имитационного моделирования подтвердили, что применение аппроксимации функции плотности вероятностей с использованием ряда гауссовых кривых обеспечивают достижение погрешности не более 5% процентов при использовании полинома с порядком, равным 3.

3.4. Исследование автокорреляционных функций (АКФ) речевых сигналов и шумовых помех.

По исследованиям ВКФ и АКФ речевых сигналов существует ряд трудов [13, 22, 70]. Однако в условиях эхосигналов и акустических помех, например, в специализированных помещениях или штабе оперативно-командной связи, необходимо учитывать факторы дополнительного влияния. В этом случае становится возможным получить более точную информацию о длительности фонем, времени корреляции и сегментов речевого сигнала, используя кратковременные АКФ.

Если предположить, что речевой сигнал — это случайный эргодический процесс, для исследования и получения экспериментальной АКФ можно задействовать следующее предложенное выражение [35, 36, 90]

$$R(kT) = \frac{1}{N} \sum_{n=0}^{N-1} U(nT)U(nT + kT), \quad 0 \leq k \leq K, \quad K < N,$$

в котором $K + N$ — количество анализируемых дискретов при расчете кратковременной АКФ, k — текущая задержка, T — период дискретизации, n — номер выборки.

Чтобы оценить форму и ширину спектра речевого сигнала необходимо применить нормирование и усреднение с помощью следующих выражений

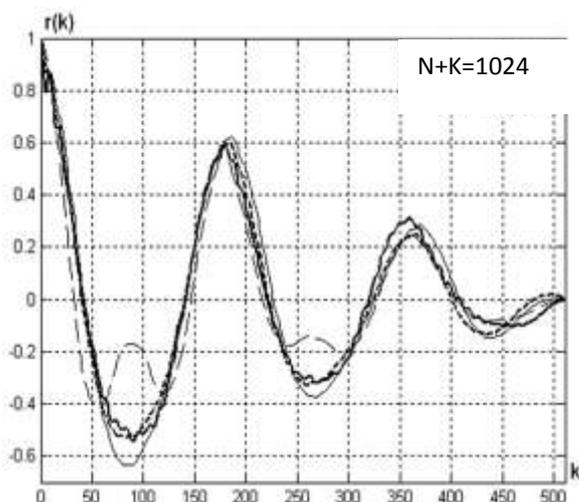
$$r(kT) = \frac{\frac{1}{N} \sum_{n=0}^{N-1} U(nT)U(nT + kT)}{\sigma^2}, \quad 0 \leq k \leq K. \quad (3.13)$$

$$R_M(kT) = \frac{1}{M} \frac{1}{N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} U(nT + mNT)U(nT + mNT + kT), \quad 0 \leq k \leq K. \quad (3.14)$$

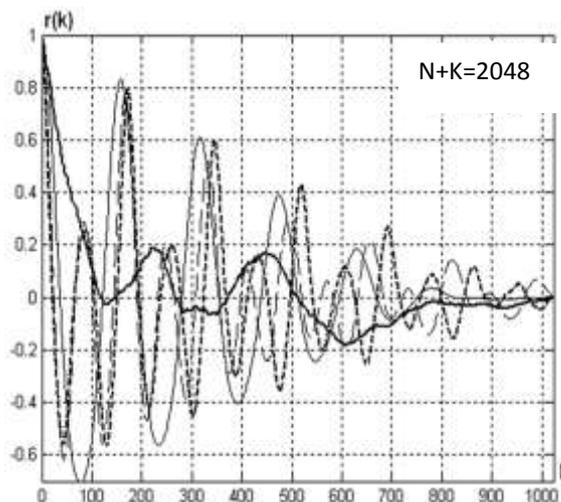
где M — количество усредненных АКФ, а m — номер сегмента речи.

На графиках, изображенных на рисунке 3.7 показаны кратковременные автокорреляционные функции, соответствующие невокализованным и вокализованным сегментам речевого сигнала. Длина интервала анализа подбирается таким образом, чтобы статистические характеристики не

изменялись в пределах его длительности.



а. Промежуток анализа = 512
дискретов



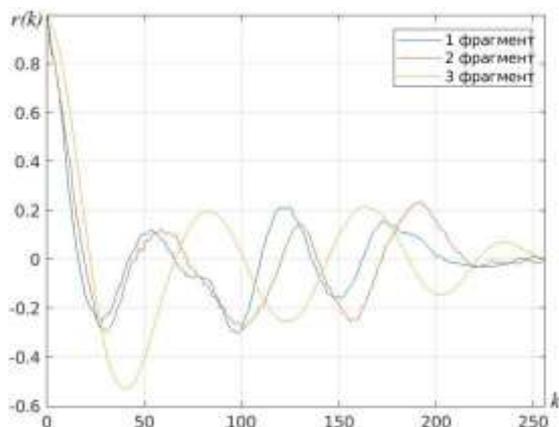
б. Промежуток анализа = 1024
дискрета

Рисунок 3.7 – Кратковременные автокорреляционные функции, 4 сегмента речи, $T=22,676 \cdot 10^{-6}$ с,

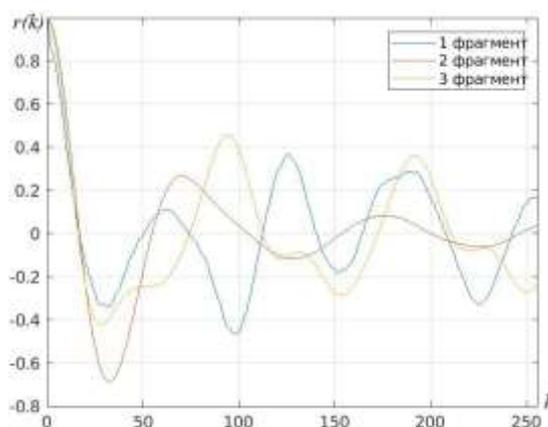
Приведенные на графиках результаты получены на промежутках 0,0116 и 0,0232 секунды, что соотносится 512 и 1024 дискретам.

Используя полученные экспериментальные данные, становится возможным выполнить расчет спектральной плотности мощности. Это позволит получить формантное распределение мощности.

В результате, удалось получить данные о характере речевого сигнала на всей реализации (рисунок 3.8).



а. K=256 отсчетов



б. K=512 отсчетов

Рисунок 3.8 – Графики усредненной АКФ для различных фрагментов речевых сигналов

Имея представление о форме усредненной АКФ возможно провести оценку полосы частот энергетического спектра, рассчитывая время корреляции речевого сигнала следующим выражением [6]

$$\Delta\tau = \int_0^{\infty} |r(\tau)| d\tau,$$

где $r(\tau)$ – нормированная АКФ.

Таким образом, был получен алгоритм для расчета времени корреляции:

$$\frac{\Delta\tau}{T} = \sum_{k=0}^{K-1} |r(kT)| = \sum_{k=0}^{K-1} \left| \frac{\sum_{n=0}^{N-1} \sum_{k=0}^{K-1} U(nT)U(nT+kT)}{\sum_{n=0}^{N-1} U^2(nT)} \right|. \quad (3.15)$$

Используя усреднённую автокорреляционную функцию, рассчитано время корреляции

$$\frac{\Delta\tau_M}{T} = \sum_{k=0}^{K-1} \left| \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} U(nT+mNT)U(nT+mNT+kT)}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} U^2(nT+mNT)} \right|. \quad (3.16)$$

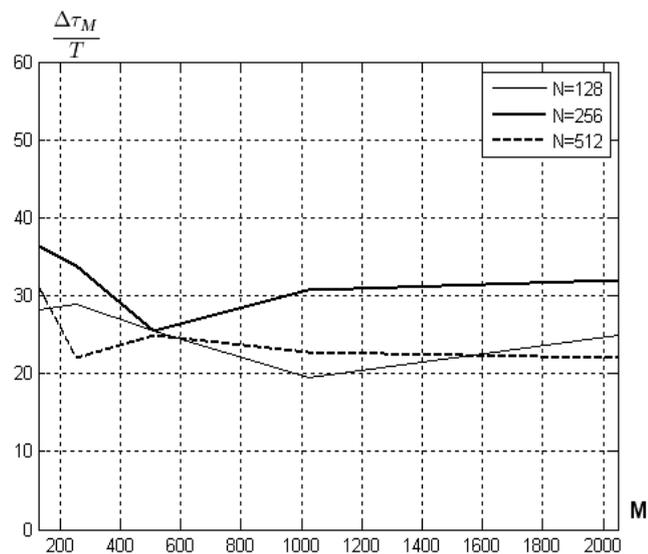


Рисунок 3.9 – Зависимость времени корреляции от размера окна данных
($T=22,676 \cdot 10^{-6}$ с)

Полученные результаты (рисунок 3.9) показывают зависимость времени корреляции речевого сигнала от размера окна данных в частотном

диапазоне от 1-2 кГц и выше. Оно находится в пределах 0,4-0,9 миллисекунды для 20-36 периодов дискретизации.

3.4.1. Исследование АКФ внешних акустических шумовых помех.

АКФ шумов и помех с усреднением и нормированием вычисляется аналогично автокорреляционной функции речевых сигналов как показано в предыдущем разделе.

На рисунке 3.10 показана блок-схема алгоритма вычисления автокорреляционной функции шумовых помех.

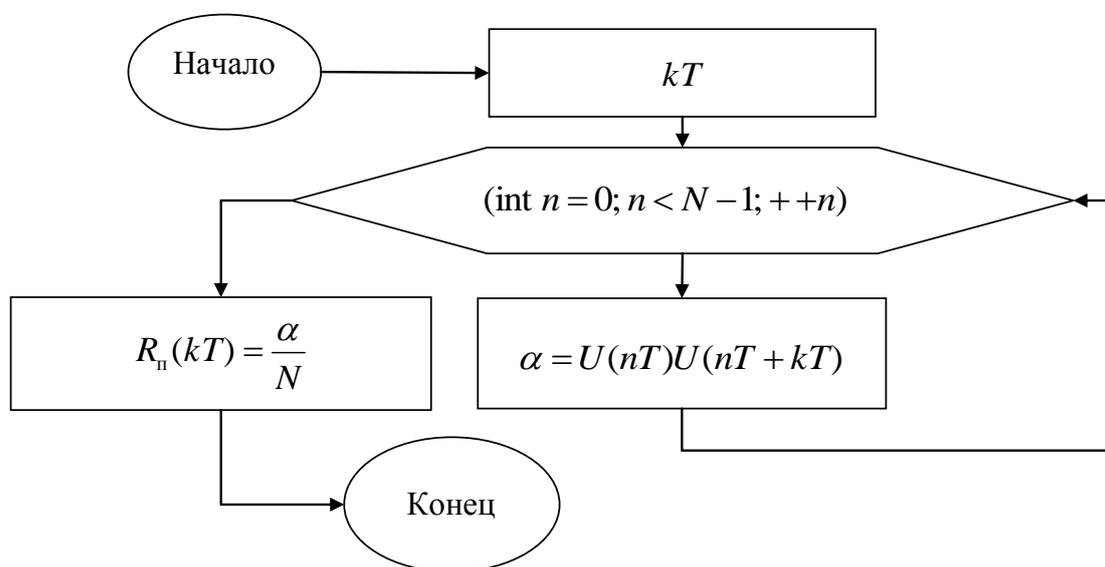
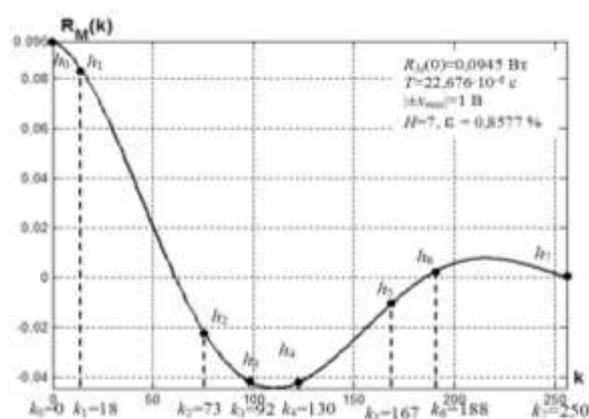
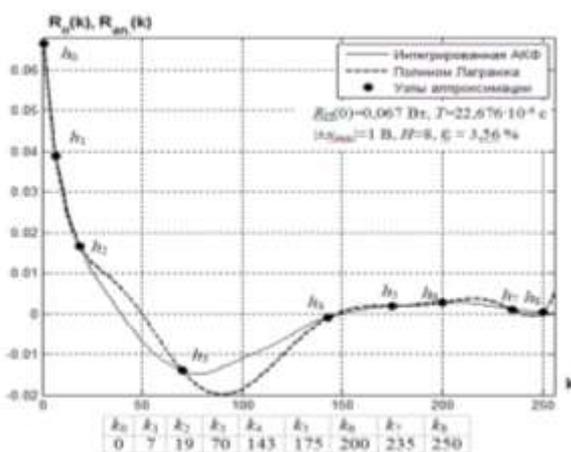


Рисунок 3.10 – Структура блок-схемы программы реализации алгоритма вычисления АКФ шумовых помех $R_n(kT)$.

Расчеты автокорреляционных функций помех выполнялись с применением статистических характеристик определенных типов помех [112, 119].

Графики усредненных $R_n(kT)$ автокорреляционных функций шумов изображены на рисунке 3.11.



а. АКФ шума моря

б. АКФ шума ветра

Рисунок 3.11 – Усредненные автокорреляционные функции внешнего шума

На рисунке 3.11 показаны результаты, которые демонстрируют, что почти не происходит изменения формы автокорреляционной функции при смене номера исследуемого фрагмента определенной помехи. Это может говорить о стационарном характере исследуемых шумов.

Автокорреляционную функцию шумовых помех предлагается аппроксимировать многочленом Лагранжа, используя следующую формулу

$$R_{\Pi}(k) = \sum_{h=0}^H R(k_h) \prod_{\substack{\ell=0, \\ \ell \neq h}}^H \frac{k - k_{\ell}}{k_h - k_{\ell}}, \quad (3.19)$$

Необходимые значения $R(k_h)$ и k_l необходимо брать подстановкой из таблиц 3.1-3.3, в которых представлены сеточные функции автокорреляционной функции шума машинного отделения, ветра и моря.

Таблица 3.1

k_h	0	7	19	70	143	175	200	235	250
$R_{\Pi}(k_h)$	0,0671	0,04	0,0134	-0,008	-0,001	0,002	0,003	0	0

Таблица 3.2

k_h	0	18	73	92	117	167	188	250
$R_{\Pi}(k_h)$	0,0945	0,0537	-0,0358	-0,0453	-0,036	0,0076	0,0123	0

Таблица 3.3

k_h	0	12	20	30	60	94	126	158	187	214	226	238	248	255
$R_h(k_h)$	0,12	0,072	0	-0,02	0,0144	-0,012	0,0072	-0,002	0	-0,0016	0,001	0	0	0

Из показанных графиков АКФ на рисунке 3.11 можно сделать вывод, что достигается погрешность менее четырёх процентов. Приведенный результат подтверждает эффективность предложенного решения аппроксимацией интерполяционным многочленом Лагранжа при порядках полинома $N=7\dots 13$.

Таким образом, результаты проведенных исследований отражают, что среднее время корреляции для шума моря составляет 40-80 $T_{дискр}$. Время корреляции для шума ветра составляет около 110 $T_{дискр}$. При исследованиях времени корреляции использовался период дискретизации $T_{дискр}=22,676\cdot 10^{-6}$ с. В результате время корреляции исследуемых видов шумов составляет $0,9\dots 3\cdot 10^{-3}$ с, а энергия концентрируется в низкочастотной области спектра. До 1200 Гц для шума машинного отделения и до 400 для морских и ветровых шумов.

3.5. Исследование спектральных характеристик речевых сигналов.

Преобразование Фурье на конечном числе выборок N имеет следующий вид [42, 70]

$$S(e^{j\omega}) = \sum_{n=0}^{N-1} u(n)e^{-j\omega nT},$$

где T, F_d – частота и период дискретизации, $F_d \geq 2f_e$.

Зачастую выражение дискретного преобразования Фурье представляется в виде

$$S(l) = \sum_{n=0}^{N-1} u(n)w_N^{ln}, \text{ если } L=N, \text{ то } 0 \leq l \leq N-1, \quad (3.21)$$

при условии что $W_N = e^{-j\frac{2\pi}{N}}$ и $S(l) = S(j \cdot \Delta\omega \cdot l) = S(j \cdot \omega_l)$.

В таком случае обратное преобразование Фурье записывается в виде

$$u(n) = \frac{1}{N} \sum_{l=0}^{N-1} S(l) W_N^{-ln}.$$

При недостаточном значении N для повышения разрешения по частотам спектральных функций, допускается вводить новую последовательность длины L ($L > N$) вида

$$u(n) = \begin{cases} u(n), & 0 \leq n \leq N-1, \\ 0, & N \leq n \leq L-1. \end{cases}$$

Таким образом, дополняя последовательность конечной длины нулевыми отсчетами становится возможным достичь необходимого разрешения по частоте спектральной функции.

3.5.1. Алгоритмы исследования характеристик спектральной плотности мощности акустических речевых сигналов

Рассчитывая коэффициенты спектральной плотности, получают распределение уровней энергии спектральных составляющих. По этим уровням определяется полоса частот и границы частотного спектра для речевых сигналов [37, 56].

Таким образом, информацию о форме энергетического спектра речевых сигналов получаем с помощью дискретного преобразования Фурье от АКФ речевых сигналов по формуле [70]

$$G(\omega) = \sum_{k=-\infty}^{\infty} R(kT) e^{-j\omega kT}, \quad (3.22)$$

в которой $R(kT)$ – автокорреляционная функция речевых сигналов.

После этого возможным становится нахождение усредненной функции спектральной плотности мощности, применяя в выражении выше автокорреляционную функцию на конечном интервале из N дискретов

$$G(\omega_l) = \sum_{l=0}^{N-1} \sum_{k=0}^{K-1} R(kT) e^{-j\omega_l kT} \quad (3.23)$$

Здесь l – номер спектральной дискретной составляющей в $\omega_l = \frac{2\pi F_\Delta}{N} \cdot l$.

После этого необходимо вычислить плотность мощности с помощью дискретного преобразования Фурье, используя вышеуказанное выражение. Тогда оно принимает вид

$$G(\omega_l) = \sum_{l=0}^{N-1} \sum_{k=0}^{K-1} r(kT) \cos \omega_l kT,$$

в котором $r_M(kT)$ – нормированная автокорреляционная функция.

На рисунке 3.12 представлена структура блок-схемы программы реализации алгоритма вычисления энергетического спектра по формуле 3.23.

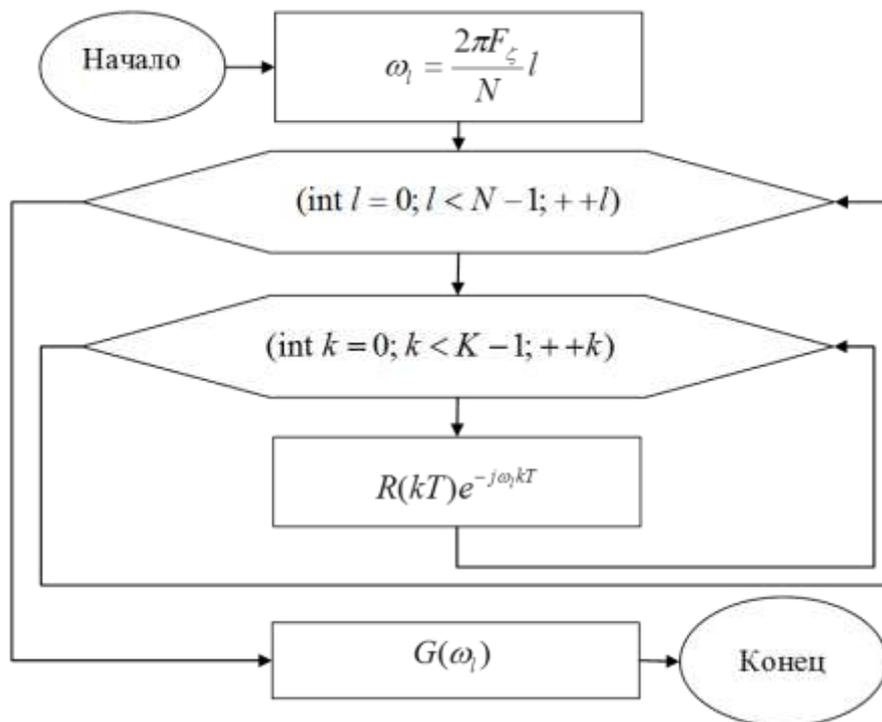


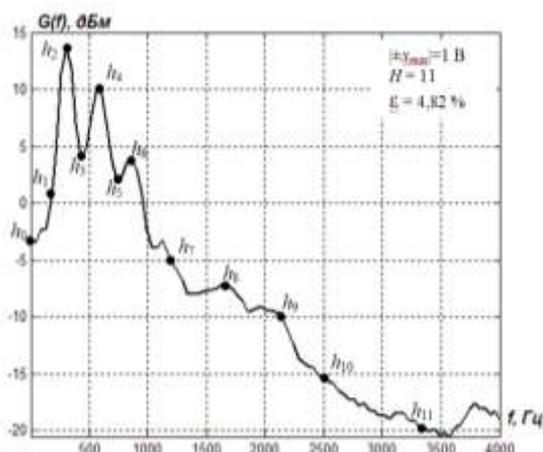
Рисунок 3.12. Структура блок-схемы программы реализации.

В результате проведенных исследований был разработан алгоритм вычисления функции спектральной плотности реализации акустических сигналов на конечных интервалах, блок-схема программы которого представлена на рис. 3.5. Представленный алгоритм может быть использован для спектрального анализа речевых сигналов, акустических шумов и помех.

