

Министерство образования Республики Беларусь

Учреждение образования
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Факультет компьютерных систем и сетей
Кафедра электронных вычислительных средств

К защите допустить:
Заведующий кафедрой ЭВС
_____ И. С. Азаров

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
к дипломному проекту
на тему

**СИСТЕМА ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ НА ОСНОВЕ
МОДЕЛИ СЛУХОВОГО ВОСПРИЯТИЯ И КОМПРЕССИИ
ДИНАМИЧЕСКОГО ДИАПАЗОНА СИГНАЛА**

БГУИР ДП 1-40 02 02 01 016 ПЗ

Студент	Н.А. Семенюк
Руководитель	М.И. Порхун
Консультанты: <i>от кафедры ЭВС</i>	М.И. Порхун
<i>по экономической части</i>	Т.А. Рыковская
Нормоконтролер	Д. С. Лихачёв
Рецензент	

Минск 2021

Министерство образования Республики Беларусь

Учреждение образования

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Факультет	<u>КСиС</u>	Кафедра	<u>ЭВС</u>
Специальность	<u>1-40 02 02</u>	Специализация	<u>01</u>

УТВЕРЖДАЮ
Зав. кафедрой ЭВС

_____ И.С. Азаров

«29» марта 2021 г.

ЗАДАНИЕ
по дипломному проекту студента

Семенюк Натальи Александровны
(фамилия, имя, отчество)

1. Тема проекта: *Система повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала*

утверждена приказом по университету от « 29 » марта 2021 г. № _____

2. Срок сдачи студентом законченного проекта 15 июня 2021 года

3. Исходные данные к проекту

3.1 Назначение разработки: повышение разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала.

3.2 Технические характеристики:

- минимальная частота дискретизации: 12 кГц;
- минимальное количество бит на отсчёт: 16;
- число каналов во входном звуковом сигнале: 1;
- минимальное количество каналов банка фильтров: 20.

4. Содержание пояснительной записки (перечень подлежащих разработке вопросов)

4.1 Введение. 4.2 Обзор существующих методов повышения разборчивости речи на основе модели слухового восприятия. 4.3 Анализ технического задания. 4.4 Разработка структуры и алгоритма работы системы повышения разборчивости речи. 4.5 Программная реализация системы повышения разборчивости речи. 4.6 Техничко-экономическое обоснование целесообразности проектирования и внедрения системы повышения разборчивости речи. 4.7 Анализ результатов тестирования системы. 4.8 Заключение. 4.9 Список используемых источников.

5. Перечень графического материала (с указанием названия чертежей и их количества в пересчёте на формат А1)

5.1 Структура системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала – 1 лист формата А1 (плакат).

5.2 Схема программы реализации банка фильтров для системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала – 1 лист формата А1.

5.3 Схема программы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала – 1 лист формата А1.

5.4 Настройка системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала по аудиограмме – 1 лист формата А1 (плакат).

5.5 Экспериментальное исследование эффективности работы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала по аудиограмме – 1 лист формата А1 (плакат).

5.6 Схема данных для экспериментального исследования эффективности работы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала по аудиограмме – 1 лист формата А1.

6. Содержание задания по технико-экономическому обоснованию

1. Расчет затрат на проведение исследования системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала.

2. Оценка уровня качества разработанной системы.

Задание выдал: _____ / Т.А. Рыковская/

КАЛЕНДАРНЫЙ ПЛАН

Наименование этапов дипломного проекта (работы)	Объём этапа в %	Срок выполнения этапа	Примечание
I этап – п.4.1 – 4.4, п.5.1 – 5.3	40	20.04.21	
II этап – п.4.5, п.5.4	20	28.04.21	
III этап – п.4.6, 4.7., п.5.5, 5.6.	40	12.05.21	
Нормоконтроль		17.05.21 – 21.05.21	
Рабочая комиссия		24.05.21 – 28.05.21	
Рецензирование		01.06.21 – 11.06.21	
Защита		15.06.21 – 30.06.21 (в соответствии с графиком заседаний ГЭК)	

Дата выдачи задания « » марта 2021 г. Руководитель _____ /М. И. Порхун/

Задание принял к исполнению _____ /Н.А. Семенюк/

РЕФЕРАТ

СИСТЕМА ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ НА ОСНОВЕ МОДЕЛИ СЛУХОВОГО ВОСПРИЯТИЯ И КОМПРЕССИИ ДИНАМИЧЕСКОГО ДИАПАЗОНА СИГНАЛА: дипломный проект / Н. А. Семенюк. – Минск: БГУИР, 2021, – п.з. – 72 с.,чертежей (плакатов) – 6 л. формата А1.

Основные задачи данной дипломной работы заключаются в исследовании методов повышения разборчивости речи и разработке системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала.

Важной особенностью системы является применение в ней банка гамматон-фильтров, имитирующего механизм частотной декомпозиции звука в ухе. Другой особенностью системы является её автоматическая настройка по аудиограмме слабослышащего человека. Недостатком системы является длинная импульсная характеристика банка фильтров в области низких частот, что увеличивает количество вычислительных операций.

Технико-экономическое обоснование определило целесообразность проектирования и внедрения системы повышения разборчивости речи.

Результаты тестирования системы, а также сравнение с уже существующими аналогами позволило заключить, что данная система обладает достаточным уровнем конкурентоспособности в своей области применения.

Ключевые слова: повышение разборчивости речи, гамматон-фильтр, компрессор динамического диапазона, аудиограмма.