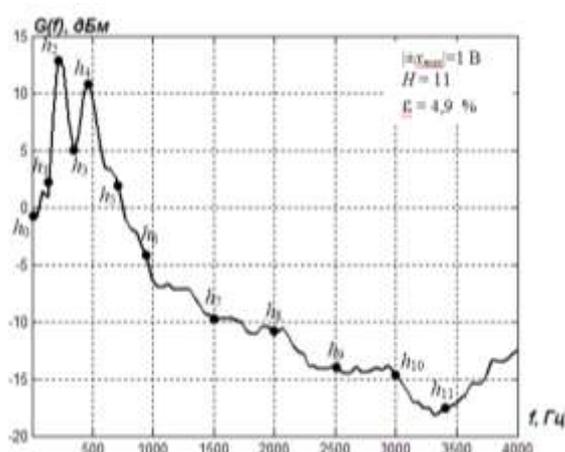
3.5.2. Применение алгоритмов кратковременного дискретного преобразования Фурье для исследования спектра речевых сигналов.

На рисунке 3.13. приведены результаты исследований спектральных

характеристик набора реализаций речевых сигналов с помощью алгоритма 3.7. Для получения результатов приведенных на рисунке 3.13 в пределах одного сегмента исследовались реализации речевых сигналов.



а. Реализация №1



б. Реализация №2

Рисунок 3.13 –Спектральная плотность мощности акустических речевых сигналов при различных реализациях

По результатам, приведенным на рисунке 3.13 можно отметить следующие выводы: спектральная плотность мощности речевой информации начинает понижаться со скоростью около 20 дБ/октава, со значения частоты ~550 Гц, при этом максимум плотности мощности сигнала концентрируется в границах частот 250...900 Гц. Спектральные составляющие выше 3400 Гц понижаются на 30 дБ.

3.5.3. Аппроксимация функций спектральной плотности мощности

Оценки спектра, приведенные в вышеуказанном разделе, являются случайными величинами. В таком случае необходимо применить сглаживание с помощью интерполяционной формулы.

Используя неравномерную сетку узлов, рассмотрены аппроксимации, основанные на кубических сплайнах и интерполяционном многочлене Лагранжа [4, 68, 85]. Многочлен Лагранжа, построенный по 12 узлам имеет вид [4]

$$G_{\text{анп.}}(f) = \sum_{h=0}^{H=11} G(f_h) \prod_{\substack{\ell=0, \\ \ell \neq h}}^{H=11} \frac{f - f_\ell}{f_h - f_\ell}, \quad (3.24)$$

в котором h – порядковый номер узла;

$G(f_h)$ – спектральная составляющая в h -том узле;

f_h – частота в рассматриваемом узле.

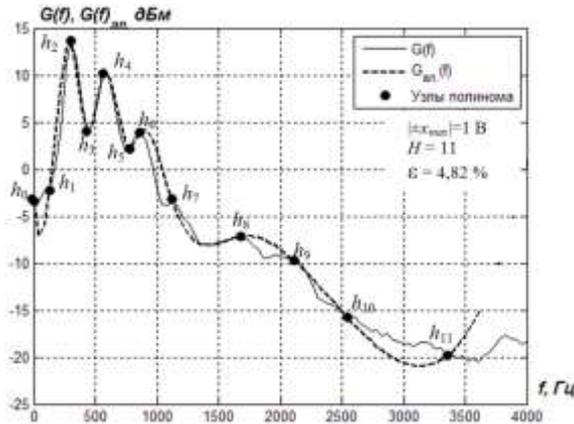


Рисунок 3.14. Аппроксимация функции спектральной плотности мощности с применением полинома Лагранжа

В таком случае для расчета функции $G_{\text{анп.}}(f)$ в виде многочлена Лагранжа 11-ого порядка, представлена сеточная функция в таблице 3.4. Значения из таблицы подставляются в (3.24).

Таблица 3.4

h	0	1	2	3	4	5	6	7	8	9	10	11
$f_h, (ГГц)$	0	180	300	450	510	770	850	1190	1680	2175	2500	3400
$G(f), (dbm)$	-3,8	1,0	13,5	4,5	10	2,2	3,7	-5	-7,5	-10,4	-15	-20

По графику из рисунка 3.14 видно, что погрешность аппроксимации лежит в пределах 5%. Результаты исследований зависимости погрешности от количества узлов аппроксимации H приведены на рисунке 3.15.

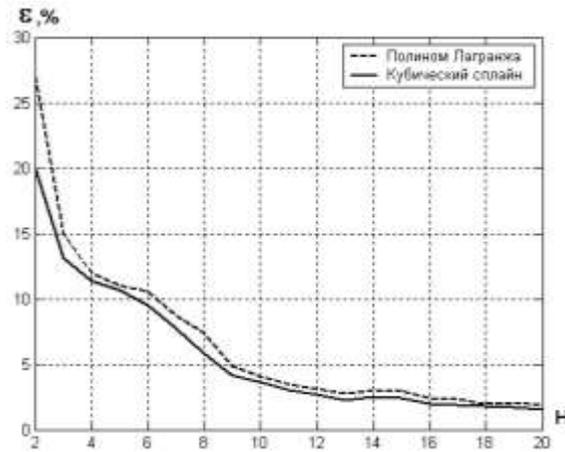


Рисунок 3.15 – Вычисленные погрешности аппроксимации при различном количестве узлов аппроксимации N

По результатам, представленным на рисунке 3.15, можно отметить, что при аппроксимации кубическими сплайнами и полиномом Лагранжа погрешности энергетического спектра речевых сигналов имеют примерно одинаковую величину.

3.6. Исследование спектральных характеристик акустических шумовых помех.

На основании данных из базы временных реализаций сигналов шумовых помех, приведенных в [112, 119] было выполнено исследование спектральной плотности мощности $G_n(f)$ ряда внешних акустических помех.

Для вычисления $G_n(f)$ применялась формула:

$$G_n(l \cdot f_l) = \sum_{l=0}^{N-1} \sum_{k=0}^{K-1} R_n(kT) e^{-j2\pi l \cdot f_l kT}, \quad (3.25)$$

в которой N – количество дискретных отсчетов наблюдаемого интервала, i – параметр, определяющий номер дискретной частотной составляющей $f_l = \frac{F_0}{L}$, $R_n(kT)$ – АКФ акустического шума, либо её аппроксимация.

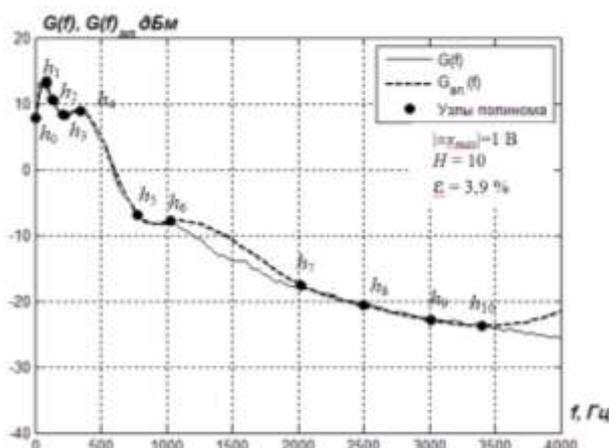
На рисунке 3.16, приведены результаты исследований спектральных функций для двух типов акустических шумовых помех. Сеточная функция для случая акустического шума «сигнал помехи в машинном отделении»,

приведенная в таблице 3.5, составлена с учетом узлов интерполяции представленных на графиках спектральных функций.

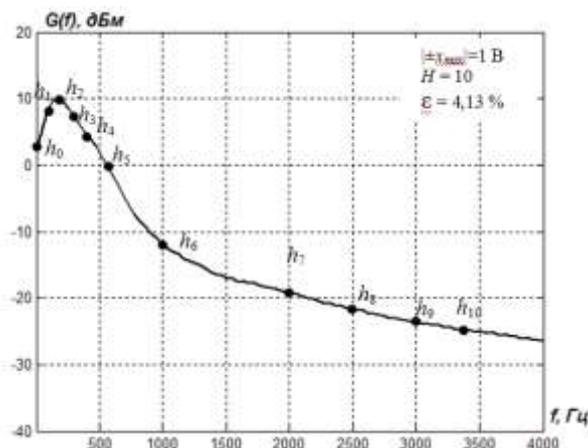
Таблица 3.5

$f_h, (Гц)$	0	110	160	250	370	700	1050	2000	2500	3000	3400
$G(f), (Bm)$	$7 \cdot 10^{-3}$	$20 \cdot 10^{-3}$	$10 \cdot 10^{-3}$	$7,2 \cdot 10^{-3}$	$8 \cdot 10^{-3}$	$0,33 \cdot 10^{-3}$	$0,125 \cdot 10^{-3}$	$1,56 \cdot 10^{-3}$	$0,9 \cdot 10^{-5}$	$0,5 \cdot 10^{-5}$	$0,39 \cdot 10^{-5}$
$G(f), (dbm)$	8,6	13	10	8,7	9	-4	-9	-18	-21	-23	-24

При подстановке в (3.25) значения переменных f_h и $G(f_h)$ из таблицы 3.5, получается аппроксимация с применением интерполяционного полинома Лагранжа спектральной функции процесса «сигнал помехи в машинном отделении». Из рисунка 3.16 видно, что при использовании многочлена Лагранжа десятого порядка (при $H=10$) для аппроксимации вышеприведенных функций, может быть обеспечена погрешность $\varepsilon < 4,13\%$.



а. Сигнал помехи в машинном отделении



б. Сигнал помехи от звука моря

Рисунок 3.16 – Спектральные плотности мощности различных видов шумовых помех

Анализ спектральных характеристик различных видов шумовых помех, представленных на рисунке 3.16, показывает, что составляющие спектральных функций имеют максимумы, которые сосредоточены в области низких частот ограниченной полосой от 0 до 900 Гц. Наиболее это выражено у шума моря, где сосредоточение составляющих с максимальным уровнем

происходит в области частоты 180 Гц и далее спадает со скоростью 30 дБ/октава. При этом максимальный уровень составляющих спектра процесса вида «шум моря» сосредоточен в окрестности частоты 100 Гц и уменьшается в среднем со значением скорости около 10 дБ/октава (рисунок 3.16 б).

Рассмотренная функция энергетического спектра помех от устройств в машинном отделении имеет более пологую и широкую полосу частот, а спектральная функция имеет два максимума: максимум на уровне 14 дБм при частоте 50 Гц, и второй максимум на уровне 9 дБм при частоте 300 Гц. При частоте 1кГц уровень составляющих функции спектральной плотности уменьшается до -22 дБ и ниже (рисунок 3.16 а). На рисунке 3.16 полученные значения составляющих отображены в децибелах относительно уровня 1 мВт, в дБм.

Рисунок 3.16 показывает, что спектр описанных акустических шумовых помех, которые присутствуют, в том числе, на морских судах, имеет частотную область от 0 Гц до 500-1000 Гц, при этом спектр речевых сигналов в основном находится в области от частот 200-500Гц до уровня 3000-3600 Гц.

При выполнении анализа спектральной плотности мощности акустических шумовых помех применялось нормирование отсчетов $x(nT)$ по правилу $|\pm x_{\max}| = 1В$, где x_{\max} – максимальное значение отсчета.

3.7. Выводы

1. Разработан метод восстановления распределения плотности вероятности аудиосигналов и акустических помех с погрешностью не более 5% в виде многочлена третьего порядка по системе экспоненциальных и гауссовых функций соответственно.

2. Применяя модель одномерной функции плотности вероятности возможно рассчитать дисперсию и математическое ожидание, аудиосигналов и акустических помех с допустимой погрешностью.

3. Разработан алгоритм вычисления функции спектральной плотности реализации акустических сигналов на конечных интервалах, который может быть использован для спектрального анализа речевых сигналов, акустических шумов и помех.

ГЛАВА 4 АЛГОРИТМЫ ПОВЫШЕНИЯ ПОМЕХОЗАЩИЩЕННОСТИ АУДИООБМЕНА В ТЕЛЕКОММУНИКАЦИЯХ

4.1. Линейная фильтрации в телекоммуникациях аудиообмена

Модель функции распределения вероятностей акустических сигналов в виде (3.9), позволяет осуществить вычисление математического ожидания процессов \bar{X} и дисперсии σ^2 в соответствии с выражениями [104]:

$$\bar{x} = \sum_{i=-2^{n-1}}^{i=2^{n-1}-1} x_i \cdot P(x_i), \quad \sigma^2 = \sum_{i=-(2^n-1)}^{i=2^n-1} x_i^2 \cdot P(x_i),$$

где x_i – эмпирические данные отсчетов акустических сигналов, с нормированием максимальных значений отсчетов в виде $|x_{max}|=1$ В.

В этом случае, дисперсия нормированного речевого сигнала принимает значение $\sigma^2 = 0,126$ Вт. Относительная интенсивность речевого сигнала тогда, относительно нулевого уровня $I_0=10^{-12}$ Вт/м², определяется как

$$\frac{I}{I_0} (\text{дБ}) = 10 \cdot \lg \frac{k\sigma^2}{SI_0} \cdot 10^{12},$$

где k – коэффициент направленности, $S = 4\pi R^2$ – площадь сферы [104].

Теперь можем вычислить относительную интенсивность акустического речевого сигнала I_p :

$$\frac{I_p}{I_0} (\text{дБ}) = 103\text{дБ},$$

при $\sigma^2 = 0,126$ Вт, $k=2$, $R=1$ м.

В таблице 4.1 сведены значения относительных интенсивностей для различных акустических сигналов, акустических шумов, акустических помех. Для их расчета также применены разработанные в разделах 3.2 и 3.3 математические модели функции плотности вероятностей. В формулах для вычисления относительной интенсивности присутствует среднеквадратичное

отклонение, для получения которого надо знать математическое ожидание.

Таблица 4.1

Вид сигнала	P (Вт)	$\frac{I}{I_0}$ (дБ)	$\frac{P_c}{P_{ш}}$ (дБ)
Речевой сигнал	0,126	103	–
Акустические помехи (ОСТ В4 Г0.005.004)	$3,5 \cdot 10^{-4}$	85	17,8
Шум моря	0,0687	90,4	12,6
Шум ветра	0,0953	91,8	11,2
Шум в машинном отделении	0,111	92,5	10,4

Отношение мощности сигнала к мощности акустической помехи, шума машинного отделения, ветра и моря ограничено пределами от 10,4 до 17,8 дБ, что отражено в таблице 4.1. Это понижает слоговую разборчивость русской речи до 64% (рисунок 2.3), что мешает воспринимать речевую информацию в указанных условиях.

На рисунке 4.1. представлен интерфейс программы вычисления разборчивости формант при различных значениях отношения сигнал / шум.

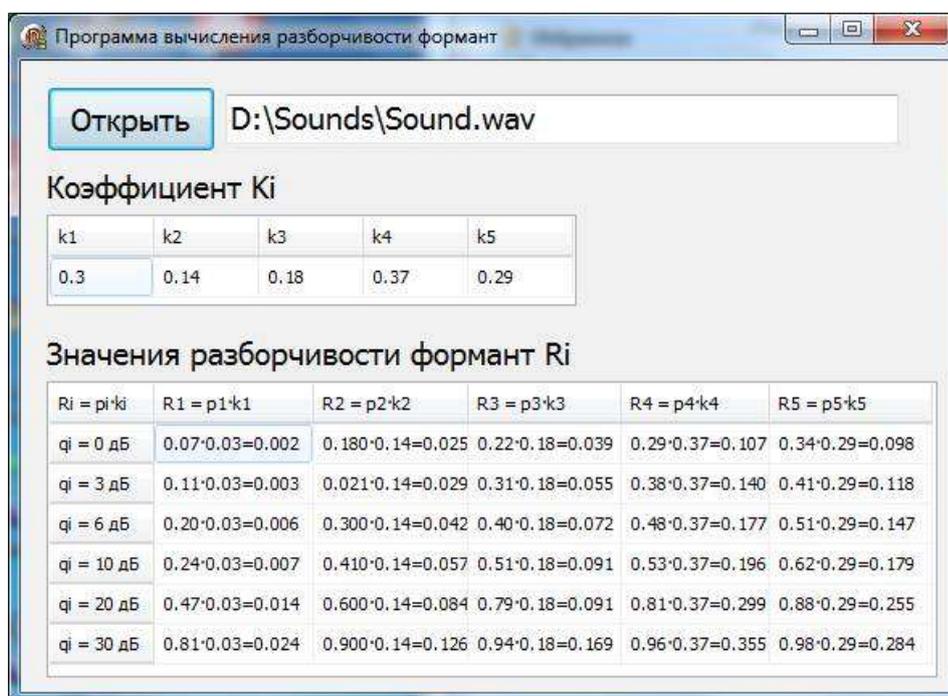
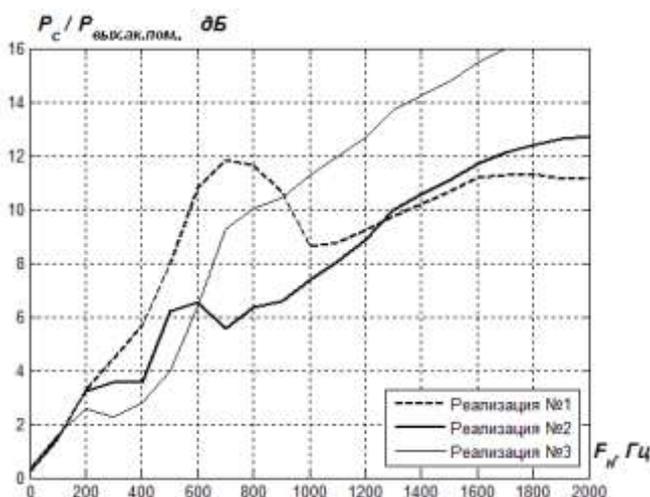


Рисунок 4.1 – Пример работы приложения по вычислению разборчивости формант.

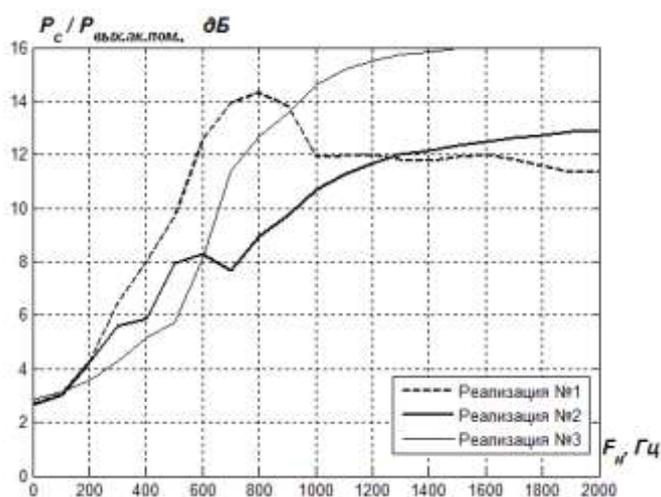
Здесь K_i – весовые коэффициенты, вычисленные для пяти октавных полос, выражающие вероятность нахождения формант речевых сообщений в i -той частотной полосе, q_i – отношение сигнал/шум в i -той формантной частотной полосе; p_i – коэффициенты восприятия формант для i -той частотной полосы при изменении ОСШ. Среда программирования Delphi, язык программирования Object Pascal.

В ходе проведения исследований выяснили, что значение $P_c / P_{ак.ном.}$, полученное на выходах трактов передачи системы телекоммуникаций аудиообмена, должно составлять 20 дБ и более, чтобы получить необходимое для полного понимания и восприятия оперативно-командной речевой информации абонентами системы телекоммуникаций, в условиях существенного воздействия внешних акустических шумов и помех значение слоговой разборчивости [104]. Поэтому нужны алгоритмы эффективного подавления этих помех до получения отношения $P_c / P_{ак.ном.} \geq 20$ дБ.

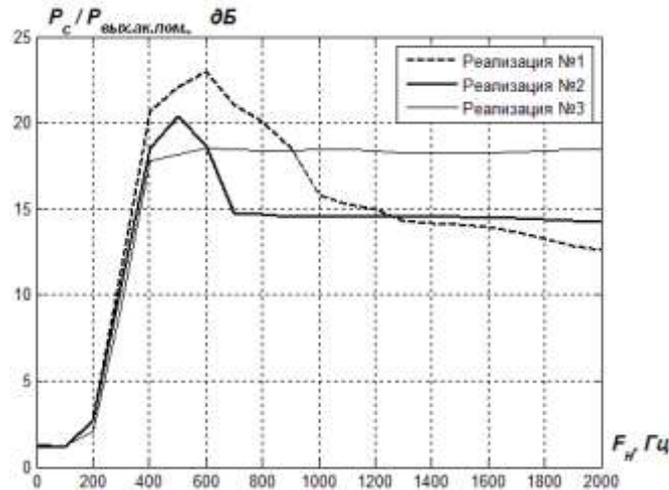
Для различных видов акустических внешних шумов на выходе фильтра с амплитудно частотной характеристикой зависимость отношения сигнал/ шум от нижней частоты спектра приведена на рис. 4.2.