СОДЕРЖАНИЕ

Введение	7
1 Обзор существующих методов повышения разборчивости речи на основе модели слухового восприятия	10
1.1 Повышение разборчивости речи в шумных реверберационных средах, основанный на увеличении энергии речевого сигнала и «обратной» фильтрации	10
1.2 Метод повышения разборчивости речи на основе новой методики уменьшения частоты для пользователей слуховых аппаратов, говорящих на северокитайском языке	12
1.3 Метод повышения разборчивости речи на основе удаления шума с помощью метода адаптивной фильтрации для людей страдающих нейросенсорной тугоухостью	14
1.4 Моделирование потери слуха с использованием компрессионных слуховых гаммачирп-фильтров	15
2 Анализ технического задания	17
2.1 Анализ требований к алгоритму	17
2.2 Анализ требований к программной реализации	17
2.3 Выбор и обоснование метода решения задачи	17
3 Разработка структуры и алгоритма работы системы повышения разборчивости речи.	19
3.1 Слух и слуховое восприятие	19
3.1.1 Строение слуховой системы	19
3.1.2 Возникновение слуховых ощущений	22
3.1.3 Влияние акустических условий на восприятие речи	23
3.2 Аудиометрия и аудиограмма	24
3.3 Принцип работы системы повышения разборчивости речи	26
3.3.1 Банк гамматон-фильтров	26
3.3.2 Настройка компрессора динамического диапазона (КДД)	30
3.3.3 Реализация КДД	33
3.4 Принцип работы системы моделирования потери слуха	35
4 Программная реализация системы повышения разборчивости речи	37
4.1 Описание пакета MATLAB	37
4.2 Программная реализация системы повышения разборчивости речи	38
4.3 Описание вспомогательных функций	39
4.3.1 Функция gammatoneFast	39
4.3.2 Функция bark2frq	40
4.3.3 Функция ttest	40
4.3.4 Функция iso226_rev	42

5 Технико-экономическое обоснование целесообразности проектирования и внедрения системы повышения разборчивости речи.....	43
5.1 Характеристика системы повышения разборчивости речи	43
5.2 Составление плана и определение трудоемкости выполнения научно-исследовательской работы.....	43
5.3 Расчёт цены научно-технической продукции	46
5.3.1 Расчёт затрат по статье «Топливо-энергетические ресурсы для научно-экспериментальных целей»	46
5.3.2 Расчёт затрат по статье «Спецоборудование для научных (экспериментальных) работ»	47
5.3.3 Расчёт затрат по статье «Основная заработная плата научно-технического персонала»	48
5.3.4 Дополнительная заработная плата научно-технического персонала	48
5.3.5 Отчисления на социальные нужды (отчисления в фонд социальной защиты населения и обязательное страхование).....	48
5.3.6 Накладные расходы	49
5.3.7 Полная себестоимость	49
5.4 Оценка уровня качества научно-технического результата	49
6 Анализ результатов тестирования системы.....	52
6.1 Подготовка данных для проведения эксперимента	52
6.2 Описание эксперимента	54
6.3 Результаты эксперимента.....	55
Заключение	57
Список использованных источников	59
Приложение А	61
А.1 Программная реализация алгоритм системы повышения разборчивости речи	61
А.2 Программная реализация алгоритм определения значимости результатов	67
Приложение Б.....	69

ВВЕДЕНИЕ

Звук возникает в результате вибрации любых объектов с определенной частотой. Среда, в которой распространяется звук, называется звуковым полем. Звуки многих музыкальных инструментов являются результатом колебаний натянутых струн (скрипка, гитара, фортепиано). Звук голоса человека – это результат колебаний голосовых складок в гортани под воздействием выдыхаемой струи воздуха. Эти звуковые колебания распространяются по воздуху и проникают в наше ухо. Структуры наружного, среднего и внутреннего уха обеспечивают передачу звуковых колебаний волосковым клеткам. После чего формируются электрические импульсы, которые передаются по слуховому нерву к подкорковым слуховым центрам мозга, а затем в слуховые корковые центры, где осуществляется основная обработка речи, узнавание, запоминание, интерпретация речевых и неречевых сигналов, и, наконец, формируются различные слуховые ощущения [1].

Слух и понимание являются разными аспектами восприятия акустических сигналов. Для того чтобы акустический сигнал (речь) стал ясным, необходимо, чтобы он был слышен. Если сигнал не слышен, понять его невозможно. При этом если сигнал слышен, не означает, что он будет понят. Таким образом, разборчивость зависит от того, способен ли человеческий мозг распознавать слышимые звуковые сигналы. Часто бывают ситуации, когда человек может слышать акустический сигнал, но не понимает той информации, которую несет этот сигнал. Следовательно, есть существенная разница между слышимостью и разборчивостью. С проблемой непонимания сталкиваются слабослышащие люди (особенно те, у кого есть патологии внутреннего уха). Они оказываются в ситуации, когда не могут расшифровать звуковые сигналы, используя имеющуюся акустическую информацию [2].

По прогнозам Всемирной организации здравоохранения более 1,1 миллиарда людей молодого возраста подвергаются риску потери слуха из-за прослушивания музыки на слишком высоком, опасном для слуха уровне громкости. Уже сейчас более 5% населения мира нуждаются в реабилитации для решения проблемы инвалидизирующей потери слуха (т.е. потеря слуха в слышащем лучше ухе превышает 35 децибел). По оценкам, к 2050 г. более 10% будут иметь инвалидизирующую потерю слуха.

В связи с этим существует задача создания методов коррекции слуха. Ученые с разных стран мира модифицируют имеющиеся методы и создают новые для решения проблемы потери слуха. Условно системы, помогающие решить данную проблему, можно разделить на следующие категории: системы моделирования потери слуха и, непосредственно, сами системы повышения

разборчивости речи и коррекции слуха. Использование систем моделирования потери слуха может быть ценным инструментом для разработки стратегий обработки сигналов для слуховых аппаратов, поскольку в предварительных экспериментах могут принимать участие люди с нормальным слухом, что часто проще и дешевле.

Основные задачи, поставленными в данной дипломной работе:

- разработать систему повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала;
- определить, является ли разрабатываемая система пригодной для решения задачи повышения разборчивости путем тестирования с помощью модели потери слуха человека на людях с нормальным слухом.

Проектирование системы повышения разборчивости речи проводилось в несколько этапов, которые отражены в структуре пояснительной записки данного дипломного проекта.

Первым этапом работы является ознакомление с используемыми в настоящее время методами повышения разборчивости речи на основе модели слухового восприятия. Для этого проводился поиск необходимого теоретического материала и их анализ. Результат этих действий приведен в первом разделе пояснительной записки. В этом разделе описан краткий обзор каждого метода.

Перед осуществлением проектирования алгоритма работы системы, проводился анализ технического задания дипломного проекта, в целях выявления метода, составляющего основу алгоритма работы системы. Данный этап отражает второй раздел пояснительной записки.

На основе выбранного метода повышения разборчивости речи было проведено проектирование системы. Этот этап описывается в третьем разделе, который содержит описание вводной информации о слухе и слуховом восприятии, аудиометрии и аудиограмме. Понимание этих понятий необходимо для описания принципа работы системы, которое представлено также в третьем разделе.

Для проверки работоспособности алгоритма работы системы проводилась его программная реализация. Особенности этого процесса приведены в четвертом разделе. Дополнительно этот раздел содержит справочную информацию о некоторых используемых функциях, которая облегчает понимание исходного кода, приведенного в приложении А.1.

Технико-экономическое обоснование целесообразности проектирования и внедрения системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала приведено в пятом разделе пояснительной записки.

Заключительным этапом является тестирование разработанной системы. После проведения тестирования необходимо провести проверку значимости полученных результатов для исключения возможности получения случайных результатов из-за небольшой выборки. Код этой функции отражен в приложении А.2. Результаты тестирования системы приведены в шестом разделе.

В заключении был проведен анализ результатов проектирования системы в целом, а также сделаны выводы о проделанной работе.

Отчет о проверке на заимствование представлен в приложении Б.

1 ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ НА ОСНОВЕ МОДЕЛИ СЛУХОВОГО ВОСПРИЯТИЯ

1.1 Повышение разборчивости речи в шумных реверберационных средах, основанный на увеличении энергии речевого сигнала и «обратной» фильтрации

В исследовании [3] предлагается метод предварительной обработки, который включает в себя «обратную» фильтрацию для улучшения разборчивости речи в средах с реверберацией и с фоновым шумом. Для использования в таких условиях метод улучшения разборчивости речи с перераспределением энергии был модифицирован, а для достижения лучшей производительности был разработан быстрый «обратный» фильтр на основе модели слухового восприятия.

На первом этапе для входного сигнала применяется частотно-временная декомпозиция с использованием банка гамматон-фильтров для получения кратковременного чистого вокализованного фрейма. Затем из фрейма необходимо получить информацию о речевой активности. Далее фрейм проходит через модуль предварительной обработки и синтеза, который состоит из двух блоков. В блоке предварительной обработки и синтеза речи на этапе усиления используется модифицированный метод улучшения разборчивости речи, основанный на увеличении энергии речевого сигнала. Затем на этапе выравнивания используется быстрая обратная фильтрация на основе гамматон-фильтра для предварительной компенсации искажений в канале передачи. Сигнал с выхода блока предварительной обработки и синтеза используется в качестве входного сигнала для вывода в звуковоспроизводящее устройство. Сигнал с искажениями (фоновым шумом) затем записывается микрофоном, и затем снова применяются частотно-временная декомпозиция на основе гамматон-фильтров для получения кратковременного фрейма с искажением.

Далее применяется модуль оценки спектральной плотности мощности для дальнейшей оценки энергии фонового шума. Наконец, функция усиления вычисляется алгоритмом увеличения разборчивости речи на основе увеличения энергии речевого сигнала, а обратные подфильтры получаются алгоритмом быстрой «обратной» фильтрации с использованием гамматон-фильтра. Оба параметра используются для настройки предварительной обработки для получения наилучшей разборчивости речи.

Структура алгоритма метода продемонстрирована на рисунке 1.1.