а. Шум в машинном отделении



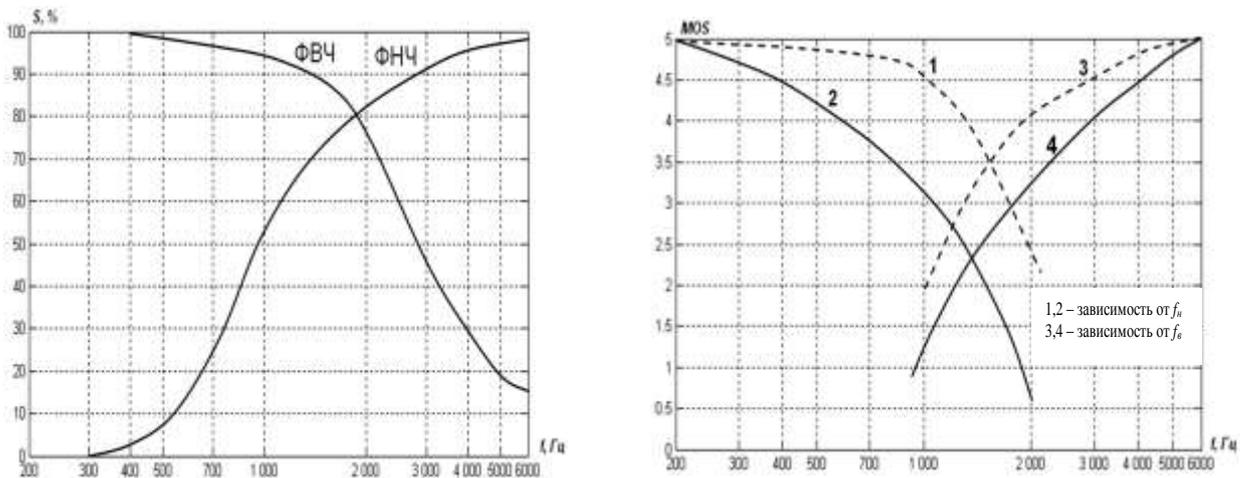
б. Шум моря



в. Шум ветра

Рисунок 4.2 – Зависимость отношения $P_c / P_{\text{вых.ак.пом.}}$.

В ходе проведенных исследований в разделах 3.5 и 3.6 получили, что спектр шумовых помех занимает более низкочастотную часть спектра, что позволяет применить для повышения отношения $P_c / P_{\text{ак.пом.}}$ методы линейной фильтрации. Или можно увеличить область режекции от $F=0$ Гц, до $F_n=300-900$ Гц в полосе канала передачи речевого сигнала.



а)

б)

Рисунок 4.3 – Зависимость оценки: а) слоговой разборчивости от ограничения полосы воспроизводимых частот, б) зависимость усредненной оценки разборчивости речи от характеристик полосы пропускания.

Из графика на рис. 4.3 а видно, что величина слоговой разборчивости S

почти не меняется при увеличении области режекции до 900 Гц, и остается более 93%, что допустимо [120,121,122].

Исходя из полученных результатов, предполагаем возможность управления областью режекции в полосе от 0 до (300÷900) Гц для подавления акустических помех с учетом особенностей шума, применяя адаптивные алгоритмы обработки речевых сигналов.

4.2. Метод корреляционно-экстремального оценивания параметров эхосигналов

Чтобы увеличить скорость работы системы выделим часть параметров, которые не меняются или мало меняются в процессе обработки. К таким долговременным параметрам можно отнести $\hat{\tau}_k = \hat{D}_k T$ – параметр запаздывания и \hat{a}_k – параметр затухания. Параметры затухания зависят от размеров и акустических характеристик помещения и мало изменяются во времени.

При нормировании входного сигнала адаптивного фильтра $u(n)$, используется параметр затухания \hat{a}_k , а параметр запаздывания $\hat{\tau}_k$ позволяет сформировать входной сигнал коррелятора $u(n)$ из $u_g(n)$ (см. рисунок 4.4).

На рисунке 4.4 приведена схема принципа оценивания вышеуказанных параметров.

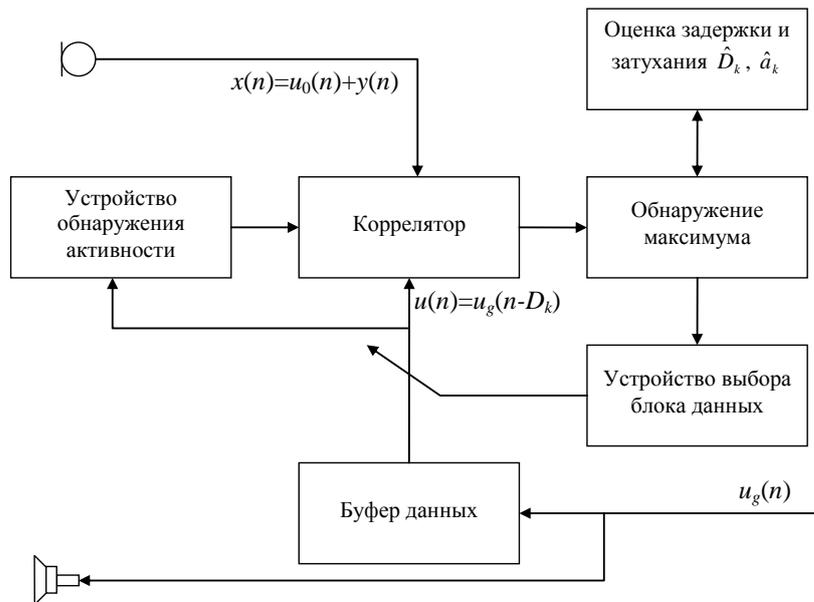


Рисунок 4.4 – Схема оценивания долговременных параметров эха

Основные компоненты системы (рисунок 4.4): устройство выбора блока данных и обнаружитель максимумов функций, представляющих собой систему регулирования. Основной задачей системы является поиск локальных максимумов выхода коррелятора, параметр этой системы: задержка D опорного сигнала, выбираемого из буфера данных. Оценка задержки и затухания – последний шаг алгоритма. Затем эти оценки затуханий используются в адаптивных фильтрах.

Оценить рассмотренные параметры можно на этапе ввода системы в эксплуатацию с применением тестовых сигналов. Далее осуществляется систематический пересчет значение параметров на основе поступающих реальных сигналов, что обеспечивает повышение точности оценивания и надежности функционирования системы.

Моделирование в Matlab выполнялось в соответствии со следующим алгоритмом.

1. Нахождение значимых каналов эха

$$\hat{D}_k T = \underset{DT \in [0, KT]}{\operatorname{argloc}} \max R_{u_g, x}(DT), \quad k = 1, 2, \dots, \quad (4.1)$$

в котором k – номер локального максимума; T – период дискретизации;

\hat{D}_k – оценка величины задержки; $x(n)$ – выходной сигнал с микрофона; $u_g(n)$

– значение опорного сигнала источника эха; $[0, KT]$ – интервал потенциальных значений задержек эха.

2. ВКФ $R_{u_g, x}(DT)$ в дискретном времени

$$R_{u_g, x}(DT) = \frac{1}{N+1} \sum_{n=0}^N x(n+D)u_g(n), \quad (4.2)$$

где $[0, NT]$ – временное окно, на котором определен сигнал $u_g(n)$, $[0, NT] \subset [0, KT]$ или $N \ll K$, т.е. временное окно более чем на порядок меньше величины максимальной задержки эха.

3. ВКФ сигналов $u_g(n)$ и $x(n)$, определенных на интервалах $[0, NT]$ и $[\tilde{D}T, (N + \tilde{D})T]$,

$$R_{u_g, x}(\tilde{D}T + dT) = \frac{1}{N+1} \sum_{n=0}^N x(n - \tilde{D} + d)u_g(n), \quad (4.3)$$

в которой $|d| < N$. Затем становится возможным расчет $d = \hat{d}$,

который при заданном смещении D позволит получить максимальное значение функции (4.3). Таким образом оценивание задержки вычисляется следующей формулой

$$\hat{D}_k = \tilde{D}_k + \hat{d}_k, k = 1, 2, \dots, \quad (4.4)$$

в которой k – номер локального максимума; \tilde{D} – номер интервала, кратный целому числу, меньшему N .

4. Расчет оценки затуханий по выражению

$$\hat{a}_k = \frac{R_{u_g, x}(\hat{D}_k T)}{R_{u_g, u_g}(0)}, k = 1, 2, \dots \quad (4.5)$$

Здесь $R_{u_g, u_g}(0)$ – автоковариационная функция сигнала $u_g(n)$ на интервале $[0, NT]$.

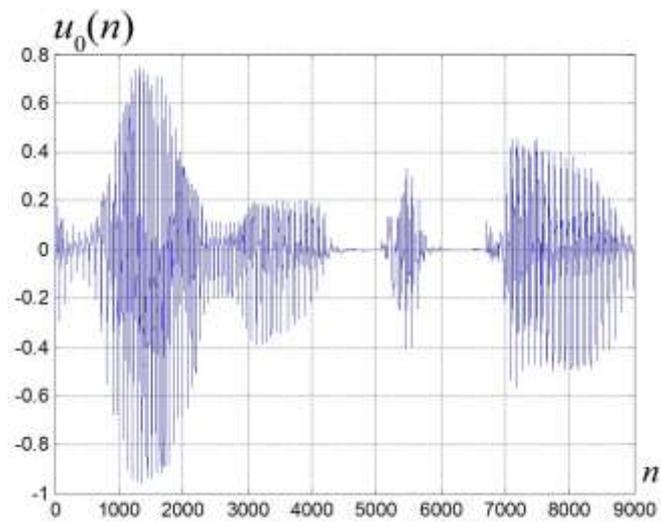


Рисунок 4.5 – Сегмент речевого сигнала

Рисунок 4.5 демонстрирует сегмент речевого сигнала не зашумленного помехами и компонентами эхо.

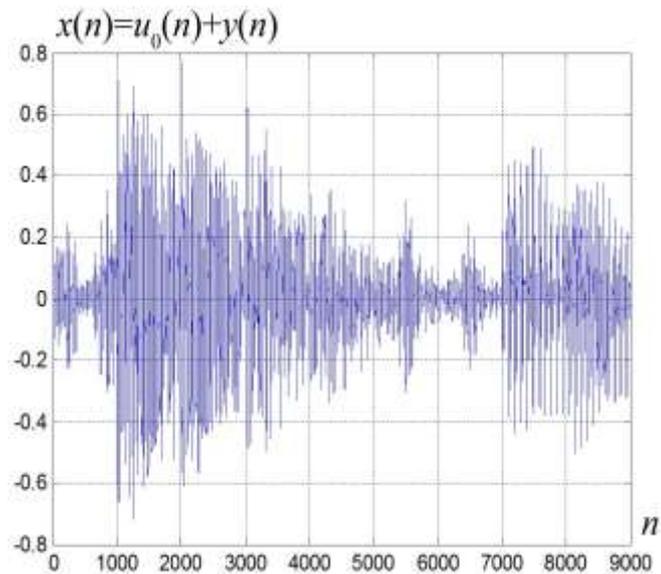


Рисунок 4.6 – Сегмент речевого сигнала на фоне эха

На рисунке 4.6 приведен сегмент сигнала с наложенными на него со сдвигом копиями, имитирующими эхосигналы,

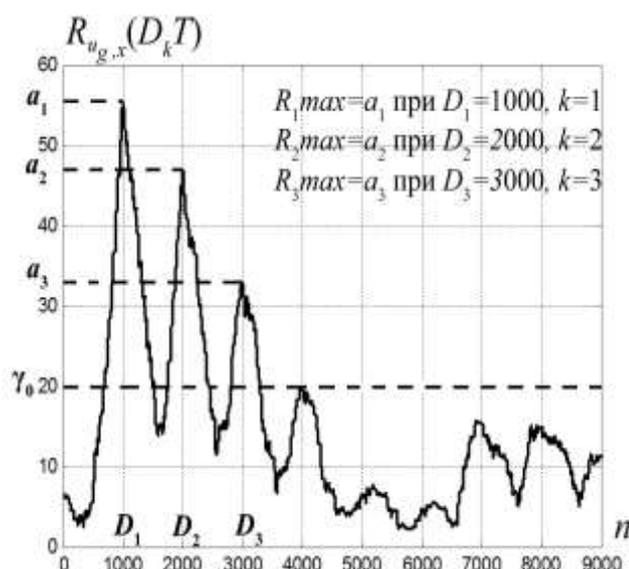


Рисунок 4.7 – Оценка влияния максимального значения ковариационной функции от сдвига интервала опорного речевого сигнала

Рисунок 4.7 демонстрирует результаты вычисления ковариационной функции модели речевого сигнала с сегментами эмуляции эхосигнала.

Проведенное моделирование подтверждает возможность использования метода корреляционно-экстремального оценивания долговременных параметров отраженных эхосигналов.

4.3. Адаптивная компенсация эхосигналов в системе телекоммуникаций громкоговорящей связи.

В свободном пространстве эхосигналы образуются в результате отражений акустических сигналов от образований поверхности, обусловленных особенностями рельефа местности. Значительные размеры территории приводят к большим задержкам распространения эхосигналов. Импульсные функции в каналах распространения эха можно при этом принять за некоторые постоянные коэффициенты передачи. Модель множественных отражений, принимает вид [97]

$$v(t) = \sum_{k=1}^r a_k u(t - \tau_k) = \sum_{k=1}^r a_k u(t - D_k T) \quad (4.6)$$

Задача при этом заключается в определении параметров затухания a_k и

запаздывания $\tau_k = D_k T$ для r каналов эхосигналов [101].

Сигнал $u(t)$, многократные отражения которого формируют сигнал $v(t)$, является известным и нестационарным. Это позволяет применить при оценивании указанных параметров технику корреляционного анализа.

Структура системы обмена информацией с компенсационным каналом эха, представлена на рисунке 4.8 [95].

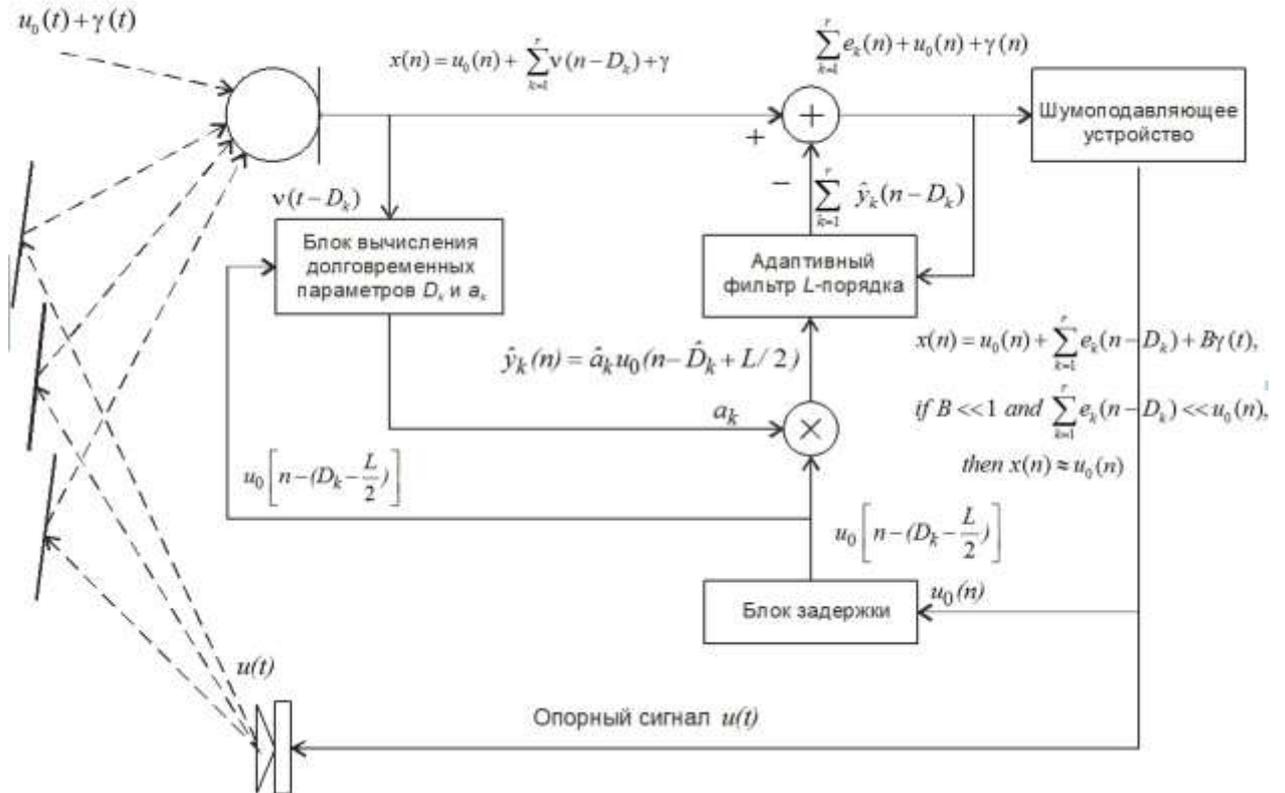


Рисунок 4.8 – Структура абонентского устройства системы связи

Поступающий на вход сигнал $x(t)$ представляет собой смесь множественных отражений $u(t)$, акустического шума $\gamma(t)$ и речевого сигнала $u_0(t)$.

$$x(t) = u_0(t) + \sum_{k=1}^r a_k u(t - D_k) + \gamma(t).$$

В блоке вычисления долговременных параметров, работу которого рассматривали в разделе 4.2, вычисляются D_k и a_k , с применением подхода корреляционного оценивания по (4.2) и (4.5) при $g = u_0$. Вычисляется опорный сигнал, для адаптивного фильтра L -того порядка.

Также в соответствии с рисунком 4.8 в структуре обмена информации имеется шумоподавляющее устройство, реализующее алгоритм адаптивного подавления акустических помех с формантным распределением полос режекции, адаптивный фильтр компенсации эха, блок идентификации долговременных параметров эха - a_k и D_k , блок формирования опорного сигнала адаптивного фильтра и сумматор. Опорный сигнал в этом случае представлен в виде формулы

$$\hat{y}_k(n) = \sum_{k=1}^r \hat{a}_k u_0(n - \hat{D}_k).$$

Опорный сигнал и эхо на выходе сумматора можно описать следующим выражением

$$u_{\text{вых}}(n) = \sum_{k=1}^r e_k(n) + u_0(n) + \gamma(n),$$

где $\sum_{k=1}^r e_k(n) = \sum_{k=1}^r v(n - D_k) - \sum_{k=1}^r \hat{y}_k(n - \hat{D}_k).$

На рисунке 4.8, в котором отражена структура обмена информацией, выходной сигнал имеет вид

$$x(n) = u_0(n) + \sum_{k=1}^r e_k(n - D_k) + B\gamma(t),$$

при условии

$$\sum_{k=1}^r e_k(n - D_k) \ll u_0(n),$$

где B меньше единицы, т.е. $B \ll 1$, поэтому сигнал на выходе имеет вид

$$x(n) \approx u_0(n).$$

В отличие от открытой территории, эхосигнал в замкнутых пространствах следует считать интегральной суммой компонентов, непрерывно распределенных по величине задержки. Достаточно хорошим приближением к действительности может явиться использование модели эхосигнала. Входящие в эти выражения импульсные функции $f(\theta)$ имитируют, при этом, непрерывное распределение задержек в пределах

относительно небольших окрестностей $[\tau_k - \varepsilon, \tau_k + \varepsilon]$, $k = 1, \dots, r$, выделенных главных каналов эха.

При моделировании системы компенсации (рисунок 4.8) вычислялся уровень подавления эхосигнала и помехи. Фильтром с равномерной АЧХ в диапазоне 300-600 Гц и спадом на 20 дБ в диапазоне 600-3400 Гц моделировалась акустическая помеха 80 дБ. Модель каскадных отражений с запаздываниями в каналах 100, 75, 50 и 25 мс формировала эхосигнал. Коэффициент затухания в таких каналах имеет значение обратное запаздыванию, т.е. 0.5, 0.25, 0.125 и 0.075. Также, в дополнение к этому, вносились дополнительные искажения по амплитуде от 5 до -10 дБ в эхосигнал основного канала и помеху.

Дополнительные фазовые искажения в основном канале находились в пределах от -90 до 90 градусов. Были использованы речевые сигналы с повторяющейся записью реального диалога. Для анализа сигнала дополнительного и основного каналов была использована оценка параметров модели эхосигнала.

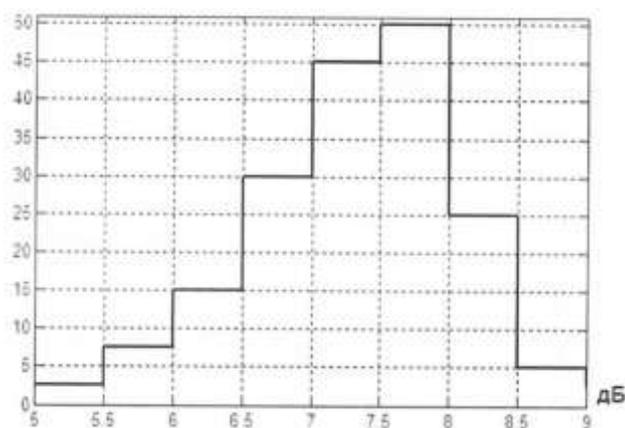


Рисунок 4.9 – Гистограмма уровня компенсации эхосигналов и помех

Для оценки уровня компенсации эхосигнала и помех был взят интервал длительностью 1 секунда. По результатам, приведенным на рисунке 4.9 видно, что уровень подавления в большинстве случаев находится в пределах от 6 до 8,2 дБ.

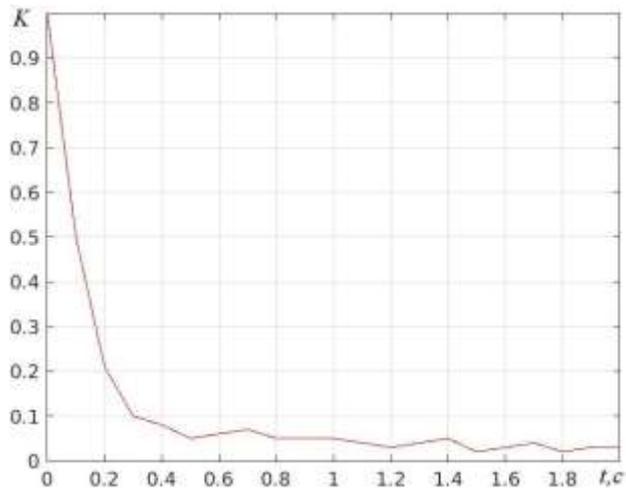


Рисунок 4.10 – Динамика подстройки компенсатора эхосигнала и помехи

На рисунке 4.10 показана динамика подстройки компенсатора помех. Получено среднее время подстройки 0,2..0,4 секунды. Оно зависит от соотношения и уровня эхо-сигналов и стационарных акустических помех. Состояние компенсатора может также оказывать влияние на величину времени подстройки в стартовый момент инициализации. Несмотря на это, полученное время в несколько раз меньше, чем в системах без оценивания долговременных параметров эхосигнала.