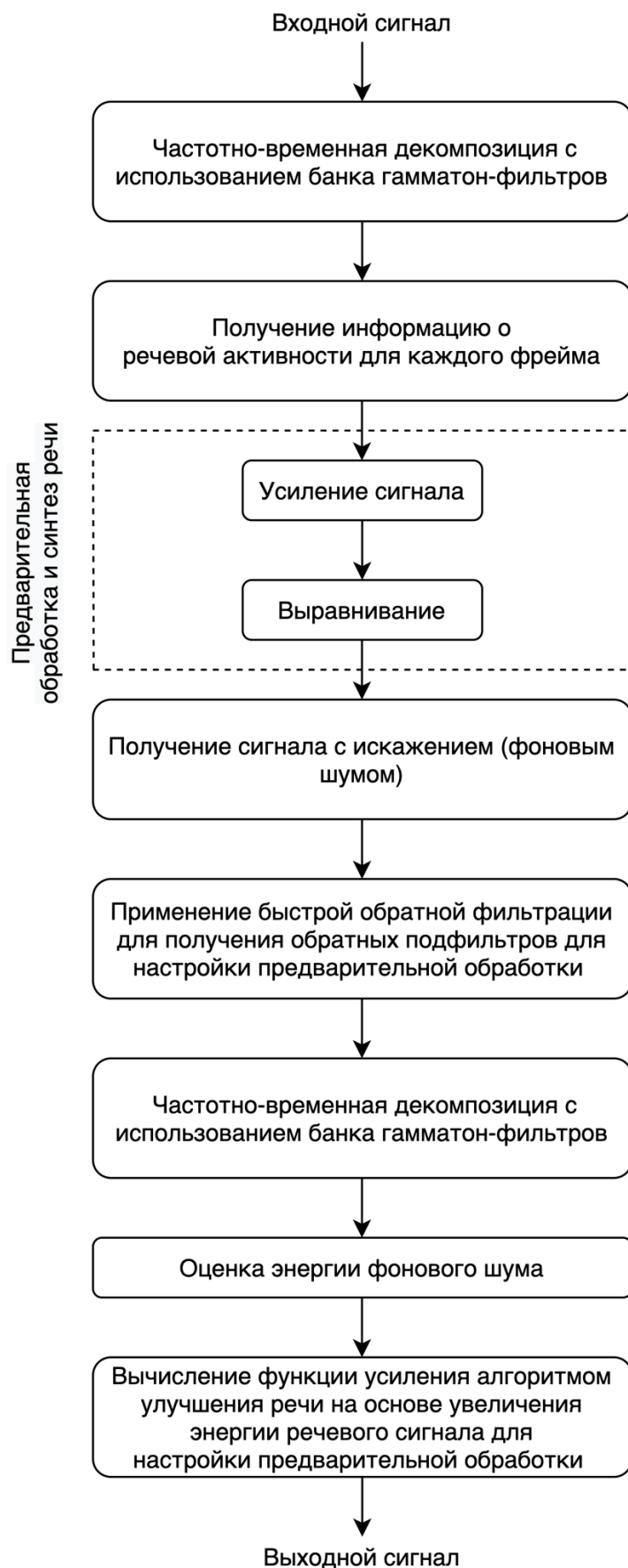


Рисунок 1.1 – Структура алгоритма метода из работы [3]

По итогам исследования был проведен эксперимент в различных шумных, реверберирующих средах, и результаты испытаний показали, что обработка звука предложенным методом значительно способствует увеличению разборчивости речи. По сравнению с отдельными методами улучшения разборчивости речи, основанных на увеличении энергии речевого сигнала и быстрой «обратной» фильтрации с использованием гамматон-фильтра, комбинированный метод может стабильно снижать искажения в речевом сигнале при наличии различных искажений в сигнале. Кроме того, субъективные тесты на прослушивание подтвердили стабильность предложенного метода, а его средний показатель увеличения разборчивости речи выше, чем у некоторых эталонных алгоритмов [3].

1.2 Метод повышения разборчивости речи на основе новой методики уменьшения частоты для пользователей слуховых аппаратов, говорящих на северокитайском языке

В работе [4], выполненной исследовательским центром инноваций в области информационных технологий в Тайване, предлагается новая схема частотной транспозиции для слуховых аппаратов, говорящих на северокитайском языке, которая основана на первоначальной идее техники частотной транспозиции, но включает в себя много новых деталей. Изначальная идея частотной транспозиции заключается в переносе высокочастотных компонент (исходная частотная полоса) в область низких частот (целевая полоса), а также в добавлении их к необработанным низкочастотным компонентам сигнала.

Предложенный метод имеет несколько особенностей. Во-первых, северокитайские согласные появляются только в начальной позиции слога. Было замечено, что гласные все еще узнаваемы даже без высокочастотной информации, и ненужная частотная транспозиция может ухудшить распознавание гласных. Поэтому использовался классификатор согласных/гласных звуков на основе нейронной сети. Частотная транспозиция выполнялась только для согласных звуков. Во-вторых, в этом методе использовалась адаптивная исходная полоса для захвата наиболее важной информации в спектре сигнала.

Структура предложенного алгоритма частотной транспозиции представлена на рисунке 1.2.

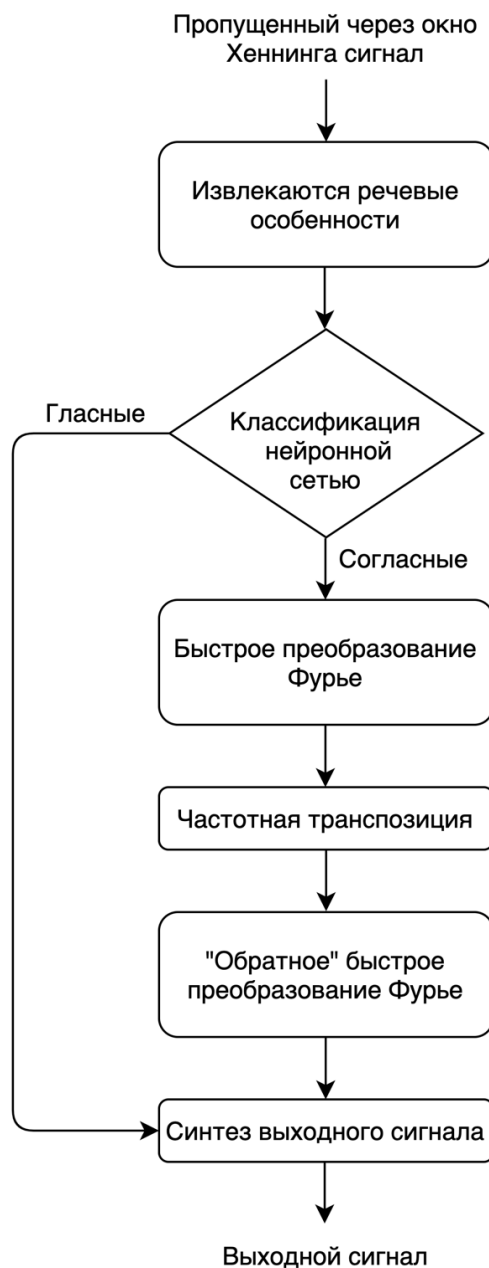


Рисунок 1.2 – Структура алгоритма метода, описанного в [4]

Для исследования эффективности метода был проведён эксперимент. Тестовые речевые сигналы были разделены на две группы. В первой группе звук обрабатывался методом моделирования потери слуха, основанным на использовании ФНЧ с частотой среза 1500 Гц (имитация тяжелой потери слуха). Во второй группе сигналы обрабатывались сначала предложенным методом коррекции слуха, а затем методом моделирования потери слуха (на основе ФНЧ). Результаты показали, что применение предложенного метода увеличивает разборчивость речи в среднем на 8%. Однако данный метод имеет явный недостаток – ориентированность на определенную группу языков, схожих с северокитайским [4].

1.3 Метод повышения разборчивости речи на основе удаления шума с помощью метода адаптивной фильтрации для людей страдающих нейросенсорной тугоухостью

Одним из существующих методов обработки звукового сигнала для слуховых аппаратов является метод, описанный в работе [5]. Метод основан на удалении из речевого сигнала шума с использованием адаптивной фильтрации.

Структура предложенного алгоритма адаптивной фильтрации для людей страдающих нейросенсорной тугоухостью представлена на рисунке 1.3.

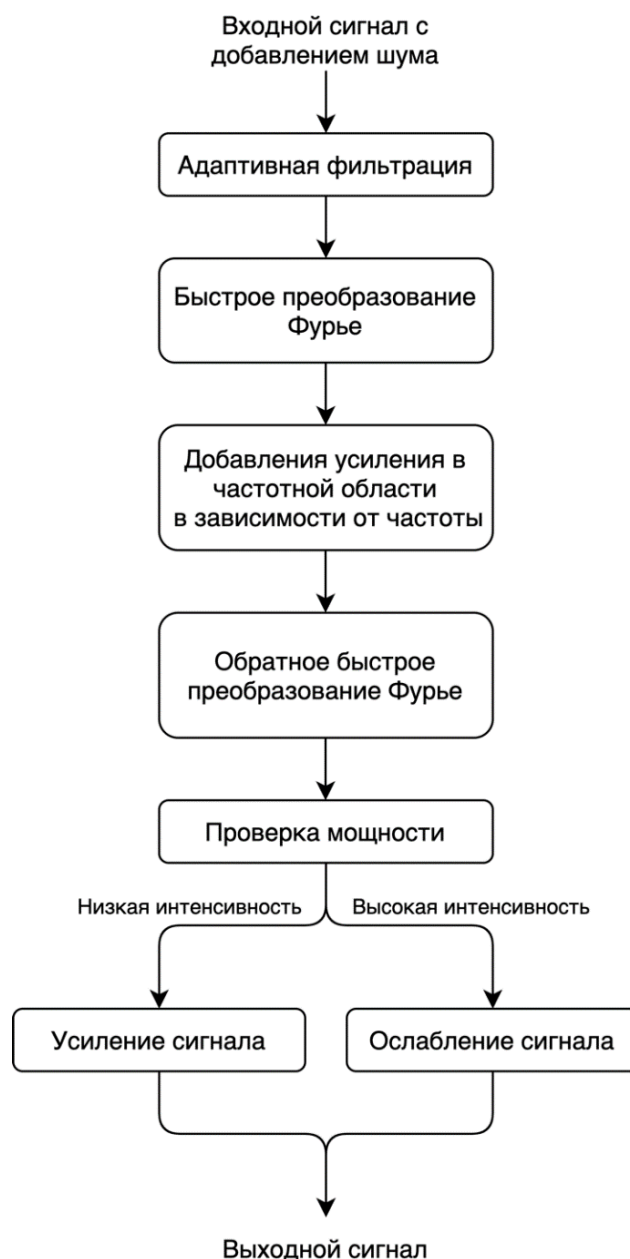


Рисунок 1.3 – Структура алгоритма метода, описанного в [5]

Цифровые слуховые аппараты являются самонастраивающимися и гибкими. Таким образом, звук, который он передает, соответствует конкретной модели потери слуха. Для реализации данного алгоритма на вход системы подается звуковой сигнал с добавлением белого шума для того, чтобы имитировать эффект многих случайных процессов, происходящих в природе. Далее сигнал подвергается адаптивной фильтрации, которая эффективно устраняет аддитивный шум путем изменения коэффициентов фильтра с течением времени и адаптируется к изменяющимся характеристикам сигнала в соответствии с алгоритмом оптимизации. Сигнал преобразуется в частотную область для упрощения вычислений с использованием быстрого преобразования Фурье. После добавления усиления в частотной области на выбранных частотах сигнал возвращается во временную область с помощью обратного быстрого преобразования Фурье. Усиление выполняется в соответствии с характеристиками слуха конкретного человека. Это позволяет удобно и детально контролировать частотную характеристику слухового аппарата. Наконец, обработанный сигнал проходит проверку мощности, чтобы убедиться, что мощность сигнала меньше болевого порога. Если сигнал имеет слишком низкую интенсивность, то он усиливается, чтобы находиться в диапазоне слышимости человека. Если сигнал имеет интенсивность выше болевого порога, тогда она ослабляется.