Использование многоканальной обработки сигналов повышает точность и упрощает алгоритм оценивания указанных параметров.

Результаты проведенного модельного эксперимента по времени настройки адаптивного фильтра показали повышение скорости подстройки компенсатора.

На рисунке 4.11 приведена зависимость времени настройки от числа настраиваемых коэффициентов адаптивного фильтра. Установлено, что число отводов существенно влияет на скорость настройки адаптивного фильтра.

За счет существенного уменьшения числа настраиваемых коэффициентов удалось достичь в четыре раза большую скорость настройки фильтров в системе с оцениванием долговременных параметров.

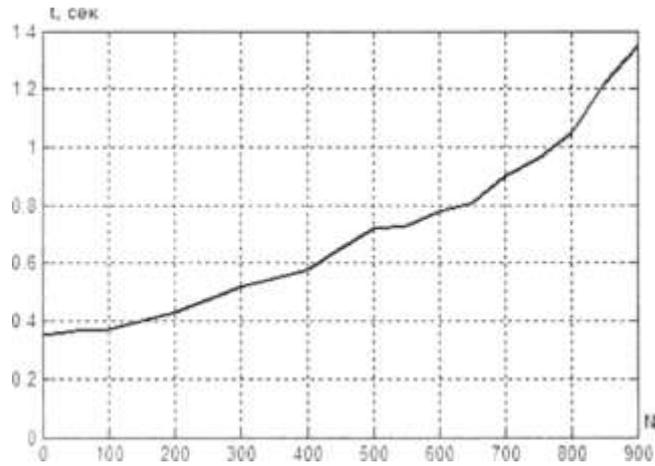


Рисунок 4.11 – Зависимость количества настраиваемых коэффициентов от времени настройки адаптивного фильтра

В результате, уровень компенсации эхосигнала и помех составил примерно 7 дБ, что более чем на 2 дБ лучше известных. Кроме того, скорость подстройки адаптивных фильтров повысилась в 3-4 раза.

4.4. Создание структуры алгоритма подавления аддитивных акустических помех

С применением гребенки фильтров, имеющих полосу пропускания Δf можно представить частотное распределение формантных полос спектра акустического речевого сигнала $F_n=300$ Гц до $F_v=3400$ [11, 73]. Таким образом, если в полосе пропускания регистрируются составляющие выше порога, с помощью приведенной структуры становится возможным подавление помех через управление затухания в фильтрах.

На рис 4.12 продемонстрировано поступление входного информационного сигнала, который можно представить случайным процессом $x(t)$ с дисперсией σ^2 и дальнейшая обработка в многоканальном тракте передачи.

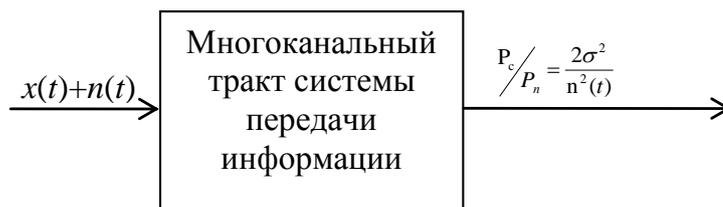


Рисунок 4.12 – Система передачи информации

Также на вход системы поступает сосредоточенные помехи в виде гармонических колебаний с амплитудой $n(t)$ и частотой f_n [2]. Выходное отношение сигнал-помеха определено выражением

$$\frac{P_c}{P_n} = \frac{2\sigma^2}{n^2(t)}. \quad (4.7)$$

Амплитудно-частотная характеристика канала передачи определяется в соответствии с [46]

$$k(f) = \begin{cases} 1, & f_1 \leq f \leq f_2, \\ 0, & \text{при других } f, \end{cases}$$

при этом ширина полосы пропускания канала определяется разностью

$$\Delta f = f_1 - f_2$$

Чтобы вычислить мощность $x(t)$, имеющую место на выходе канала передачи возможно применение выражения вида:

$$P_\Delta = \int_{f_1}^{f_2} k^2(f)G(f)df \approx \Delta f \cdot G'(f),$$

в котором $G(f)$ – преобразование Фурье от автокорреляционной функции процесса $x(t)$, $G'(f)$ – постоянная (при $\Delta f \ll \Delta F$, $\Delta F = F_g - F_n$) спектральная плотность мощности процесса $x(t)$ в пределах от f_1 до f_2 , ΔF – ширина полосы пропускания.

Структура алгоритма расчета отношения сигнал-помеха на выходной линии узкополосного канала приведена на рис. 4.13

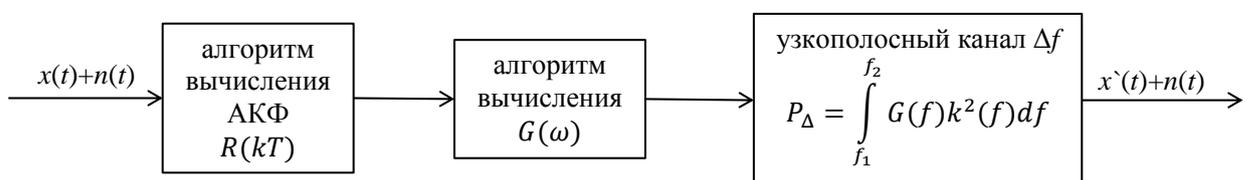


Рисунок 4.13 – Структура алгоритма расчета отношения сигнал-помеха на выходной линии

Если принять с допустимой для технических решений погрешностью, что $G(f)$ в пределах от F_n до F_g – постоянная величина, то

$$G(f) = \frac{\sigma^2}{\Delta F}, \quad P_{\Delta} = \Delta f \cdot G(f) = \frac{\Delta f \sigma^2}{\Delta F}.$$

В таком случае на выходе узкополосного канала отношение $\frac{P_c}{P_n}$ уменьшается в $\frac{\Delta f}{\Delta F}$ раз к входному отношению $\frac{P_c}{P_n}$, то есть

$$\frac{P_c}{P_n} \text{ в вых. узк. кан.} = \frac{\Delta f}{\Delta F} \cdot \frac{2\sigma^2}{n^2(t)}.$$

Значение коэффициента подавления сигналов $k_{подавл.}$, находящихся вне полосы Δf для сосредоточенной помехи определяется технической реализацией узкополосного фильтра и принимает значение $k_{подавл.} \ll 1$. Это вызвано вследствие подчинения условиям $f_1 > f_n > f_2$, а также тем, что в полосу прозрачности сосредоточенной помехи не попадает средняя частота f_n .

При этом, возникновение области режекции в амплитудно-частотной характеристике всего канала понижает полезный сигнал на величину P_{Δ} . Потери определяются следующим коэффициентом в выражении

$$K_{\Pi} = 1 - \frac{P_{\Delta}}{\sigma^2} = 1 - \frac{\Delta f}{\Delta F}.$$

На рисунке 4.14 представлен график зависимости потерь от полосы области режекции относительно полосы пропускания всего тракта. Как видно из рисунка 4.14., если все параметры соответствуют условию $\Delta f \ll \Delta F$, то потери могут составлять доли децибела [37].

Таким образом, выключение канала, пораженного сосредоточенной помехой, приводит к тому, что отношение $\frac{P_c}{P_n}$ определяется выражением [43]

$$\frac{P_c}{P_n} \text{ в вых.} = \frac{2\sigma^2}{k_{подавл.} n^2(t)} \cdot K_{\Pi} = \frac{2\sigma^2}{k_{подавл.} n^2(t)} \cdot \left(1 - \frac{\Delta f}{\Delta F}\right). \quad (4.8)$$

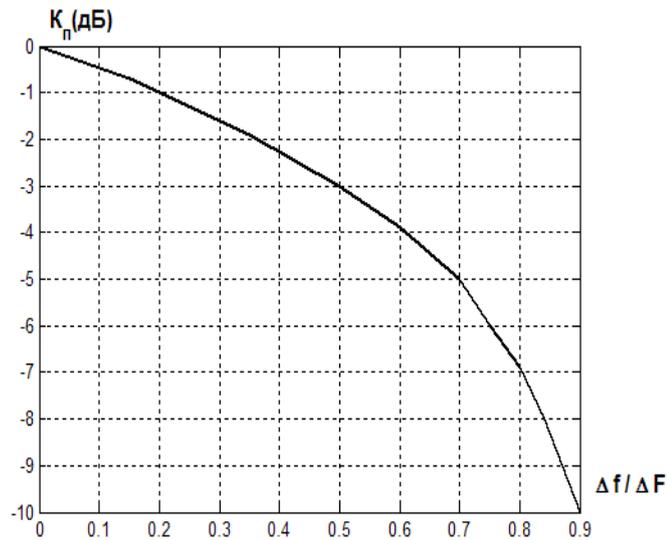


Рисунок 4.14 – График зависимости потерь от отношения $\Delta f / \Delta F$

Например, накладывание узкополосных фильтров на исходный речевой сигнал в канале с коэффициентом подавления вне полосы пропускания $k_{подавл.} = -30$ дБ и при нормировании максимальных отсчетов по правилу $|x_{max}|=1$, $|n(t)|=1$, дисперсии процесса $\sigma^2 = 0,126$ Вт, мощности сосредоточенной помехи $\frac{n^2(t)}{2} = 0,5$ Вт, приводит к получению входных и выходных отношений $\frac{P_c}{P_n}$ в виде [2]

$$\frac{P_c}{P_n}_{вх.} = 10 \cdot \lg \frac{2\sigma^2}{n^2(t)} = -6 \text{ дБ}, \quad (4.9)$$

$$\frac{P_c}{P_n}_{вых.} = -6 \text{ дБ} + 30 \text{ дБ} - 0.5 \text{ дБ} = 23.5 \text{ дБ}. \quad (4.10)$$

Согласно представленным выше результатам, если на систему происходит воздействие сосредоточенной помехи, то это может привести к потерям полезной информации и понизить отношение сигнал-помеха до -6 дБ.

Вышеуказанную проблему можно устранить, применяя многоканальную гребенку фильтров с адаптивным управлением в канале передачи речи. Учитывая технические особенности реализации полосовых фильтров возможно повысить отношение сигнал-шум с помощью автоматического отключения пораженных помехами каналов передачи [37,

43,45].

Алгоритм формирования управляющего сигнала (алг.2) для каждого канала представлен на рисунке

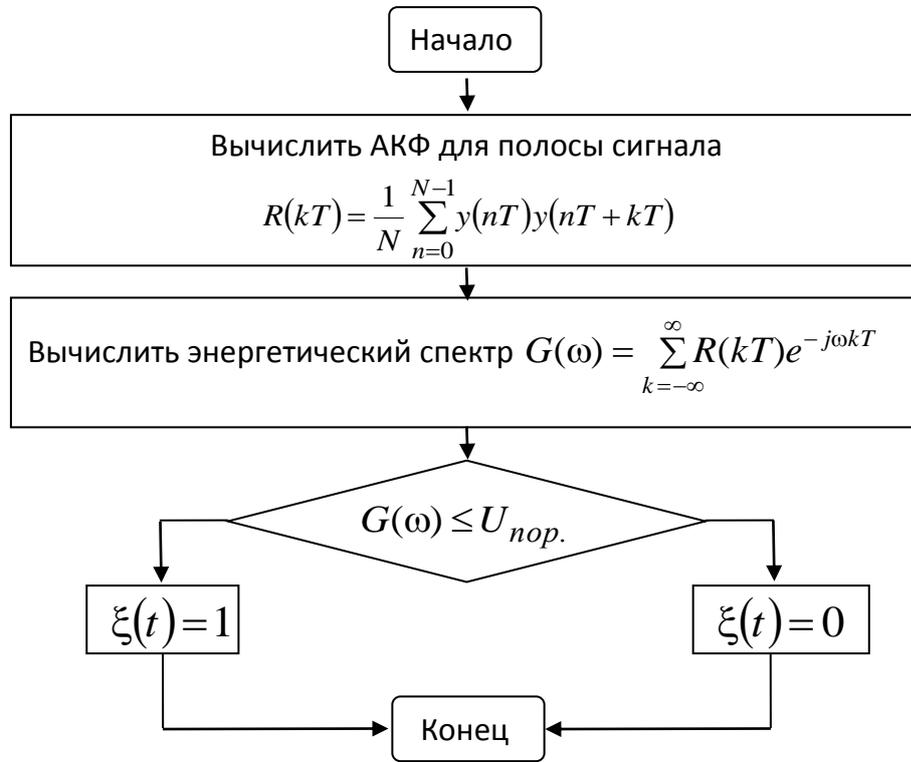


Рисунок 4.15 – Алгоритм формирования управляющего сигнала для канала

Как видно из рисунка 4.15 алгоритм формирования систем управления состоит из действий по вычислению АКФ $R(kT)$ процесса $x(t) + n(t)$, блока вычисления функции энергетического спектра процесса путем формирования преобразования Фурье от АКФ, алгоритма порогового устройства – блока вычисления сигнала управления выключением канала, порешенного сосредоточенной акустической помехой, и блоков вычисления значения порога, для подачи его на пороговое устройство.

В соответствии с рисунком 4.15 сигнал управления формируется по следующей схеме: узкополосный канала с полосой Δf вычисляет уровень $x'(t)$ на его выходе путем интегрирования спектральной функции сигнала в пределах полосы пропускания канала [37].

Теперь рассмотрим алгоритм вычисления порогового сигнала $U_{nop.}$ для порогового устройства (рисунок 4.16), путем вычисления значения σ^2 с

помощью модели функции распределения плотности вероятности, с последующим вычислением значения $U_{пор.} = k\sigma$, где $k = 0,5 \div 1$.

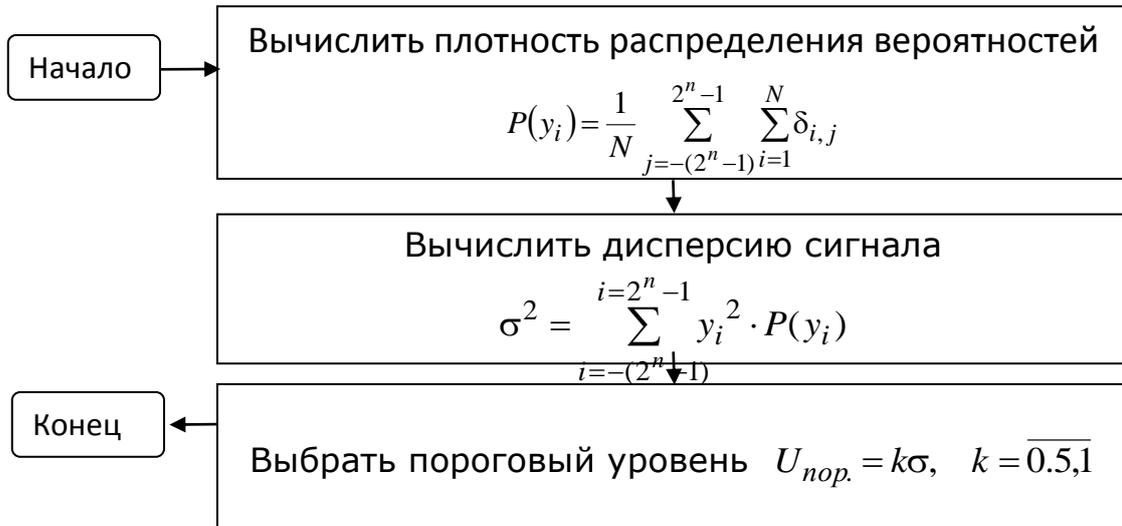


Рисунок 4.16 – Алгоритм вычисления порогового уровня.

Вычисленное значение порога $U_{пор.}$ подается на первый вход порогового устройства. Второй вход порогового устройства принимает сигнал $x'(t) + n(t)$, а на его выходе выдаются сигналы управления $\xi(t)$, формирование которых происходит по правилу

$$\xi(t) = \begin{cases} 0, & x'(t) + n(t) > U_{пор.} \\ 1, & x'(t) + n(t) \leq U_{пор.} \end{cases}$$

На рисунке 4.17 представлена блок-схема программы реализации алгоритма вычисления дисперсии последовательности $x(k)$.

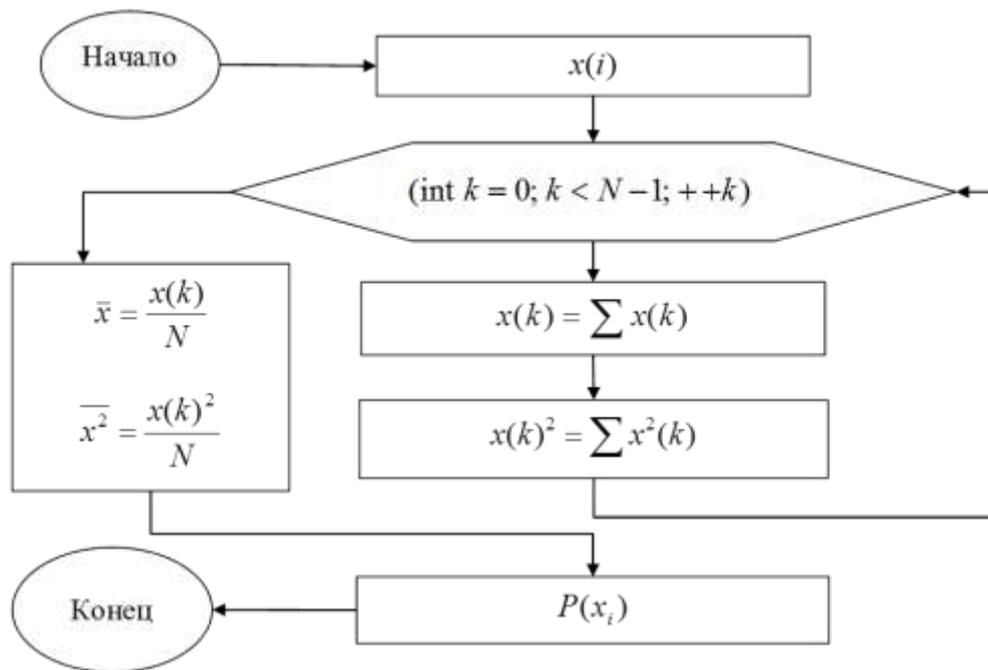


Рисунок 4.17 – Структура блок-схемы программы реализации алгоритма вычисления дисперсии последовательности $x(k)$

Для апробации разработанных алгоритмов было разработано программное приложение, одна из форм которого приведена на рисунке 4.19

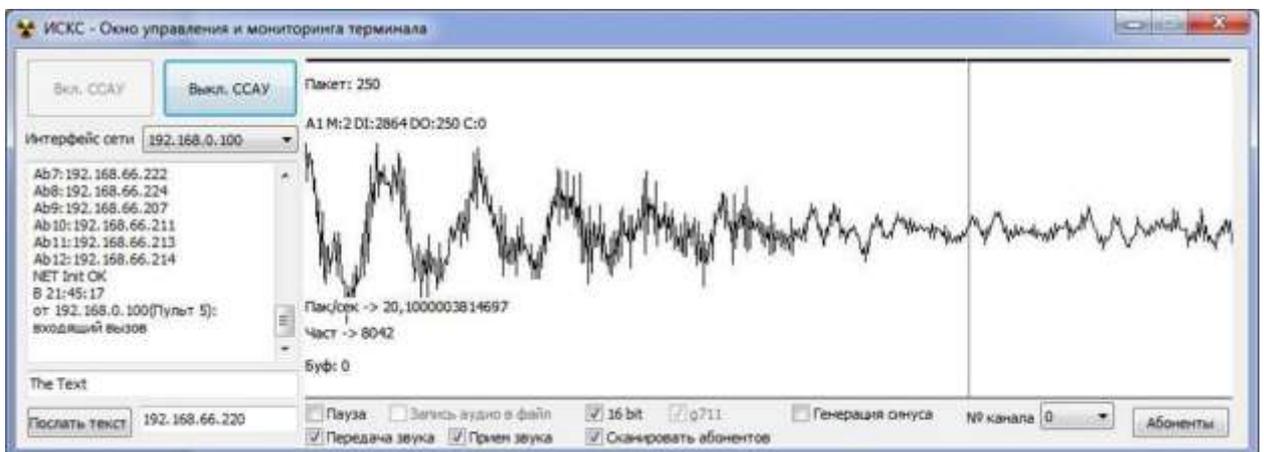


Рисунок 4.19 – Интерфейс модуля программного обеспечения.

На рисунке 4.19 приведен пример вывода визуализации речевой информации, полученной по локальной сети. После распаковки принятого пакета речевая информация проходит обработку в модуле шумоподавления и эхокомпенсации, после чего поступает на звуковыводящее устройство.

4.5. Разработка адаптивного многоканального алгоритма подавления аддитивных и сосредоточенных акустических помех

Из представленных данных по частотному распределению формантных полос речи можно отметить, что гребенка из 16 фильтров с полосой пропускания 200 Гц позволяет получить достаточное приближение. При этом гребенка реализуется с помощью прямоугольной формы амплитудно-частотной характеристики в диапазоне от 300...3500.

Реализация такой системы позволяет более точно подавлять акустические помехи.

На рисунке 4.20 представлена структурная схема устройства подавления сосредоточенных помех, реализующая адаптивный многоканальный алгоритм передачи речи с управлением области режекции [88, 127].

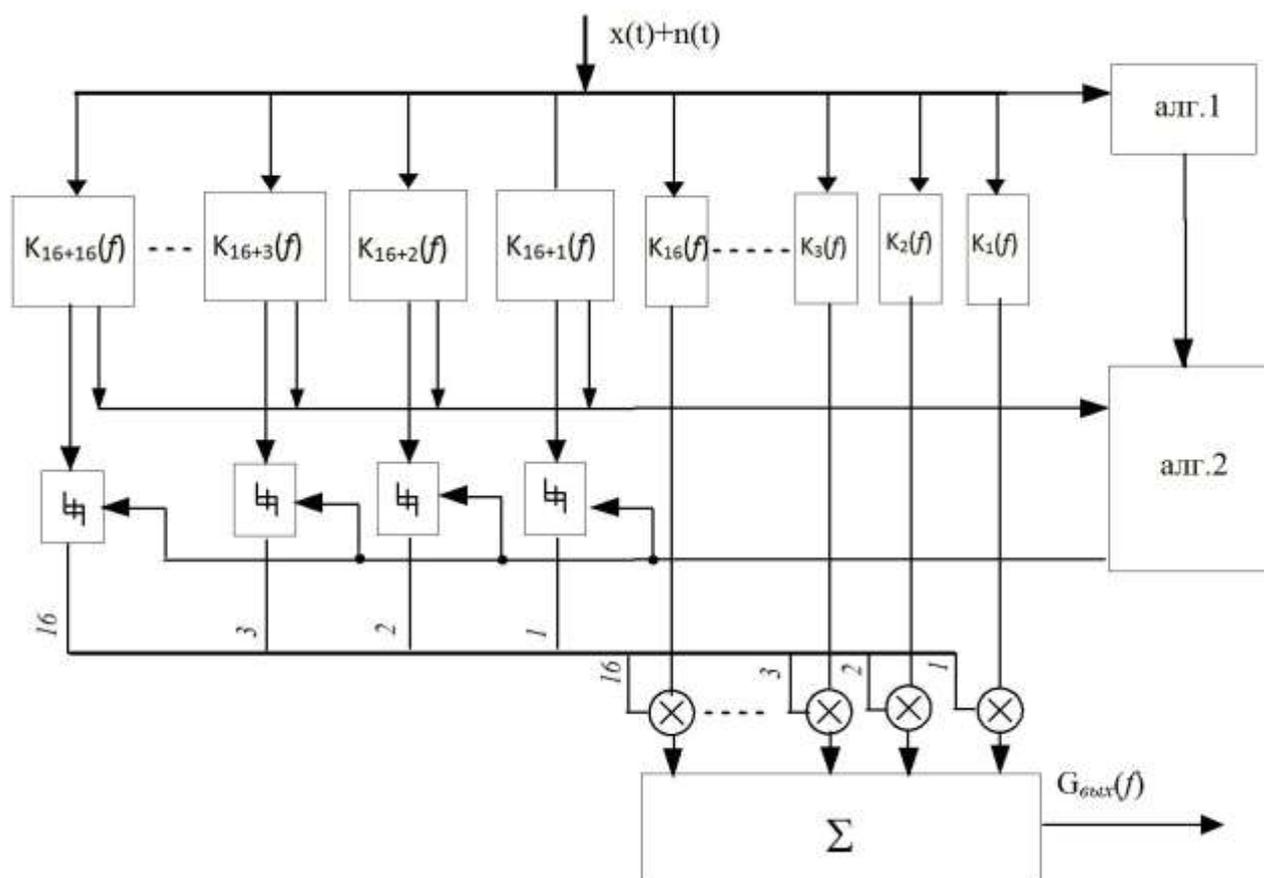


Рисунок 4.20 – Схема подавления помех адаптивно коммутируемой системой узкополосных фильтров.

Из рисунка 4.20 видно, что в состав разработанного устройства входят

основные каналы, которые осуществляют формирование общего фильтра прямоугольной формы. При поступлении соответствующего сигнала управления имеется возможность отключения одного или нескольких каналов в области режекции [41]. Алгоритмы алг.1 и алг.2, представленные на схеме разбирались в разделе 4ю5 и представлены на рисунках 4.15 и 4.16.

Для формирования адаптивного управляющего сигнала областью режекции многоканальной системы используется энергетическая соизмеримость акустических помех с речевыми сигналами. При этом, в силу большей концентрации акустических помех в спектральных составляющих по отношению к речевым сигналам можно разрабатывать алгоритмы амплитудной селекции для адаптивного управления многоканальной фильтрации [3].

При этом, в рассматриваемом методе применяется расчет порогового уровня по правилу

$$\sigma \leq U_{nop.} \leq |x_{\max}|.$$

Также пороговый уровень $U_{nop.}$ в пороговом устройстве может быть сформирован с учетом значения дисперсии речевого сигнала и выбран по следующему правилу:

$$\left[\sum_{i=-(2^n-1)}^{2^n-1} x_i^2 \cdot P(x_i) \right]^{\frac{1}{2}} \leq U_{nop.} < |\pm x_{\max}|,$$

где $P(x_i)$ - закон распределения вероятности отсчетов речевого сигнала.

Вычисление дисперсии σ^2 можно получить по выражению $\sigma^2 = \sum x_i^2 P(x_i)$ и с применением плотности вероятности $P(x_i)$, где x_i – дискретные интервалы речи, квантованные на 2^n интервалах и нормированные относительно максимальных значений в виде $|x_{\max}| = 1B$. В этом случае вычисленное значение $\sigma^2 = 0,126$ Вт, $U_{nop.} = 0,355 \div 1$.

При рассмотрении структуры адаптивного устройства шумоподавления, приведенного рисунком 4.18, вычисление получения

функции спектральной плотности мощности возможно с применением выражения следующего вида

$$G_{вых}(f) = \sum_{n=1}^{16} k_n^2(f) G_{вх}(f)$$

в котором $G_{вх}(f)$ - является функцией спектральной плотности мощности рассматриваемого речевого сигнала, $k_n(f)$ - амплитудно-частотная характеристика n -го фильтра, включенного в состав гребенки.

Амплитудно-частотная характеристика (АЧХ) рассматриваемого единичного фильтра может быть записана в следующем виде

$$k_n(f) = \begin{cases} 1, & n\Delta f \leq f \leq (n+1)\Delta f \\ 0, & n\Delta f > f > (n+1)\Delta f \end{cases}$$

в котором n - является номером фильтра, причем, $n \in Z$ и $1 \leq n \leq N$.

Рассматривая диаграммы распределения мощностей, имеющих место на выходных линиях применяемых единичных фильтров (приведенных на рисунке 4.21 «а», «б»), при условии поступления на устройство акустических речевых сигналов и аддитивных помех, сосредоточенных в полосе частот около $f = 1000$ Гц, можно прийти к заключению, что при наихудших условиях, воздействующие на систему связи сосредоточенные шумовые помехи имеют мощность, которая выше мощности составляющих спектра речевого сигнала более, чем на 12 дБ.

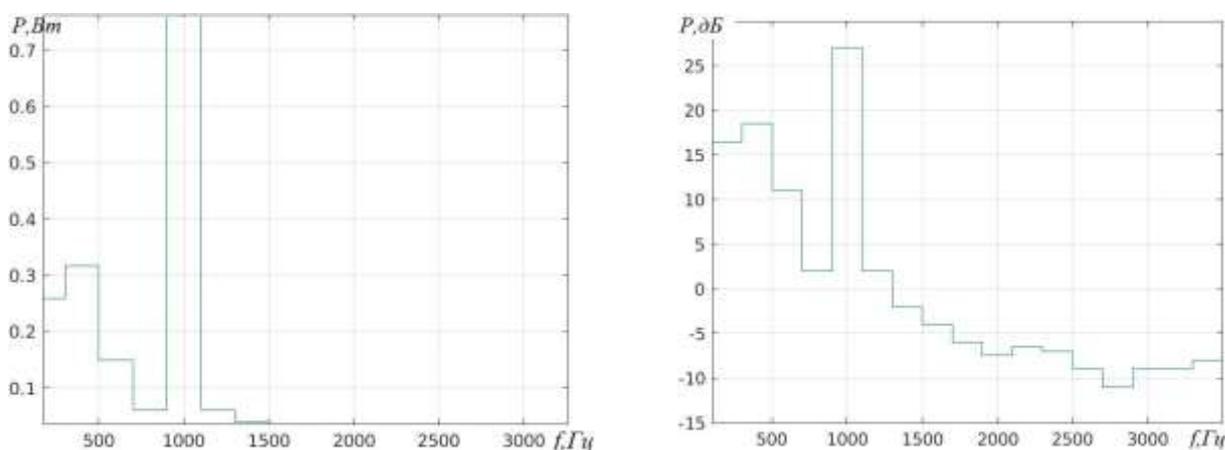


Рисунок 4.21 – Зависимость мощности выходного сигнала от частоты

Получение значений мощностей на диаграммах определяется следующим образом:

Для рисунка 4.21 «а» мощности определяются в соответствии с выражением:

$$P_{n\text{вых}} = \int_{100+n\Delta f}^{100+(n+1)\Delta f} k_n^2(f)G_{\text{ex}}(f)df,$$

Для рисунка 4.21 «б» мощности определяются в соответствии с выражением:

$$P_{n\text{ввб}}(\partial\text{БМ}) = 30 + 10 \cdot \log \int_{100+n\Delta f}^{100+(n+1)\Delta f} k_n^2(f)G_{\text{ex}}(f),$$

$$P_{n\text{ввх}}(\partial\text{БМ}) = 30 + 10 \log \int_{100+n\Delta f}^{100+(n+1)\Delta f} k_n^2(f)G_{\text{ex}}(f)df$$

Представленные диаграммы получены с учетом следующих значений параметров:

$$1 \leq n \leq N, \quad N = 16, \quad f = 100 + n\Delta f, \quad \Delta f = 200 \text{Гц}.$$

Итоговые результаты проведенного исследования зависимости эффективности передачи речевых сообщений от наличия сосредоточенных помех приведены на рисунке 4.21, при этом, если в канале нет средств подавления вышеуказанных сосредоточенных акустических помех, то данный канал передачи речевой информации полностью «забит» помехой, в связи с чем происходит понижение отношения сигнал/помеха и слоговой разборчивости. Это обуславливает необходимость проведения процедур шумоподавления и разработки соответствующих алгоритмов и структур программного обеспечения.

Кроме того рисунок 4.20 демонстрирует, что возможно применение различных методов, обеспечивающих фильтрацию сосредоточенных помех, минимизацию их влияния с применением подходов частотной селекции при применении гребенки фильтров, либо проведение амплитудной селекции сигналов. Если поставить пороговое устройство обработки на выходную линию n -го (параллельного) фильтра, то это обеспечит возможность формирования сигналов управления основным каналом с номером n , которое позволит создание алгоритма адаптивного шумоподавления сосредоточенных акустических помех, которые наблюдаются в канале n .

Таким образом, представленная на рисунке 4.20 структурная схема реализации адаптивного шестнадцатиканального алгоритма позволяет эффективно подавлять сосредоточенные акустические помехи. Степень подавления вышеуказанных помех определяется уровнем задержания выключенного фильтра в канале.

4.6. Алгоритм адаптивного подавления акустических помех с формантным распределением полос режекции

Алгоритм многоканальной адаптивной обработки и подавления помех, с использованием формантного распределения полос режекции можно представить в виде набора фильтров с узкой полосой пропускания (гребенки из фильтров). Все фильтры гребенки обеспечивают формирование общей полосы ΔF , которая делает возможной реализацию алгоритма шумоподавления с применением адаптивного управления диапазоном режекции сигнала в низкочастотной области амплитудно-частотной характеристики тракта передачи речевых сообщений. Это возможно за счет отключения одного или нескольких каналов передачи в низкочастотной части спектра сигнала.

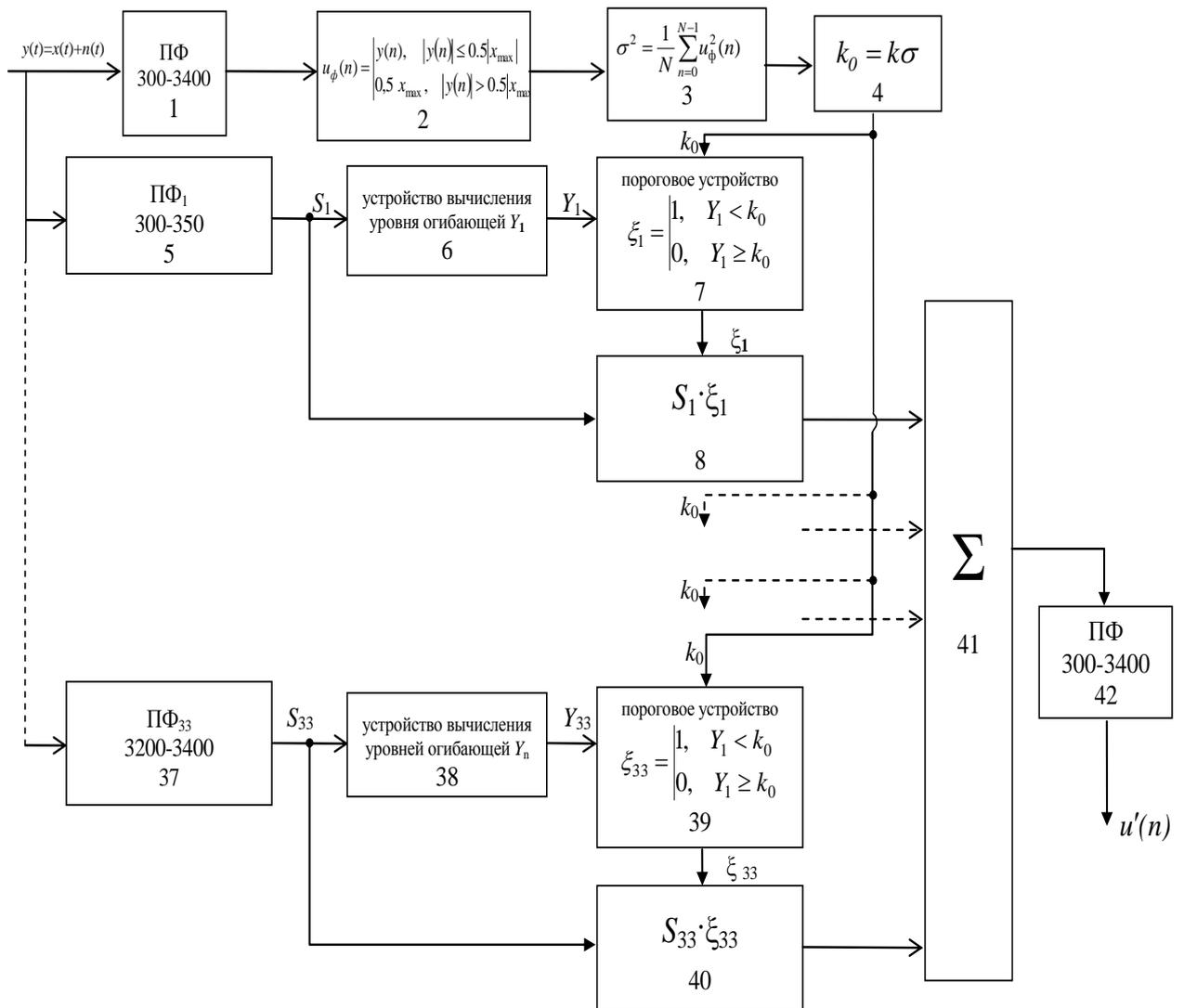


Рисунок 4.22 – Структура описанного многоканального устройства подавления помех.

Рисунок 4.22 представляет структуру устройства подавления аддитивных и сосредоточенных акустических помех. В состав устройства включены: (1) - полосовой фильтр, имеющий заданную ΔF , ($F_n = 300$ Гц, $F_s = 3400$ Гц) [40], (2) - блок получения сигнала $u_\phi(n)$, (3) - блок определения σ^2 , блок (4) производящий расчет пороговых уровней k_0 , гребенка узкополосные фильтров (5)-(37), блоки нахождения значений уровней огибающей Y_l (6)-(38), блоки формирования соответствующих сигналов управления ξ_l (7)-(39), блоки в которых осуществляется умножение значений составляющих S_l и полученных сигналов управления ξ_l (8)-(40), а также сумматор (41) и полосовой фильтр для выходного сигнала,

настроенный на частоту 300-3400 Гц (42).

В таблице 4.2. представлены частотные характеристики полосовых фильтров многоканального устройства шумоподавления. При этом Δf (ширина полосы пропускания) в соответствующем канале с номером l определено в зависимости от характеристики распределения определенных формантных частот в спектре акустического речевого сигнала [39].

Таблица 4.2. Соответствие характеристик полосовых фильтров с номерами каналов передачи

№ канала пропускания, l	№ блока многоканального устройства	Частотные характеристики фильтров
1	5	$f_n=300, f_g=350, \Delta f = 50$ Гц
2	6	$f_n=350, f_g=400, \Delta f = 50$ Гц
...		
14	18	$f_n=950, f_g=1000, \Delta f = 50$ Гц
15	19	$f_n=1000, f_g=1100, \Delta f = 100$ Гц
16	20	$f_n=1100, f_g=1200, \Delta f = 100$ Гц
...		
28	32	$f_n=2300, f_g=2400, \Delta f = 100$ Гц
29	33	$f_n=2400, f_g=2600, \Delta f = 200$ Гц
30	34	$f_n=2600, f_g=2800, \Delta f = 200$ Гц
...		
33	37	$f_n=3200, f_g=3400, \Delta f = 200$ Гц

Многоканальное устройство обработки акустических речевых сигналов, искаженных шумовыми помехами и компонентами эха функционирует следующим образом. На вход устройства подается сигнал $y(t)$, который можно представить в виде суммы полезного «истинного»

речевого сигнала, представленного $x(t)$, а также шумовой помехи, представленной $n(t)$. Смешанный зашумленный сигнал поступает на входную линию для полосового фильтра с полосой пропускания на частотах 300-3400 Гц (1), а также на множество входных линий для узкополосных частотных фильтров (5)-(37). Чтобы осуществить определение уровня пороговой обработки k_0 , пропущенный с выходной линии полосового фильтра (1) речевой сигнал, поступает на устройство (2), задачей которого является формирование сигнала $u_\phi(n)$ в соответствии с указанным ниже соотношением

$$u_\phi(n) = \begin{cases} y(n), & |y(n)| \leq 0,5|x_{\max}|, \\ 0,5x_{\max}, & |y(n)| > 0,5|x_{\max}|. \end{cases}$$

С выходной линии устройства (2) акустический речевой сигнал $u_\phi(n)$ подается на входную линию блока (3). В данном блоке осуществляется нахождение параметра σ^2 в соответствии с выражением

$$\sigma^2 = \frac{1}{N} \sum_{n=0}^{N-1} u_\phi^2(n),$$

в котором N – это количество полученных отсчетов в отдельно выбранном сегменте акустического речевого сигнала, имеющего длительность τ .

Полученное в блоке (3) σ^2 поступает на входную линию устройства (4). В нем осуществляется нахождение порогового уровня для обработки сигналов $k_0 = k\sigma$. Определенное значение k_0 передается на первый из входов блоков (7)-(39), в них осуществляется нахождение сигнала управления работой выбранного канала l , с учетом следующего выражения:

$$\xi_l = \begin{cases} 1, & Y_l < k_0 \\ 0, & Y_l \geq k_0 \end{cases},$$

в котором Y_l – является значением уровня огибающей составляющей S_l .

Результирующие значения S_l , полученные в (5) - (37), соответственно

передаются на первые входные линии блоков умножения (8) - (40), а также на входные линии блоков (6) - (38), в которых определяются уровни огибающих составляющих Y_l . Полученные Y_l подаются на вторые входы элементов (7) - (39), в которых происходит нахождение значений сигналов обеспечивающих управление обработкой ξ_l .

Полученные в (7) - (39) значения ξ_l передаются на входы (со вторыми номерами) блоков перемножения (8) - (40). В (8) - (40), реализуется перемножение управляющих сигналов ξ_l и значений S_l . Это позволяет получить результирующие сигналы каналов $S_l \cdot \xi_l$, которые в дальнейшем будут переданы на вход сумматора (41). Результирующий сигнал на выходной линии сумматора $U(n)$ находится в соответствии с выражением

$$U(n) = \sum_{l=1}^{33} S_l \cdot \xi_l,$$

в котором $\xi_l \in \{0,1\}$.

Те сигналы каналов с номером l , которые ранее были поражены искажающими полезный сигнал шумовыми составляющими [40], не пропускаются на выход устройства сумматора (41). При этом должно быть обеспечено равенство $\xi_l = 0$.

Полученные на выходе сумматора сигналы передаются на входную линию блока (42), который служит полосовым фильтром, имеющим ΔF с полосой 300-3400 Гц. На выходе (42) будет сформирован очищенный от помеховых составляющих акустический речевой сигнал $u'(n)$. Результирующий сигнал на выходе устройства может быть представлен выражением

$$u'(n) = k(f) \sum_{l=1}^{33} S_l \cdot \xi_l,$$

в нем $k(f)$ – полученная аппроксимация амплитудно-частотной характеристики полосового фильтра на выходе устройства.

$$k(f) = \begin{cases} 1, & \text{если } 300 \text{ Гц} \leq f \leq 3400 \text{ Гц} \\ 0, & \text{в случае других } f \end{cases}.$$

Полученная в работе и представленная рисунком 4.20 схема обеспечивает реализацию алгоритма позволяющего проводить адаптивное многоканальное подавление сосредоточенных акустических шумовых помех. При этом обеспечивается более эффективное понижение влияния помех (подавление на 8,8 - 15 дБ) за счет настраиваемого отключения определенных каналов передачи, расположенных в низкочастотной части гребенки фильтров, что обеспечивает создание возможности проведения адаптивной режекции в спектре сигналов, устраняющих сосредоточенные помехи в определенной частотной полосе.