По итогам исследования было выявлено, что адаптивная фильтрация обеспечивает лучшее шумоподавление, чем вейвлет-фильтрация. Также с помощью данного метода эффективно удаляется шум. Кроме того, каждой полосе частот может быть задана разная величина усиления. Следовательно, система является гибкой [5].

1.4 Моделирование потери слуха с использованием компрессионных слуховых гаммачирп-фильтров

Важными инструментами для решения задачи повышения разборчивости речи являются системы, моделирующие потерю слуха. Данные методы могут применяться при проверке эффективности методов коррекции слуха без непосредственного участия слабослышащих людей, что часто делает процесс проведения тестирования проще и дешевле.

Предложенный в работе [6] метод представляет собой систему моделирования потери слуха, которая сочетает психоакустически-мотивированный компрессионный банк гаммачирп-фильтров с традиционными алгоритмами обработки сигналов для моделирования потери слуха.

Структура предложенного в работе [6] алгоритма моделирования потери слуха с использованием компрессионных слуховых гаммачирп-фильтров представлена на рисунке 1.4.

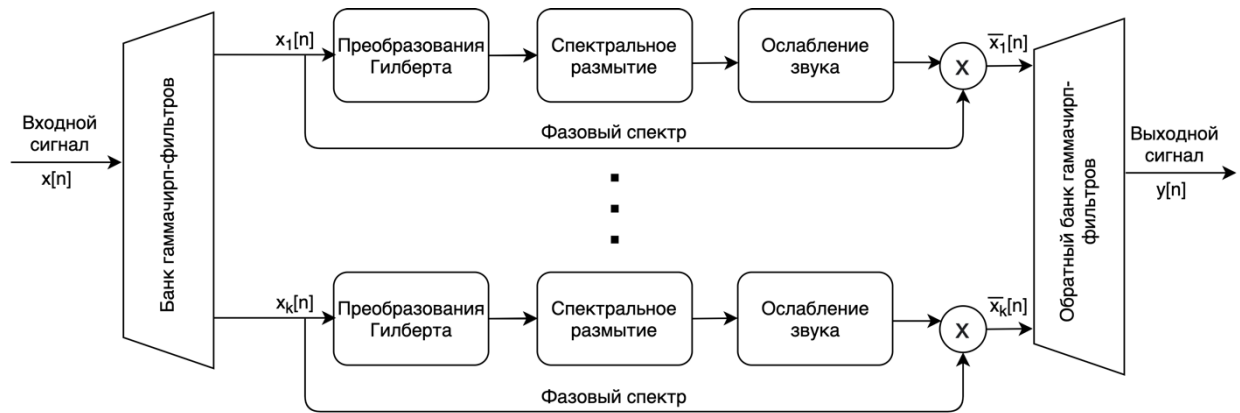


Рисунок 1.4 – Структура алгоритма метода, описанного в [6]

Входной сигнал декомпозируется на 26 субполос, а затем в каждой субполосе применяются методы цифровой обработки сигналов, имитирующие ослабление частотного разрешения уха, увеличение порога слышимости, а также частотно-зависимого ослабления сигнала. Указанные на рисунке 1.4 этапы обработки звукового сигнала необходимы для демонстрации вышеперечисленных аспектов тугоухости нормально слышащему человеку.

Огибающая и информация о фазе в каждом канале разделяются с помощью преобразования Гилберта. Спектральное размытие применяется для имитации ослабления частотного разрешения уха и частотно-зависимого ослабления сигнала – для изменения громкости и увеличение порога слышимости.

Результаты моделирования работы данного метода демонстрируют, что спектральная детализация в выходном звуковом сигнале размыта (эффект ослабления частотного разрешения уха), а сам сигнал имеет меньшую энергию в результате частотно-зависимого ослабления. Соответственно, такая система может имитировать различные перцептивные аспекты нейросенсорной тугоухости [6]

2 АНАЛИЗ ТЕХНИЧЕСКОГО ЗАДАНИЯ

2.1 Анализ требований к алгоритму

Основными требованиями к алгоритму работы системы повышения разборчивости речи на основе модели слухового восприятия являются:

- независимость от степени потери слуха;
- декомпозиция сигнала на основе «слухового» банка фильтров;
- выбор менее ресурсоёмких алгоритмов цифровой обработки сигналов.

Для удовлетворения этих требований необходимо определить способы их решения.

Независимость от степени потери слуха для системы повышения разборчивости речи является важным требованием, т. к. с ним связана возможность использования разработанной системы людьми с разной степенью потери слуха. Данное требование можно удовлетворить путем применения системы, которая автоматически настраивается согласно характеристикам слуха человека.

Декомпозиция сигнала на основе «слухового» банка фильтров также является важным требованием в разработанной системе. Такой банк согласован с работой человеческого уха и имитирует механизм частотной декомпозиции звука в ухе.

Выбор менее ресурсоёмких алгоритмов цифровой обработки сигналов обеспечит достаточно быструю работу программной реализации.

2.2 Анализ требований к программной реализации

Техническим заданием предъявляются следующие требования к программной реализации системы повышения разборчивости речи:

- минимальная частота дискретизации: 12 кГц;
- число каналов во входном звуковом сигнале: 1;
- минимальное количество каналов банка фильтров: 20.

Требования были учтены при выполнении программной реализации проекта системы.