4.7. Выводы

В настоящее время системы громкоговорящей связи и оповещения на многофункциональных объектах, работают в условиях активного воздействия различных акустических шумов, ревербераций и эхоотражений. Указанные выше нежелательные составляющие, искажающие акустические речевые сигналы, зачастую обладают очень высокой интенсивностью, вплоть до 90 дБ и более относительно I_0 . Так как в объектовых системах связи важное значение имеет качество передачи речевой информации, требуется передавать акустические речевые сигналы и сообщения с достаточной степенью достоверности и понимания абонентами. Поэтому в системы связи требуется интегрировать устройства, реализующие эффективные методы и алгоритмы обработки сигналов с целью подавления внешних акустических шумовых помех, а также устройства эхокомпенсаторов.

Существующие подходы к проектированию нелинейных алгоритмов понижения влияния акустических помех являются более подходящими для шумоподавления широкополосных помех, которые имеют спектральные характеристики, близкие к «белому шуму». Однако их вычислительная сложность существенно выше относительно линейных алгоритмов обработки. Известные алгоритмы обеспечивают подавление в диапазоне 6 - 8 дБ. Например, имеющие широкую полосу акустические шумовые помехи могут быть подавлены на 8 дБ, а такой вид шума, как «шум конференции» на

значение до 6 дБ.

В ходе проведенных при работе над кандидатской диссертацией исследований были получены новые алгоритмы обеспечивающие подавление акустических помех и эхосигналов, обладающие более высокой эффективностью. Полученный многоканальный алгоритм адаптивной обработки обеспечивает подавление внешних сосредоточенных помех на 30 дБ и более, при этом результаты могут зависеть от конкретной технической реализации полосовых фильтров в каналах передачи.

При рассмотрении результатов применения многоканального алгоритма подавления помех, обладающих формантным распределением полос режекции можно достичь следующих показателей эффективности подавления: при обработке сигналов, содержащих низкочастотные внешние помехи уровень подавления достигает $-(11\div 15)$ дБ при виде шума «шум моря», $-(17\div 23)$ дБ при виде шума «шум ветра», что обеспечивает отношение сигнал/помеха на уровне $22.5\div 34,2$ дБ.

Алгоритмы многоканальной обработки, использующие метод фильтрации с применением кратковременного дискретного преобразования Фурье (ДПФ), позволяют обеспечить подавление аддитивных и сосредоточенных помех до 15 дБ, при этом возможно получение результирующего отношения сигнал/шум в более 25 дБ. Отмеченное подтверждает, что разработанные алгоритмы и реализующее их программное обеспечение обеспечивают подавление помеховых составляющих сигналов на $3\div 15$ дБ больше по отношению к существующим системам и аналогичным алгоритмам, формируя результирующее отношение сигнал/помеха на уровне более 20 дБ, что обеспечивает необходимую для полного восприятия речевой информации слоговую разборчивость не менее 93%.

Разработанный алгоритм компенсации эхосигналов, в котором минимизировано количество каналов компенсации (оно соответствует количеству определенных каналов распространения эха), обеспечивает большую скорость сходимости, а также высокую степень устойчивости фильтров адаптивной обработки, что достигается за счет минимизации порядка применяемых адаптивных фильтров, а также сокращения числа настраиваемых в них коэффициентов. Это также позволяет повысить скорость их настройки,

обеспеченное более точным оценивания долговременных параметров эхосигналов, позволяя достичь компенсацию компонент эха на 1 - 4 дБ больше, по сравнению с известными алгоритмами.

ЗАКЛЮЧЕНИЕ

1. Разработана методика получения функции спектральной плотности мощности сигналов на конечных интервалах в телекоммуникационных системах, который позволил разработать алгоритм линейной фильтрации сигналов и может быть использован в исследованиях спектральных характеристик акустических сигналов и помех.

2. Получено, что спектральная плотность мощности речевых сигналов начинает уменьшаться со скоростью 20 дБ/октава, со значения частоты 300 – 500 Гц, при этом максимум плотности мощности приходится на диапазон частот 300..900 Гц. На частотах выше 3400 Гц спектральные составляющие понижаются более чем на 30 дБ.

Полученная информация позволила определить область режекции в низкочастотной части спектра линейной фильтрации, что существенно повысило эффективность подавления акустических помех до 12 дБ.

3. Для борьбы с сосредоточенными помехами применяется как частотная селекция гребенкой фильтров, так и одновременно амплитудная селекция. Для этого разработан адаптивный алгоритм подавления помех находящихся в n -ом канале.

4. Результаты корреляционно-экстремальной оценки параметров эхосигналов, позволили создать новый корреляционный алгоритм вычисления долговременных параметров эха – затухания и задержки, применение которого позволяет сократить число отводов в адаптивном фильтре, что приводит к уменьшению времени настройки эхокомпенсатора и к увеличению подавления эхосигналов на 1.9-4.1 дБ больше, по сравнению с известными алгоритмами.

5. Разработана структурная схема реализации алгоритма компенсации и программное обеспечение адаптивного управления областями режекции, которые позволили эффективно подавлять сосредоточенные акустические помехи, что повышает помехоустойчивость информационных

телекоммуникаций. Степень подавления помех определяется уровнем задержания выключенного фильтра в канале и находится в пределах более 12 дБ.

В результате проведенных исследований, получена модель, методы и алгоритмы повышения эффективности обработки акустических речевых сигналов под воздействием интенсивных помех и эхосигналов.

Путем создания устройств подавления внешних акустических помех и адаптивных эхокомпенсаторов, удалось получить отношение сигнал/помеха более 20 дБ и добиться слоговой разборчивости более 93%.

ЛИТЕРАТУРА

1. Адаптивные фильтры / Под ред. К.Ф.Н. Коуэна и П.М. Гранта. – М.: Мир, 1988. – 392 с.
2. Азаров И.С. .Вычисление мгновенных гармонических параметров речевого сигнала / И.С. Азаров, А.А. Петровский // Речевые технологии, 2008, № 1, с. 67 –
3. Аоки М. Оптимизация стохастических систем. М.: Наука, 1971. – 424 с.
4. Бахвалов Н.С. Численные методы.—9-е изд., электрон / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. - Москва : БИНОМ. Лаборатория знаний, 2020. - 636 с. - ISBN 978-5-00101-836-0. - URL: <https://ibooks.ru/reading.php?productid=366176> (дата обращения: 29.06.2020). - Текст: электронный.
5. Беллман,Р. Дифференциально-разностные уравнения/ Р. Беллман, К.Л. Кук. -М.: Мир, 1967. -548 с.
6. Бендат Дж. Прикладной анализ случайных данных/ Дж. Бендат, А. Пирсол. –М.: Мир, 1989. – 540 с.
7. Боровков А.А. Математическая статистика. – М.: Наука, 1984. – 472 с.
8. Бриллинджер Д. Временные ряды. Обработка данных и теория. – М.: Мир, 1980. – 536 с.
9. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. – М.: Наука, 1979. – 448 с.
10. Василенко Г.И. Теория восстановления сигналов: О редукции к идеальному прибору в физике и технике. – М.: Советское радио, 1979. – 272 с.
11. Вемян Г. В. Передача речи по сетям электросвязи. – М.: Радио и связь, 1985. – 272 с.
12. Гитлин В.Б. Совместный алгоритм выделения основного тона речи по методам GS и автокорреляционной функции спектра/ В.Б. Гитлин, Д.А. Лузин// Речевые технологии, 2008, № 3, с. 39 – 49.
13. Голд Б., Рэйдер Ч. Цифровая обработка сигналов / Б. Голд, Ч. Рэйдер – М.: Советское радио, 1973. – 368 с.

14. Горовиц А.М. Синтез систем с обратной связью. – М.: Советское радио, 1970. – 600 с.
15. Граничин О.Н. Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах / О.Н. Граничин, Б.Т. Поляк – М.: Наука, 2003. – 296 с.
16. Гурецкий Х. Анализ и синтез систем управления с запаздыванием. – М.: Машиностроение, 1974. – 328 с.
17. Ермолаев В.А. Вопросы параметрического представления нестационарных сигналов / В.А. Ермолаев, Ю.А. Кропотов // Проектирование и технология электронных средств, 2010, № 1, с. 31 – 35.
18. Ермолаев В.А. Методы оптимизации в задачах локальной обработки сигналов / В.А. Ермолаев, Ю.А. Кропотов // В мире научных открытий, 2010, № 6.1, с. 44 – 47.
19. Ермолаев В.А. Об оценивании функций распределения и стабилизации характеристик случайных процессов релейными системами // Приборы и системы управления, 1997, №4, с.36 – 37.
20. Ермолаев В.А. Оценивание акустических тональных сигналов методами линейной и нелинейной регрессии / В.А. Ермолаев, Ю.А. Кропотов // Радиотехнические и телекоммуникационные системы, 2011, №1 (1), с.71–73.
21. Ермолаев В. А. Устройство адаптивного подавления акустических шумов и акустических сосредоточенных помех / В.А. Ермолаев, Ю.А. Кропотов :пат. № 2502185. 2013. Рос. Федерация: бюл. № 35.
22. Журавлев В. Н. Анализ параметров автокорреляционной функции фрагментов фонем в задачах идентификации абонента / В.Н. Журавлев, В.С. Кабак, В.О. Рыбин // Радіоелектроніка. Інформатика. Управління, № 2, 2004, с. 29 – 33.
23. Кайно Г. Акустические волны: устройство, визуализация и аналоговая обработка сигналов . Мир,1990.-655с.
24. Катковник В.Я. Непараметрическая идентификация и сглаживание данных: метод локальной аппроксимации. М.: Главная редакция физико-

математической литературы, 1985. – 336 с.

25. Кашьяп Р.Л. Построение динамических стохастических моделей по экспериментальным данным / Р.Л. Кашьяп, А.Р. Рао – М.: Наука, 1983. – 384 с.
26. Кей С.М. Современные методы спектрального анализа: Обзор / С.М. Кей, С.Л. Марпл мл. // ТИИЭР, т. 69, №11, 1981, с. 5 – 51.
27. Кропотов Ю.А. Алгоритм вычисления сигнала управления каналом режекции многоканальной системы передачи акустических сигналов // Вопросы радиоэлектроники. Серия ОТ, 2010, вып. 1, с 57 – 59.
28. Кропотов Ю.А. Алгоритм определения параметров экспоненциальной аппроксимации закона распределения вероятности амплитуд речевого сигнала // Радиотехника, 2007, № 3, с. 44 – 47.
29. Кропотов Ю.А. Анализ факторов снижения разборчивости речи в системах громкоговорящей связи / Ю.А. Кропотов, Я.Ю. Кульков // Известия Орловского ГТУ, сер. «Фундаментальные и прикладные проблемы техники и технологии: информационные системы и технологии», 2008, № 1-3/269 (544), с. 129 – 133.
30. Кропотов Ю.А. Модель одномерной функции плотности вероятности речевых сигналов // Системы управления, связи и безопасности, 2015, № 4, с. 158 – 170.
31. Кропотов Ю.А. Аппроксимация закона распределения вероятности амплитуд речевого сигнала // Proceedings of 16-th International Crimean Conference “Microwave & Telecommunication Tecnology”. Sevastopol, Ukraine, 2007, v.1, p. 422 – 425. IEEE Catalog, Namber 06EX1376, ISBN 966-322-006-6.
32. Кропотов Ю.А. Аппроксимация закона распределения вероятности отсчетов сигналов акустических помех / Ю.А. Кропотов, А.А. Быков // Радиотехнические и телекоммуникационные системы, 2011. - №2(2). – С.61-63.
33. Кропотов Ю.А. Аппроксимация огибающей энергетического спектра речевого сигнала / Ю.А. Кропотов, А.А. Быков //18-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные

технологии» (КрыМиКо'2008). Севастополь, Украина, 8-12 сентября 2008 г.:
Материалы конф. в 2т. – Севастополь, Украина: Вебер.2008. – ISBN 966-322-
006-6. – С.305-307.

34. Кропотов Ю.А. Временной интервал определения закона
распределения вероятности амплитуд речевого сигнала // Радиотехника,
2006, № 6, с.97 – 98.

35. Кропотов Ю.А. Исследование автокорреляционных функций речевых
сигналов / Ю.А. Кропотов, А.А. Быков // Радиотехника, 2008, № 9, с. 107 –
109.

36. Кропотов Ю.А. Исследование и разработка систем громкоговорящей
связи в условиях воздействия акустических помех: монография /
Ю.А. Кропотов, А.А. Быков // Муром: Изд.-полиграфический центр МИ
ВлГУ, 2011. – 107 с.: ил. + 3 табл. - Библиогр.: 156 назв. ISBN 978-5-8439-
0288-9

37. Кропотов Ю.А. Повышение отношения сигнал/акустическая помеха в
телекоммуникациях аудиообмена методами адаптивной фильтрации /
Ю.А. Кропотов, А.А. Белов, А.Ю. Проскуряков // Сборник трудов ИТНТ-
2019, 2019, с. 979-985

38. Кропотов Ю.А. Математическая модель закона распределения
вероятности амплитуд речевых сигналов в базисе экспоненциальных
функций / Ю.А. Кропотов, А.Ю. Проскуряков // Proceedings of 17-th
International Crimean Conference “Microwave & Telecommunication
Technology”. Sevastopol, Ukraine. 2007, v.1, p. 364 – 366. IEEE Catalog, Number
07EX1742, ISBN 978-966-335-014-1.

39. Кропотов Ю. А. Методы проектирования алгоритмов обработки
информации телекоммуникационных систем аудиообмена: монография /
Ю.А. Кропотов, А.А. Парамонов - Москва-Берлин: Директ-Медиа, 2015. –
226 с.

40. Кропотов Ю.А. Модели и методы обработки аудиосигналов
телекоммуникационных систем в сложной помеховой обстановке [Текст]:

дис. доктора технических наук: 05.12.13: защищена 25.10.13 / Автор Кротов Юрий Анатольевич – М., 2013. - 273 с.

41. Кротов Ю.А. Об алгоритме вычисления сигнала управления каналом режекции адаптивной многоканальной системы передачи акустических сигналов // Proceedings of 20-th International Crimean Conference “Microwave & Telecommunication Technology”. Sevastopol, Ukraine. 2010, v.1, p. 305 – 307. IEEE Catalog, Number CEP10788, ISBN 978-966-335-329-6.

42. Кротов Ю.А. О методах локально-оптимального представления сигналов на конечных интервалах: спектральный анализ и характеристики качества / Ю.А. Кротов, А.А. Быков, В.А. Ермолаев // 19-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии» (КрыМиКо'2009). Севастополь, Украина, 14-18 сентября 2009 г.: Материалы конф. в 2т. – Севастополь, Украина: Вебер.2009. – ISBN 966-322-006-6. – С.387 – 390.

43. Кротов Ю.А. Повышение отношения сигнал/помеха в многоканальной системе передачи акустических сигналов при воздействии узкополосных помех // Информационные системы и технологии, 2011, № 1 , с. 102-106.

44. Кротов Ю.А. Погрешность при определении закона распределения вероятности амплитуд речевого сигнала / / Ю.А. Кротов, А.А. Быков // Математическое и программное обеспечение вычислительных систем: Межвуз. сб. науч. тр. / Под ред. А.Н.Пылькина – М.: Горячая линия-Телеком, 2006. – 70 – 74 с.

45. Кротов Ю.А. Повышение отношения сигнал/акустическая помеха в телекоммуникациях аудиообмена методами адаптивной фильтрации / Ю.А. Кротов, А.А. Белов, А.Ю. Проскуряков // Сборник трудов ИТНТ-2019, 2019, с. 979-985

46. Кротов Ю.А. Экспериментальные исследования закона распределения вероятности амплитуд сигналов системы передачи речевой информации // Проектирование и технология электронных средств, 2006, №4,

с. 37 – 42.

47. Лайонс, Ричард Цифровая обработка сигналов; пер. с англ. под ред. А. А. Бритова. - 2-е изд. - Москва : БИНОМ : БИНОМ-Пресс, 2006. - 652 с. : ил. ; - Пер. изд. : Understanding digital signal processingJ. / R. Lyons. - 2000 экз. - ISBN 5-9518-0149-4
48. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической обработки наблюдений. – М.: Физматлит, 1958. – 336 с.
49. Лоран П.-Ж. Аппроксимация и оптимизация. – М.: Мир, 1975. – 496 с.
50. Лоусон Ч. Численное решение задач метода наименьших квадратов / Ч. Лоусон, Р. Хенсон – М.: Наука, 1986. – 232 с.
51. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990. – 584 с.
52. Методы исследования нелинейных систем автоматического управления / Под ред. Р.А. Нелепина. – М.: Наука, 1975. – 448 с.
53. Назаров М.В. Методы цифровой обработки и передачи речевых сигналов / М.В. Назаров, Ю.Н. Прохоров - М.: Радио и связь, 1985. - 176 с., ил.
54. Нестационарные системы автоматического управления: анализ, синтез и оптимизация / Под ред. К.А. Пупкова и Н.Д. Егупова. – М.: Издательство МГТУ им. Н.Э. Баумана, 2007. – 632 с.
55. Нестеров Ю.Е. Введение в выпуклую оптимизацию. – М.: МЦНМО, 2010. – 280 с.
56. Никиас Х.Л. Биспектральное оценивание применительно к цифровой обработке сигналов / Х.Л. Никиас, М.Р. Рагувер // ТИИЭР, т. 75, № 7, 1987, с. 5 – 30.
57. Оппенгейм А. Цифровая обработка сигналов. Издание 3-е, исправленное / А. Оппенгейм, Р. Шафер. - Москва : Техносфера, 2012. - 1048 с. ISBN 978-5-94836-329-5. - URL: <https://ibooks.ru/reading.php?productid=339869> (дата обращения: 29.06.2019). - Текст: электронный.
58. Ортега Дж. Итерационные методы решения нелинейных систем

- уравнений со многими неизвестными / Дж. Ортега, В. Рейнболдт – М.: Мир, 1975. – 560 с.
59. Перов В.П. Прикладная спектральная теория оценивания. – М.: Наука, 1982. – 432 с.
60. Петровский А. Обработка речи на основе дискретного преобразования Фурье с неравномерным частотным разрешением / А. Петровский, А. Борович, М. Парфенюк // Речевые технологии, 2008, № 3, с. 3 – 15.
61. Петровский А. Дискретное преобразование Фурье с неравномерным частотным разрешением в перцептуальных системах редактирования шума в речи разрешением / А. Петровский, А. Борович, М. Парфенюк // Речевые технологии, 2008, № 3, с. 16 – 26.
62. Петровский А.А. Субполосная обработка сигналов: эффективность и применение в речевых технологиях // Речевые технологии, № 1, 2010, с. 13 – 26.
63. Полак Э. Численные методы оптимизации. Единый подход. – М.: Мир, 1974. – 376 с.
64. Поляк Б.Т. Алгоритмы матричного оценивания /Щербаков П.С. // Автоматика и телемеханика, 1995, № 11, с. 122 – 139.
65. Поляк Б.Т. Введение в оптимизацию. – М.: Наука, 1983. – 384 с.
66. Покровский Н.Б. Расчет и измерение разборчивости. – М.: Связьиздат, 1962. – 472 с.
67. Попов О.Б. Цифровая обработка сигналов в трактах звукового вещания. Учебное пособие для вузов. - 2-е изд., стереотип. / О.Б. Попов, С.Г. Рихтер. - Москва : Горячая Линия–Телеком, 2015. - 342 с. - ISBN 978-5-9912-0289-3. - URL: <https://ibooks.ru/reading.php?productid=354354> (дата обращения: 29.06.2019). - Текст: электронный.
68. Поршнева С. В. Вычислительная математика: курс лекций. - Санкт-Петербург: БХВ-Петербург, 2014. - 320 с. - ISBN 978-5-9775-2002-7. - URL: <https://ibooks.ru/reading.php?productid=18464> - Текст: электронный.
- 69 Паспорта научных специальностей – ВАК [сайт]

- URL=<http://arhvak.minobrnauki.gov.ru/316> (дата обращения: 29.06.2019). -
Текст: электронный.
70. Рабинер Л.Р. Теория и применение цифровой обработки сигналов / Л.Р. Рабинер, Б. Гоулд – М.: Мир, 1975. – 835 с., ил.
71. Рашевский, Я. И. Обзор зарубежных методов определения разборчивости речи. / Я. И. Рашевский, В. Л. Каргашин // М.: Специальная техника, № 4, 2002. – С. 37-40.-Библиогр.:с.40.
72. Резван В. Абсолютная устойчивость автоматических систем с запаздыванием. – М.: Наука, 1983. – 360 с.
73. Сапожков М.А. Вокодерная связь / М.А. Сапожков, В.Г. Михайлов – М.: Радио и связь, 1983. – 248 с.
74. Сапожков, М.А. Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1962.– 452 с.
75. Смит С. Цифровая обработка сигналов практическое руководство для инженеров и научных работников. М.: Додэка XXI, 2012. -718 с.
76. Солодов А.В. Системы с переменным запаздыванием / А.В. Солодов, Е.А. Солодова – М.: Наука, 1980. – 384 с.
77. Сондхи М.М. Методы подавления эха в телефонных сетях / М.М. Сондхи, Д.А. Беркли // ТИИЭР, т. 68, №8, 1980, с. 5 – 24.
78. Сорока А.М. Алгоритм двухэтапного распознавания фонем русского языка // Речевые технологии, № 1, 2010, с. 35 – 41.
79. Сорока А.М. Комплексная верификация ключевых слов на основе метода опорных векторов // Речевые технологии, № 1, 2010, с. 27 – 34.
80. Сорокин В.Н. Верификация диктора по спектрально-временным параметрам речевого сигнала / В.Н. Сорокин, А.И. Цыплихин // Информационные процессы, 2010, Том 10, № 2, с. 87 – 104.
81. Сорокин В.Н. Фундаментальные исследования речи и прикладные задачи речевых технологий // Речевые технологии, 2008, № 1, с. 18 – 48.
82. Сюзев В.В. Основы теории цифровой обработки сигналов. - РТСофт,2014. -754с.

83. Тихонов А.Н. Методы решения некорректных задач / А.Н. Тихонов, В.Я. Арсенин. – М.: Наука, 1979. – 288 с.
84. Тихонов В.И. Статистический анализ и синтез радиотехнических устройств и систем / В.И. Тихонов, В.Н. Харисов – М.: Радио и связь, 1991. – 608с
85. Троян В.Н. Методы аппроксимации геофизических данных на ЭВМ / В.Н. Троян, Ю.М. Соколов – Л.: Издательство Ленинградского университета, 1989. – 304 с.
86. Уилкс С. Математическая статистика. – М.: Наука, 1967. – 632 с.
87. Умняшкин С.В. Основы теории цифровой обработки сигналов. – М.: Техносфера, 2018. - 528 с. - ISBN 978-5-94836-508-4. - URL: <https://ibooks.ru/reading.php?productid=359195>. - Текст: электронный.
88. Пат. 91490 Российская Федерация, МПК7 Н04В1/10 (2006.01). Устройство подавления сосредоточенных акустических помех. / А.А. Быков, В.А. Ермолаев, Ю.А. Кропотов, Л.С. Лазарев, А.Ю. Проскуряков Заявители и патентообладатели Быков А.А., Ермолаев В.А., Кропотов Ю.А., Лазарев
89. Формалев В.Ф. Численные методы / В.Ф. Формалев, Д.Л. Ревизников. - М.: Физматлит, 2004. – 400с. ISBN 5-9221-0479-9.
90. Френкс Л. Теория сигналов. – М.: Советское радио, 1974. – 344 с.
91. Хейл, Дж. Теория функционально-дифференциальных уравнений. – М.: Мир, 1984. – 421 с.
92. Холкина Н.Е. Идентификация и оценивание параметров акустических сигналов в телекоммуникационных системах аудиообмена / Н.Е. Холкина, Ю.А. Кропотов, А.Ю. Проскуряков, Д.В. Бейлекчи Сборник трудов. IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» ИТНТ-2018, 2018. С. 1815-1821.
93. Холкина Н.Е. Идентификация и оценивание параметров акустических сигналов в телекоммуникационных системах аудиообмена / Н.Е. Холкина, Ю.А. Кропотов // Вестник Брянского государственного технического университета. 2018. № 2 . С. 34-41.

94. Холкина Н.Е. Исследование статистических характеристик сосредоточенных акустических помех в системах телекоммуникаций аудиообмена/ Н.Е. Холкина, Ю.А. Кропотов, С.В. Савинов // Инновационные, информационные и коммуникационные технологии: сборник трудов XV Международной научно-практической конференции. / под. ред. С.У.Увайсов – Москва: Ассоциация выпускников и сотрудников ВВИА им. проф. Жуковского, 2019, С. 413-416.
95. Холкина Н.Е. Моделирование сигналов в телекоммуникациях аудиообмена в условиях акустических помех / А.А. Белов, Ю.А. Кропотов, А.Ю. Проскуряков, Н.Е. Холкина // Информационные системы и технологии. 2019 №5 (115). С. 92-102.
96. Холкина Н.Е. Модель помехи с медленно изменяющимися параметрами в системах обмена акустическими сигналами / Н.Е. Холкина, Ю.А. Кропотов, А.А. Белов, А.Ю. Проскуряков // Инновационные, информационные и коммуникационные технологии: Сборник трудов XIII Международной научно-практической конференции. /под редакцией С.У.Увайсов - Москва: Ассоциация выпускников и сотрудников ВВИА им. проф. Жуковского, 2016.-с.280-283. ISSN 2500-1248
97. Холкина Н.Е. Оценивание моделей сигналов и акустических помех в телекоммуникациях аудиообмена / Н.Е. Холкина, Ю.А. Кропотов, А.А. Белов, А.Ю. Проскуряков // Системы управления, связи и безопасности. 2018. № 3. С. 1-13. URL: <http://sccs.intelgr.com/archive/2018-03/01-Kropotov.pdf>
98. Холкина Н.Е. Оценивание нестационарных сигналов в системах телекоммуникационного аудиообмена/ Н.Е. Холкина, Ю.А. Кропотов // Наука и образование в развитии промышленной, социальной и экономической сфер регионов России. XI Всероссийские научные Зворыкинские чтения: сб. тезисов докладов Издательско-полиграфический центр Ми ВлГУ. 9.02.2018. – Муром: Издательско-полиграфический центр Ми ВлГУ, 2018

99. Холкина Н.Е. Оценивание характеристик стационарных акустических сигналов в телекоммуникационных системах / Н.Е. Холкина, Ю.А. Кропотов // Наука и образование в развитии промышленной, социальной и экономической сфер регионов России. XI Всероссийские научные Зворыкинские чтения: сб. тезисов докладов Издательско-полиграфический центр Ми ВлГУ. 8.02.2019. – Муром: Издательско-полиграфический центр Ми ВлГУ, 2019
100. Холкина Н.Е. Функции регрессии и метод наименьших квадратов в задачах параметрического моделирования в системах обмена информацией акустическими сигналами / Н.Е. Холкина, Ю.А. Кропотов // Вестник Брянского государственного технического университета. 2018. № 1 . С. 35-40.
101. Холкина Н.Е., Моделирование сигналов и акустических помех в телекоммуникациях аудиообмена / Н.Е. Холкина, Ю.А. Кропотов, А.А. Белов, А.Ю. Проскуряков // Вестник Брянского государственного технического университета. 2018. № 9 . С. 39-46.
102. Холкина Н.Е., Расширенное микширование аудиопотоков для многопроцессорных устройств в телекоммуникациях / Н.Е. Холкина, А.А. Колпаков, Ю.А. Кропотов, А.А. Белов // Системы управления, связи и безопасности. 2016. №3. С. 227-238.
103. Холкина Н.Е., Функционально-дифференциальные и дифференциально-разностные модели систем с акустической обратной связью / Н.Е. Холкина, В.А. Ермолаев, Ю.А. Кропотов Вестник Брянского государственного технического университета. 2019. № 4 . С. 60 - 67.
104. Холкина Н.Е., Эффективность передачи информации систем оповещения и телекоммуникаций аудиообмена в условиях помех // Вестник Брянского государственного технического университета. 2020 № 5 (90). С. 45-55.
105. Яглом А.М. Корреляционная теория стационарных случайных функций. - Л.: Гидрометеиздат, 1981. – 280 с.
106. Янушевский Р.Т. Теория линейных оптимальных многосвязных систем управления. – М.: Наука, 1973. – 464 с.

107. Янь Цзинбинь Классификация аудиосигналов с использованием одноклассового метода опорных векторов для систем поиска информации в мультимедиаархивах / Янь Цзинбинь, У Ши, А.М. Сорока, А.А. Трус // Речевые технологии, 2008, № 4, с. 15 – 24.
108. Agarwal, R.P. Nonoscillation theory of functional differential equations/ R.P. Agarwal, L. Berezansky, E. Braverman, G. Domoshitsky. – New York: Springer, 2012. – 520 p.
109. Ben-Tal A. Lecture on modern convex optimization. Analysis, Algorithms and engineering applications / Tal A. Ben-, A. Nemirovski Philadelphia: SIAM, 2001. – 504 p.
110. Ben-Tal A. Robust optimization. / Tal A. Ben-, L.E. Ghaoui, A. Nemirovski Princeton and Oxford, Princeton University Press, 2009. – 564 p.
111. Boyd S. Convex optimization / S. Boyd, L. Vandenberghe Cambridge University Press, 2004. – 718 p.
112. Browse for Sounds - findsounds.com [Электронный ресурс]: FindSounds - Электрон. текстовые дан., [2009].- URL= findsounds.com, Режим доступа: свободный.-Загл. с экрана.
113. Cohen, I. Speech processing in modern communication / I. Cohen, J. Benesty, S. Gannot. – Berlin, Heidelberg: Springer, 2010. – 342 p.
114. Hansler, E. Topics in acoustic echo and noise control: Selected methods for the cancelation of acoustic echoes, the reduction of background noise, and speech processing / E. Hansler, G. Schmidt. – Berlin, Heidelberg: Springer, 2006. – 642 p.
115. Kholkina N.E. Identification and estimation of acoustic signals parameters in telecommunication systems of audio exchange /N.E. Kholkina, Y.A. Kropotov, A.Y. Proskuryakov, D.V. Beilekchi // The IV International Conference on Information Technology and Nanotechnology. IOP Conf. Series: Journal of Physics: Conf. Series, 1096 (2019) 012192, doi:10.1088/1742-6596/1096/1/012192 (ИТНТ-2018)
- Kholkina, N.E., Proskuryakov, A.Y., Belov, A.A. Development of an Algorithm for Suppression of Concentrated Acoustic Noise in Telecommunication Systems

/N.E. Kholkina, A.Y. Proskuryakov, A.A. Belov, 2021 23rd International Conference on Digital Signal Processing and its Applications, DSPA 2021, DOI 10.1109/DSPA51283.2021.9535906

116. Kolmanovskii, V. Introduction to the theory and applications of functional differential equations / V. Kolmanovskii, A. Myshkis. – Dordrecht, Boston, London: Kluwer Academic Publishers, 1999. – 648 p.

117. Kropotov Y.A. Applicatio Method of Barrier Functions in the Problem of Estimating the Probability Density of the Parameterized Approximations / Y.A. Kropotov, A.A. Belov // 2016 13th International Scientific-Technical Conference APEIE – 39281, Vol. 1, part 2, 2016, pp. 69-72

118. Kropotov Y.A. Methods Assessment the Probability Density of Discrete Signals in Telecommunications / Y.A. Kropotov, A.A. Belov // Supplementary Proceedings of the 9th International Conference on Discrete Optimization and Operations Research and Scientific School (DOOR 2016) Vol-1623, 2016, pp. 745-754

119. Sound Ideas Sound Effects Library - footage.3dn.ru [Электронный ресурс]: FindSounds - Электрон. текстовые дан., [2009].- Режим доступа: findsounds.com, свободный.-Загл. с экрана.

120. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости.

121. ГОСТ 16600-72. Передача речи по трактам радиотелефонной связи. Требования к разборчивости речи и методы артикуляционных измерений.

122. ГОСТ Р 51061-97. Система низкоскоростной передачи по цифровым каналам. Параметры качества речи и методы измерений.

123. ITU-T Recommendation P.340. Transmission characteristics and speech quality parameters of hands-free terminals. (05/2000).

124. ITU-T Recommendation P.342. Transmission characteristics for telephone band (300 – 3400 Hz) digital loudspeaking and hands-free telephony terminals. (08/1996).

125. ITU-T Recommendation P.800. Methods for subjective determination of

transmission quality.

126. ITU-T Recommendation P.862. Perceptual Evaluation of Speech Quality. (02/2001).

127. Свидетельство о государственной регистрации программы для ЭВМ № 2021613972. Программное обеспечение телекоммуникационной системы громкоговорящей связи и оповещения с шумоподавлением и эхокомпенсацией. Правообладатели и авторы: Холкина Н.Е., Проскуряков А.Ю., Бейлекчи Д.В., Белов А.А., Колпаков А.А. Заявка 2021612824. Дата поступления 05 марта 2021. Зарегистрировано в реестре программ для ЭВМ 17 марта 2021 г.

ПРИЛОЖЕНИЕ А

Акты внедрения



УТВЕРЖДАЮ

Генеральный директор
АО «Муромский радиозавод»

Денисов С.Н.

«03» 09 2022 г.

АКТ ВНЕДРЕНИЯ

результатов диссертационной работы Холкиной Натальи Евгеньевны на тему «Алгоритмы обработки речевых сигналов телекоммуникационных систем в условиях помех» на соискание ученой степени кандидата технических наук по специальности 2.2.15 – «Системы, сети и устройства телекоммуникаций».

Научно-технический совет АО «Муромский радиозавод», г. Муром, составил настоящий акт в том, что теоретические и практические результаты диссертационной работы Холкиной Натальи Евгеньевны такие как:

- разработанные алгоритмы повышения помехозащищенности аудиообмена в телекоммуникационных системах связи;
- разработанная методика оценивания слоговой разборчивости в системах передачи речевых сообщений с применением сетчатой функции, обеспечивающая практическое оценивание слоговой разборчивости по отношению сигнал/помеха;
- разработанный адаптивный алгоритм подавления аддитивных и сосредоточенных акустических помех, позволяющий осуществить подавление помех на 30 дБ.

внедрены в процессе разработки аппаратного и программного обеспечения оперативно-командных телекоммуникационных систем ГГС, трансляции и оповещения в ходе выполнения ОКР «Модернизация комплекса оперативно-командной громкоговорящей и телефонной связи КГС-01ЦС», выполняемой АО «Муромский радиозавод». Приведенные выше результаты диссертации в исполнении указанного проекта повышают тактико-технические характеристики устройств комплекса.

Главный инженер

Волков А.В.

УТВЕРЖДАЮ

Директор Муромского института (филиала)
ФГБОУ ВО «Владимирский государственный
университет имени Александра Григорьевича и
Николая Григорьевича Столетовых»



д.т.н., профессор Жизняков А.Л.
« 05 » сентября 2022 г.

АКТ ВНЕДРЕНИЯ

результатов диссертационной работы Холкиной Натальи Евгеньевны на тему:
«Алгоритмы обработки речевых сигналов телекоммуникационных систем в условиях помех» на соискание ученой степени кандидата технических наук по специальности 2.2.15 – « Системы, сети и устройства телекоммуникаций » в учебном процессе Муромского института (филиала) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых» (МИВлГУ)

Мы, нижеподписавшиеся, начальник учебного отдела МИВлГУ Педя Т.Н., декан факультета информационных технологий и радиоэлектроники МИВлГУ к.т.н., доцент Храмов К.К. и заведующий кафедрой «Электроники и вычислительной техники» доцент Белов А.А. составили настоящий акт в том, что результаты диссертационной работы Холкиной Н.Е. внедрены в учебный процесс кафедры «Электроники и вычислительной техники». В частности, в учебном процессе использованы такие результаты исследований:

- разработанная табличная функция погрешности гистограммного оценивания оцифрованных сигналов в зависимости от времени анализа, с помощью которой возможно определение длительности кадра в алгоритмах обработки данных;
- разработанная модель функции плотности вероятности речевого сигнала в виде многочлена по системе экспоненциальных функций, построенная на основании гистограммной оценки речевых сигналов в системах телекоммуникаций аудиообмена;
- разработанный метод формирования сигнала управления областью режекции в многоканальном адаптивном алгоритме подавления аддитивных и сосредоточенных помех,
- разработанный алгоритм формирования сигнала управления автоматическим выключением пораженных сосредоточенными помехами каналов системы связи.

Вышеуказанные модели и алгоритмы используются в лекционных курсах по дисциплинам «Теория передачи цифровой информации», «Современные компьютерные методы исследования сигналов», «Теория информации», а также рекомендованы к применению в ходе курсового и дипломного проектирования для студентов направлений подготовки 09.03.01 (уровень бакалавриата), 09.04.01 (уровень магистратуры) «Информатика и вычислительная техника».

Начальник учебного отдела

Педя Т.Н.

Декан факультета информационных технологий и радиоэлектроники, к.т.н., доцент

Храмов К.К.

Заведующий кафедрой ЭиВТ, к.т.н., доцент

Белов А.А.

ПРИЛОЖЕНИЕ В

Свидетельство о внедрении интеллектуальной собственности

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ
№ 2021613972

**Программное обеспечение телекоммуникационной
системы громкоговорящей связи и оповещения с
шумоподавлением и эхокомпенсацией**

Правообладатели: *Холкина Наталья Евгеньевна (RU), Колпаков Александр Анатольевич (RU), Бейлекчи Дмитрий Владимирович (RU), Белов Алексей Анатольевич (RU), Проскуряков Александр Юрьевич (RU)*

Авторы: *Холкина Наталья Евгеньевна (RU), Колпаков Александр Анатольевич (RU), Бейлекчи Дмитрий Владимирович (RU), Белов Алексей Анатольевич (RU), Проскуряков Александр Юрьевич (RU)*

Заявка № 2021612824

Дата поступления 05 марта 2021 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 17 марта 2021 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Иелиев

ПРИЛОЖЕНИЕ С

Листинг модулей программного обеспечения телекоммуникационной системы громкоговорящей связи и оповещения с шумоподавлением и эхокомпенсацией

Модуль mc_mainprog.cpp

```
#include <QtGlobal>
#include <QMap>
#include <QVariant>
#include <QDebug>
#include "mc_mainprog.h"

#if QT_VERSION < 0x050000
#define qInfo() qDebug()
#endif

mc_mainprog::mc_mainprog(QObject *parent) : QObject(parent)
{
    cfgdev = new MC_configdevs(CONFIG_FILE);
    checkgpio = new QTimer();
    exec_process = NULL;
}

void mc_mainprog::Init() {

    //LoadKSettings();
    cfgdev->ConfigMCSwitch();
    cfgdev->ConfigMCSensors();
    connect(checkgpio, SIGNAL(timeout()), this, SLOT(CheckGpioProcess()));
    checkgpio ->start(TIMER_IMP_TIME);

    qInfo()<<"---- init done. main cycle working";
}
QProcess *process=NULL;

void mc_mainprog::ExecCmd() {
    if (process) {
        qDebug() << "Already started";
        return;
    }

    exec_process = new QProcess(this);
    QMap<QString,QVariant> videocfgr = cfgdev->mc_conf["video"].toMap();
    QString file = videocfgr["cmdstr"].toString();
    qDebug() << file;
    //QProcess::startDetached(file);
    exec_process->setProcessChannelMode(QProcess::ForwardedChannels);
    exec_process->start(file);
    exec_process->waitForStarted(5000);
}

void mc_mainprog::CheckGpioProcess() {
    for(int i=0;i<this->cfgdev->mc_sensorslist->length();i++) {

        if (cfgdev->mc_sensorslist->at(i)->CheckSensor()) {
            qInfo() << "Detect " <<cfgdev->mc_sensorslist->at(i)-
>pin_sensor;

            if (cfgdev->mc_sensorslist->at(i)->name=="video") {
                if (!exec_process) {
                    qDebug() << "Start exec";
                    ExecCmd();
                    for(int j=0; j<cfgdev->mc_switchlist-
>count();j++)

                        if(cfgdev->mc_switchlist->at(j)->name == "video_sw") {
                            cfgdev->mc_switchlist->at(j)->Set(1);
                            break;
                        }
                } else {
                    qDebug() << "Stop exec";
                }
            }
        }
    }
}
```

```

        for(int      j=0;      j<cfgdev->mc_switchlist-
>count();j++)
            if(cfgdev->mc_switchlist->at(j)->name == "video_sw") {
                cfgdev->mc_switchlist->at(j)->Set(0);
                break;
            }
            QMap<QString,QVariant>   videocfgr   =   cfgdev-
>mc_conf["video"].toMap();
            exec_process-
>write(videocfgr["cmdexit"].toString().toLocal8Bit().data());
            exec_process->waitForBytesWritten(3000);
            exec_process->waitForFinished(3000);
            if (exec_process->state() !=QProcess::NotRunning)
{
                qDebug() << "Process not stopped. Try kill
exec";

                exec_process->kill();
            }
            delete exec_process;
            exec_process = NULL;
        }
    }
}

```

Модуль mc_mainprog.h

```

#ifndef MC_MAINPROG_H
#define MC_MAINPROG_H

#include <QObject>
#include <QProcess>
#include "mc_configdevs.h"

#ifdef Q_OS_UNIX
#define LOG_DIR "/var/log/kopv/"
#define LOG_FILE "kopv"
#define CONFIG_FILE "/etc/kpv.json"
#elif defined(Q_OS_WIN)
#define LOG_DIR "../etc/log/"
#define LOG_FILE "kopv"
#define CONFIG_FILE "../etc/kpv.json"
#endif
#define LOG_DEBUG

class mc_mainprog : public QObject
{
    Q_OBJECT
public:
    explicit mc_mainprog(QObject *parent = 0);

    void Init();
    void ExecCmd();

    MC_configdevs* cfgdev;
    QTimer* checkgpio;
    QProcess *exec_process;

signals:

public slots:
    void CheckGpioProcess();
};