2.3 Выбор и обоснование метода решения задачи

Основная задача разрабатываемой системы – компенсировать нарушение слуховой функции. Тугоухий человек, как правило, имеет повышенный порог слышимости, что препятствует восприятию звуков низкой

интенсивности. Однако восприятие звуков высокой интенсивности близко к тому, как их воспринимают люди с нормальным слухом. Таким образом, одной из основных функций системы является отображения широкого динамического диапазона речевого сигнала в суженный динамический диапазон остаточного слуха. Кроме того, нарушение слуха обычно имеет частотно-зависимый характер, поэтому для разных частотных диапазонов требуются разные степени компенсации. Решение этой проблемы заключается в построении многоканальной системы, такой как банк фильтров, с различной степенью компенсации в каждом канале. Устройство, выполняющее описанную обработку сигнала, называется многоканальным компрессором динамического диапазона (КДД).

Одной из проблем, связанных с проектированием многоканального компрессора, является согласование частотного разрешения цифровой системы со слуховой системой человека. Например, многие процедуры настройки слухового аппарата, основанные на регулировке громкости в поврежденном ухе, включают в себя проведение частотного анализа, согласованного со слуховой системой. Частотный анализ, используемый в цифровой обработке сигналов, такой как дискретное преобразование Фурье, имеет равномерное разрешение во всем диапазоне частот. Частотное же разрешение слуховой системы человека более точно моделируется неравнополосным банком фильтров. Исходя из того, что слуховая система человека более чувствительна к изменениям сигнала в низкочастотной области, банк фильтров должен обладать следующим свойством: ширина полос должна возрастать с увеличением частоты [2].

Основываясь на приведенной выше информации и предъявленным требованиям к алгоритму целесообразно выбрать подход к построению системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала с использованием неравнополосного банка гамматон-фильтров, имитирующего механизм частотной декомпозиции звука в ухе. Важной особенностью системы является её автоматическая настройка по аудиограмме слабослышащего человека. Такая система способна справляться с задачей повышения разборчивости лучше, чем ряд существующих методов.

3 РАЗРАБОТКА СТРУКТУРЫ И АЛГОРИТМА РАБОТЫ СИСТЕМЫ ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ.

3.1 Слух и слуховое восприятие

Всю звуковую информацию, которую человек получает из внешнего мира (это около 25% от общей), он распознает с помощью слуховой системы и работы высших отделов мозга, переводит в мир своих ощущений, и принимает решения, как надо на нее реагировать. Прежде чем приступить к разработке алгоритма работы метода повышения разборчивости речи на основе слухового восприятия и компрессии динамического диапазона сигнала, необходимо коротко остановиться на механизме работы и строении слуховой системы.

3.1.1 Строение слуховой системы

Слуховая система является своеобразным приемником информации. Она состоит из периферической части и высших отделов слуховой системы. Наиболее изучены процессы преобразования звуковых сигналов в периферической части слухового анализатора.

Обычно периферическую слуховую систему делят на три части:

- наружное;
- среднее;
- внутреннее ухо [1].

Наружное ухо состоит из ушной раковины и слухового канала, заканчивающегося тонкой мембраной, называемой барабанной перепонкой. Наружные уши и голова – это компоненты внешней акустической антенны, которая соединяет барабанную перепонку с внешним звуковым пространством [8].

Основные функции наружных ушей:

- пространственное восприятие;
- усиление звуковой энергии, особенно в области средних и высоких частот;
- локализация звукового источника;
- защитная функция.

Слуховой канал представляет собой изогнутую цилиндрическую трубку длиной 22,5 мм. В этой области частот он существенно усиливает звуковой сигнал, и именно здесь находится область максимальной чувствительности слуха.

Барабанная перепонка – тонкая пленка толщиной 74 мкм, имеет вид конуса, обращенного острием в сторону среднего уха. На низких частотах она движется как поршень, на более высоких - на ней образуется сложная система узловых линий, что также имеет значение для усиления звука.

Среднее ухо – заполненная воздухом полость, соединенная евстахиевой с носоглоткой трубой для выравнивания атмосферного давления. При изменении атмосферного давления воздух может входить или выходить из среднего уха, поэтому барабанная перепонка не реагирует на медленные изменения статического давления. В среднем ухе находятся три маленькие слуховые косточки: молоточек, наковальня и стремечко. Молоточек прикреплен к барабанной перепонке одним концом, вторым он соприкасается с наковальней, которая при помощи маленькой связки соединена со стремечком. Основание стремечка соединено с овальным окном во внутреннее ухо.

Среднее ухо выполняет следующие функции:

- согласование импеданса воздушной среды с жидкой средой улитки внутреннего уха;
- защита от громких звуков (акустический рефлекс);
- усиление передаваемого во внутреннее ухо звукового давления.

Внутреннее ухо находится в лабиринте каналов в височной кости, и включает в себя орган равновесия (вестибулярный аппарат) и улитку.

Улитка выполняет основную роль в слуховом восприятии. Она представляет собой трубку переменного сечения, свернутую три раза подобно хвосту змеи. В развернутом состоянии она имеет длину 3,5 см. Внутри улитка имеет чрезвычайно сложную структуру. По всей длине она разделена двумя мембранами на три полости: лестница преддверия, срединная полость и барабанная лестница. Сверху срединная полость закрыта мембраной Рейсснера, снизу – базилярной мембраной. Все полости заполнены жидкостью. На внутренней стороне базилярной мембраны находится кортиев орган, а в нем - специализированные слуховые рецепторы – волосковые клетки. В поперечном направлении кортиев орган состоит из одного ряда внутренних волосковых клеток и трех рядов наружных волосковых клеток. Между ними образуется тоннель. Волокна слухового нерва пересекают тоннель и контактируют с волосковыми клетками.