```

```
#endif // MC_MAINPROG_H
```

Модуль main.cpp

```
//#define QT_MESSAGELOGCONTEXT

#include <QCoreApplication>
#include "mc_mainprog.h"

#include <QFile>
#include <QtGlobal>
#include <QDir>
#include <QTextStream>
#include <QDateTime>
#include <QDebug>

QScopedPointer<QFile> m_logfile;
#if QT_VERSION < 0x050000

QtMsgHandler old_mhnd;

void messagehandler(QtMsgType type, const char *msg);

#define qInfo() qDebug()

#else

#include <QLoggingCategory>

QtMessageHandler old_mhnd;

void messagehandler(QtMsgType type, const QMessageLogContext &content, const
QString &msg);

#endif

int main(int argc, char *argv[])
{
    QCoreApplication a(argc, argv);

    /*QString log_file = QString(LOG_DIR) + QString(LOG_FILE) +
    QDateTime::currentDateTime().toString("_yyyy_MM_dd_hh_mm_ss") + ".log";
    m_logfile.reset(new QFile(log_file));
    if (!m_logfile.data()->open(QIODevice::WriteOnly|QFile::Text)) {
        qDebug() << "error open log file";
        return -1;
    }
    old_mhnd=qInstallMsgHandler(messagehandler);*/
    qInfo() << "-----programm started-----";

    mc_mainprog* mcprg = new mc_mainprog();
    mcprg->Init();
    return a.exec();
}

#if QT_VERSION < 0x050000
void messagehandler(QtMsgType type, const char *msg) {
#else
void messagehandler(QtMsgType type, const QMessageLogContext &content, const
QString &msg) {
#endif
    QTextStream out(m_logfile.data());
    QString sout = QDateTime::currentDateTime().toString("yyyy-MM-dd
```

```

hh:mm:ss.zzz ");
    switch(type) {
#ifdef QT_VERSION > 0x050000
        case QtInfoMsg: sout += "INF "; break;
#endif
        case QtDebugMsg: sout += "DBG "; break;
        case QtWarningMsg: sout += "WRN "; break;
        case QtCriticalMsg: sout += "CRT "; break;
        case QtFatalMsg: sout += "FTL "; break;
    }
#ifdef !defined(LOG_DEBUG)
    if (type!=QtDebugMsg)
#endif
    {
        out << sout << " : "<< msg << endl;
        out.flush();
    }
#ifdef QT_VERSION < 0x050000
    old_mhnd(type,msg);
#else
    old_mhnd(type,content,msg);
#endif
}

```

Модуль mc_configdevs.cpp

```

#include <QtGlobal>
#include <QFile>
#include <QDebug>
#include "mc_configdevs.h"
#include "mc_switch_imx.h"
#include "mc_sensors_imx.h"
#include "mc_sensor_type_imx.h"
#include "json.h"

#ifdef QT_VERSION < 0x050000
#define qInfo() qDebug()
#endif

MC_configdevs::MC_configdevs(QString file, QObject *parent) : QObject(parent)
{
    this->mc_switchlist = new QList<mc_switch_imx*>();
    this->mc_sensorslist = new QList<mc_sensors_imx*>();
    this->switchenc = new QMap<QString, mc_switch_imx_type>();
    this->sensorenc = new QMap<QString, mc_sensors_imx_type>();
    bool ok;
    switchenc->insert("CONST",MC_SWIMX_CONST);
    switchenc->insert("PULSE",MC_SWIMX_PULSE);
    switchenc->insert("PULSELOOP",MC_SWIMX_PULSELOOP);
    sensorenc->insert("LEVAL",MC_SIMX_LEVEL);
    sensorenc->insert("PULSE",MC_SIMX_PULSE);
    QFile* files =new QFile(file);
    files->open(QIODevice::ReadOnly);
    this->mc_conf = QtJson::parse(files->readAll(), ok).toMap();
    qInfo()<<"load config: "<<ok;
}

void MC_configdevs::ConfigMCSwitch() {
    QList<QVariant> lsswitch= mc_conf["switch"].toList();
    for(int i=0;i<lsswitch.length();i++) {
        QMap<QString,QVariant> switchel = lsswitch[i].toMap();
        this->mc_switchlist->append(new
mc_switch_imx(switchel["pin"].toString(),switchenc-
>value(switchel["type"].toString()),switchel["pulsetime"].toInt(),switchel["p
ausestime"].toInt(),switchel["name"].toString()));
    }
}

```

```

void MC_configdevs::ConfigMCSensors() {
    QList<QVariant> lssensors= mc_conf["sensors"].toList();
    for(int i=0;i<lssensors.length();i++) {
        QMap<QString,QVariant> sensorelel = lssensors[i].toMap();
        mc_switch_imx* ind = NULL;
        for(int i=0; i<this->mc_switchlist->count();i++) {
            if(this->mc_switchlist->at(i)->name
sensorelel["indicator"].toString()) {
                ind = this->mc_switchlist->at(i);
                break;
            }
        }
        this->mc_sensorslist->append(new
mc_sensors_imx(sensorelel["name"].toString(),sensorelel["pin"].toString(),sen
sorenc-
>value(sensorelel["type"].toString()),sensorelel["pulsetime"].toInt(),sensore
lel["scena"].toString(),ind));
    }
}

```

Модуль mc_configdevs.h

```

#ifndef MC_CONFIGDEVS_H
#define MC_CONFIGDEVS_H

#include <QObject>
#include <QMap>
#include <QList>
#include <QVariant>
#include "mc_sensors_imx.h"
#include "mc_switch_imx.h"

class MC_configdevs : public QObject
{
    Q_OBJECT
public:
    explicit MC_configdevs(QString file, QObject *parent = 0);
    void ConfigMCSwitch();
    void ConfigMCSensors();
    QList<mc_switch_imx*>* mc_switchlist;
    QList<mc_sensors_imx*>* mc_sensorslist;
    QMap<QString, mc_switch_imx_type*>* switchenc;
    QMap<QString, mc_sensors_imx_type*>* sensorenc;
    QMap<QString,QVariant> mc_conf;
signals:

public slots:
};

#endif // MC_CONFIGDEVS_H

```

Модуль mc_modem.cpp

```

#include "mc_modem.h"

#include <QDebug>
#include <QCoreApplication>
#include <QUuid>

#include "gsmmodem.hpp"
#include "gsmsmsreader.hpp"
#include "gsmsmssender.hpp"
#include "qtelephonynamespace.h"

mc_modem::mc_modem(const QString &serialPort,QString mainnumber, QString

```

```

reservnumber, QObject *parent) :
    QObject(parent),
    m_modem (new GsmModem(this)),           // соединение с модемом
    m_reader(new GsmSmsReader(m_modem, this)), // для чтения СМС
    m_sender(new GsmSmsSender(m_modem, this)) // для отправки СМС
{
    /*quint32 data[3];
    data[0] = 0xFF0A;
    data[1] = 0xA00A;
    data[2] = 0xBFFB;
    Send_KStatus(2,data,true);
    Send_KControl(4,0x0FA4,true);
    */
    this->init = false;
    m_modem->init(serialPort, 115200);
    this->mainnumber = mainnumber;
    this->reservnumber = reservnumber;
    connect(m_reader,          SIGNAL(messageCount(int)),          this,
    SLOT(onMessageCount(int)));
    connect(m_sender,          SIGNAL(startSendingSms(QString)),  this,
    SLOT(onStartSendingSms(QString)));
    connect(m_sender,          SIGNAL(finished(QString,int)),     this,
    SLOT(onFinished(QString,int)));
    connect(m_modem,           SIGNAL(initialized(bool)),         this,
    SLOT(onInitialized(bool)));
    connect(m_modem,           SIGNAL(operatorName(QString)),     this,
    SLOT(onOperatorName(QString)));
    connect(m_modem,           SIGNAL(signalLevels(int,int)),     this,
    SLOT(onSignalLevel(int,int)));
}

void mc_modem::send_SMS(const QString &text, const QString &recipient, bool
sr){
    QSMSMessage sms;
    sms.setValidityPeriod(-1);
    sms.setDataCodingScheme(QSMS_UCS2Alphabet);
    sms.setText(text);
    sms.setRecipient(recipient);
    sms.setStatusReportRequested(sr);
    m_sender->send(QUuid::createUuid().toString(), sms);
    qDebug() << "Send SMS: " << text;
}

void mc_modem::send_to_main(const QString &text) {
    send_SMS(text,this->mainnumber,false);
}

void mc_modem::send_to_reserve(const QString &text) {
    send_SMS(text,this->reservnumber,false);
}

void mc_modem::send_sms_to_arm(const QString &text,bool main) {
    if(main == true) {
        send_SMS(text,this->mainnumber,false);
    } else {
        send_SMS(text,this->reservnumber,false);
    }
}

void mc_modem::Send_KStatus(quint8 command,quint32* status,bool main) {
    QString data = "KSTAT";
    QString data2 = "";
    data2.sprintf("%.2X",command);
    data += data2;
    data2.sprintf("%.8X",status[0]);
    data += data2;
    data2.sprintf("%.8X",status[1]);
}

```

```

        data += data2;
        data2.sprintf("%.8X",status[2]);
        data += data2;
        send_sms_to_arm(data,main);
    }
void mc_modem::Send_KControl(quint8 command,quint32 status,bool main) {
    QString data = "KCTRL";
    QString data2 = "";
    data2.sprintf("%.2X",command);
    data += data2;
    data2.sprintf("%.8X",status);
    data += data2;
    send_sms_to_arm(data,main);
}
void mc_modem::onInitialized(bool initialized)
{
    if (initialized) {
        qDebug() << "Modem is initialized";
        m_reader->check();
        this -> init = true;
    } else {
        qDebug() << "Failed to initialize modem";
        this -> init = false;
    }
}
void mc_modem::onBalanceCash(double balance)
{
    qDebug() << "Balance:" << balance;
}
void mc_modem::onOperatorName(QString name)
{
    qDebug() << "operator name:" << name;
}
void mc_modem::onSignalLevel(int signalLevel, int errorLevel)
{
    qDebug() << QString("Signal level: %0, error level:
%1").arg(signalLevel).arg(errorLevel);
}
void mc_modem::onMessageCount(int count)
{
    Q_UNUSED(count)
    readSmsMessages(false);
}
void mc_modem::onStartSendingSms(const QString &id)
{
    qDebug() << QString("Start sending sms whit id %0").arg(id);
}

void mc_modem::onFinished(const QString &id, const int &code)
{
    if (code == QTelephony::OK) {
        qDebug() << QString("Sms with id %0 is sended")
            .arg(id);
    } else if (code == QTelephony::Delivered) {
        qDebug() << QString("Sms with id %0 is delivered")
            .arg(id);
    } else {
        qDebug() << QString("something terrible happens while sending sms
with id %0")
            .arg(id);
    }
}
void mc_modem::decodeCTRL(QString data,bool main) {
    if (data.mid(0,5)=="KCTRL") {

```

```

        emit
        Recive_KControl(data.mid(5,2).toShort(0,16),data.mid(7,8).toShort(0,16),main)
    ;
    }
}
void mc_modem::readSmsMessages(const bool deleteSmsAfterRead)
{
    QStringList smsIdList;
    for (int i = 0; i < m_reader->count(); ++i) {
        const QSMSTaggedMessage *message = m_reader->at(i);
        if (message) {
            qDebug() << QString("Message '%0' from '%1' received")
                .arg(message->message.text())
                .arg(message->message.sender());
            if(message->message.sender() == this->mainnumber)
                decodeCTRL(message->message.text(),true);
            if(message->message.sender() == this->reservnumber)
                decodeCTRL(message->message.text(),false);
            if (deleteSmsAfterRead)
                smsIdList << message->identifier;
        }
    }
    for (int i = 0; i < smsIdList.count(); ++i){
        m_reader->deleteMessage(smsIdList.at(i));
    }
}

```

Модуль mc_sensors.cpp

```

#include "mc_sensors.h"

MC_sensors::MC_sensors(int id,QObject *parent) : QObject(parent)
{
    this->pin_sensor = id;
    this->state = false;
}

bool MC_sensors::CheckSensor() {
    bool test = this->HadwareCheckSensors();
    if(test != this->state) {
        this->state = test;
        return true;
    }
    return false;
}

bool MC_sensors::HadwareCheckSensors() {
    return !this->state;
}

bool MC_sensors::GetState() {
    return this->state;
}

```

Модуль mc_sensors.h

```

#ifndef MC_SENSORS_H
#define MC_SENSORS_H

#include <QObject>

class MC_sensors : public QObject
{
    Q_OBJECT
public:
    explicit MC_sensors(int id, QObject *parent = 0);

```

```

        bool CheckSensor();
        bool GetState();
signals:

public slots:
protected:
    int pin_sensor;
    bool state;
    bool HadwareCheckSensors();
};
#endif // MC_SENSORS_H

```

Модуль mc_sensors_imx.cpp

```

#include "mc_sensors_imx.h"
#include "mc_sensor_type_imx.h"
#include <QDebug>
#include <QTimer>
mc_sensors_imx::mc_sensors_imx(QString _name,QString id,mc_sensors_imx_type
mctype,int pulsetime,QString scena,mc_switch_imx* indicator,QObject
*parent):QObject(parent)
{
    this->name = _name;
    this->pin_sensor = id;
    this->state = false;
    this->timeout = false;
    this->pulsetime = (pulsetime/TIMER_IMP_TIME);
    this->mctype = mctype;
    io.setDirection(id,true);
    this->scena = scena;
    this->indicator = indicator;
    qDebug() << "Sensor" << id << " init" <<HadwareCheckSensors();
}
bool mc_sensors_imx::GetState() {
    return this->state;
}
void mc_sensors_imx::AutoStateReset() {
    this->timeout = false;
    qDebug()<< "Reset Sensor " << this->pin_sensor;
}
bool mc_sensors_imx::CheckSensor() {
    if (this->timeout) return false;
    bool test = this->HadwareCheckSensors();
    switch(this->mctype) {
        case MC_SIMX_LEVEL:
            if(test == this->state) { // состояние изменилось
                this->state = !test;
                if (this->indicator) this->indicator->Set(this->state);
                this->timeout = true;

                QTimer::singleShot(SENSOR_RESET_TIMEOUT,this,SLOT(AutoStateReset()));
                return true;
            }
            break;
        case MC_SIMX_PULSE:
            if (test == false) // сработка
            {
                if (this->state==false) // не было сработки
                {
                    pin_cntr++;
                    if(pin_cntr >= pulsetime) {
                        pin_cntr = 0;
                        this->state = true;
                        if (this->indicator) this->indicator->Set(this->state);
                        this->timeout = true;
                    }
                }
            }
    }
}

```

```

QTimer::singleShot(SENSOR_RESET_TIMEOUT,this,SLOT(AutoStateReset()));
    return true;
}
}
} else { // нет сработки
    if (this->state==true)
    {
        this->state = false; // сброс сработки
        if (this->indicator) this->indicator->Set(this->state);
    }
    pin_cntr = 0;
}
break;
}
return false;
}
bool mc_sensors_imx::HadwareCheckSensors() {
    bool res;
    io.value(this->pin_sensor,res);
    return res;
}

```

Модуль mc_sensors_imx.h

```

#ifndef MC_SENSORS_IMX_H
#define MC_SENSORS_IMX_H

#include <QObject>
#include "mc_switch_imx.h"
#include "mc_sensor_type_imx.h"
#include "qtgpio.h"

#define SENSOR_RESET_TIMEOUT 3000

class mc_sensors_imx : public QObject
{
    Q_OBJECT
public:
    mc_sensors_imx(QString _name, QString id,mc_sensors_imx_type mctype,int
pulsetime,QString scena,mc_switch_imx* indicator, QObject *parent = 0);
    IoController io;
    mc_sensors_imx_type mctype;
    bool CheckSensor();
    bool GetState();
    QString scena;
    QString pin_sensor;
    mc_switch_imx* indicator;
    QString name;
protected:

    int pin_cntr;
    int pulsetime;
    bool state;
    int timeout;
    bool HadwareCheckSensors();
public slots:
    void AutoStateReset();
};
#endif // MC_SENSORS_IMX_H

```

Модуль mc_switch_imx.cpp

```

#include "mc_switch_imx.h"
#include <QtDebug>
mc_switch_imx::mc_switch_imx(QString id,mc_switch_imx_type mctype,int

```

```

pulsetime,int pausetime,QString name,QObject *parent):QObject(parent)
{
    this->pin_id = id;
    io.setDirection(id,false);
    this->hadwareset(0);
    this->state = 0;
    this->mctype = mctype;
    this->pulsetime = pulsetime;
    this->pausetime = pausetime;
    this->name = name;
    if(this->mctype == MC_SWIMX_PULSELOOP ) {
        this->timeshot = new QTimer();
        connect(timeshot, SIGNAL(timeout()), this, SLOT(ShotLoop()));
    }
}
void mc_switch_imx::Shot() {
    this->hadwareset(0);
    this->state = 0;
}
void mc_switch_imx::ShotLoop() {
    if(this->state == 1) {
        this->hadwareset(0);
        this->state = 0;
        this->timeshot->setInterval(this->pausetime);
    } else {
        this->hadwareset(1);
        this->state = 1;
        this->timeshot->setInterval(this->pulsetime);
    }
}
void mc_switch_imx::Set(int value) {
    switch(this->mctype) {
        case MC_SWIMX_CONST:
            this->hadwareset(value);
            this->state = value;
            break;
        case MC_SWIMX_PULSE:
            if(value == 1) {
                this->hadwareset(value);
                this->state = value;
                QTimer::singleShot(this->pulsetime,this,SLOT(Shot()));
            }
            break;
        case MC_SWIMX_PULSELOOP:
            if(value == 1) {
                this->hadwareset(value);
                this->state = value;
                this->timeshot->start(this->pulsetime);
            } else {
                this->timeshot->stop();
                this->hadwareset(value);
            }
            break;
    }
}
void mc_switch_imx::hadwareset(int value) {
    if(this->pin_id < 0) return;
    if(value == 0) {
        io.setValue(this->pin_id,false);
    } else {
        io.setValue(this->pin_id,true);
    }
    qDebug() <<"IMX Set Pin: "<< this->pin_id << " Value: " << value;
}

```

```
int mc_switch_imx::Get() {
    return this->state;
}
```

Модуль mc_switch_imx.h

```
#ifndef MC_SWITCH_IMX_H
#define MC_SWITCH_IMX_H

#include <QObject>
#include <QTimer>
#include "qtgpio.h"
#include "mc_switch_imx_type.h"

class mc_switch_imx : public QObject
{
    Q_OBJECT
public:
    explicit mc_switch_imx(QString id,mc_switch_imx_type mctype,int
    pulsetime,int pausetime,QString name,QObject *parent = 0);
    IoController io;
    mc_switch_imx_type mctype;
    int pulsetime;
    int pausetime;
    void Set(int value);
    int Get();
    QString name;
private slots:
    void Shot();
    void ShotLoop();

protected:
    QString pin_id;
    int state;
    QTimer* timeshot;
    void hadwareset(int value);

};

#endif // MC_SWITCH_IMX_H
```

Модуль qtgpio.cpp

```
// =====
#include <QDebug>
#include <QString>
#include <QFile>
#include <QTextStream>

#include "qtgpio.h"

//! Port directory and file constants
#ifdef Q_OS_UNIX
    QString IoController::baseGPIOPath_ = "/sys/class/gpio/";
#elif defined(Q_OS_WIN)
    QString IoController::baseGPIOPath_ = "./gpio/";
#endif

QString IoController::gpioExport_ = "export";
QString IoController::gpioUnexport_ = "unexport";
QString IoController::gpioDirection_ = "direction";
QString IoController::gpioValue_ = "value";
IoController::IoController(QObject *parent) : QObject(parent)
{
}
```

```

bool IoController::writeToGPIOFile( const QString file, QString outputData )
{
    QFile exportFile( file );

    if ( exportFile.open( QIODevice::WriteOnly ) )
    {
        QTextStream exportStream( &exportFile );

        exportStream << outputData;

        exportFile.close();

        return true;
    }
    //qDebug() << "Error writing " << file;

    return false;
}

bool IoController::readFromGPIOFile( const QString file, QString &inputData )
{
    QFile inputFile( file );

    if ( inputFile.open( QIODevice::ReadOnly ) )
    {
        QTextStream inputStream( &inputFile );

        inputStream >> inputData;

        inputFile.close();

        return true;
    }
    //qDebug() << "Error reading " << file;
    return false;
}

bool IoController::exportGPIO( int pinNumber )
{
    return writeToGPIOFile( baseGPIOPath_ + gpioExport_,
        QString::number( pinNumber ) );
}

bool IoController::unexportGPIO( int pinNumber )
{
    return writeToGPIOFile( baseGPIOPath_ + gpioUnexport_,
        QString::number( pinNumber ) );
}

bool IoController::setDirection( QString pinName, bool read )
{
    bool result = false;
    QString path = baseGPIOPath_ + pinName + "/" + gpioDirection_;

    if ( read )
        result = writeToGPIOFile( path, "in" );
    else
        result = writeToGPIOFile( path, "out" );

    return result;
}

bool IoController::setValue( QString pinName, bool value )
{
    bool result = false;
    QString path = baseGPIOPath_ + pinName + "/" + gpioValue_;

```

```

    if ( value )
        result = writeToGPIOFile( path, "1" );
    else
        result = writeToGPIOFile( path, "0" );

    return result;
}
bool IoController::value( QString pinName, bool &value )
{
    bool result = false;
    QString data = "";
    QString path = baseGPIOPath_ + pinName + "/" + gpioValue_;

    result = readFromGPIOFile( path, data );

    if ( result )
        value = (data == "1");

    return result;
}

```

Модуль qtgpio.h

```

#ifndef QTGPIO_H
#define QTGPIO_H

#include <QObject>

class IoController : public QObject
{
    Q_OBJECT

public:
    explicit IoController(QObject *parent = 0);

    ///! Export Port for using
    bool exportGPIO( int pinNumber );
    ///! Unexport Port
    bool unexportGPIO( int pinNumber );
    ///! Set the direction (read/write) of the specified port
    bool setDirection( QString pinName, bool read );
    ///! Write to the specified port
    bool setValue( QString pinName, bool value );
    ///! Read from the specified port
    bool value( QString pinName, bool &value );

signals:

public slots:

private:
    static QString baseGPIOPath_;
    static QString gpioExport_;
    static QString gpioUnexport_;
    static QString gpioDirection_;
    static QString gpioValue_;

    ///! Helper function for writing to the GPIO directory
    bool writeToGPIOFile(const QString file, QString outputData);
    ///! Helper function for reading from the GPIO directory
    bool readFromGPIOFile(const QString file, QString &inputData);
};

#endif // QTGPIO_H

```