Слуховой нерв представляет собой перекрученный ствол, сердцевина которого состоит из волокон, отходящих от верхушки улитки, а наружные слои – от нижних ее участков. Войдя в ствол мозга, нейроны взаимодействуют с клетками различных уровней, поднимаясь к коре и перекрещиваясь по пути так, что слуховая информация от левого уха поступает в основном в правое

полушарие, где происходит главным образом обработка эмоциональной информации, а от правого уха в левое полушарие, где в основном обрабатывается смысловая информация. В коре основные зоны слуха находятся в височной области, между обоими полушариями имеется постоянное взаимодействие.

Высшие отделы слуховой системы (включая слуховые зоны коры), можно рассматривать как логический процессор, который выделяет (декодирует) полезные звуковые сигналы на фоне шумов, группирует их по определенным признакам, сравнивает с имеющимися в памяти образами, определяет их информационную ценность и принимает решение об ответных действиях [1, 8].

Строение уха представлено на рисунке 3.1 [1].

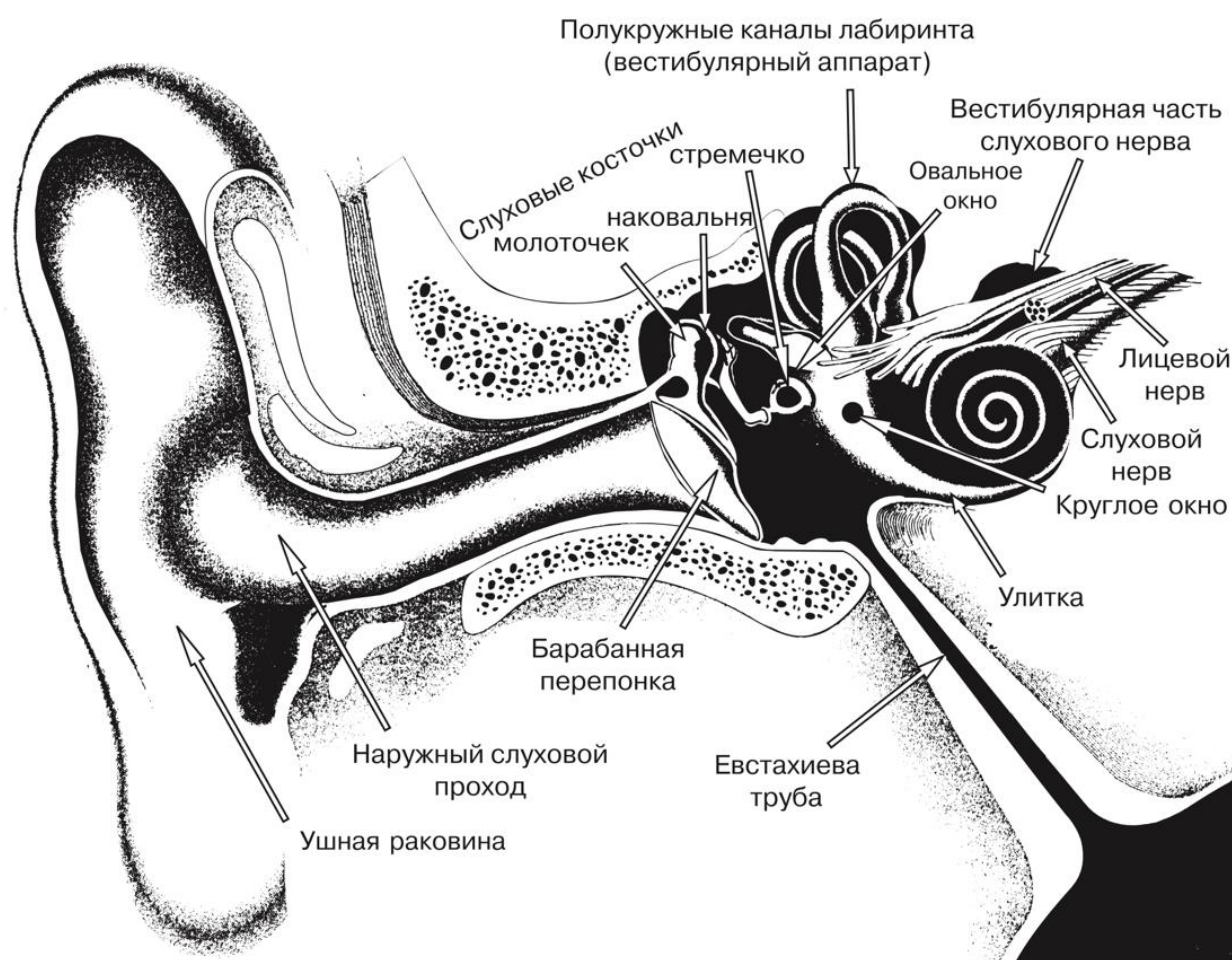


Рисунок 3.1 – Строение уха [1]

3.1.2 Возникновение слуховых ощущений

Существуют 2 способа передачи звуковых колебаний волосковым клеткам в улитке [8,9]:

- воздушная проводимость – передача звуков по воздуху в наружном и среднем ухе;

- костная проводимость – передача звуков по костям черепа, которые колеблются под действием звука и передают эти колебания жидкости в улитке), минуя наружное и среднее ухо.

Общий механизм передачи звука можно разбить на несколько этапов:

1. Звуковые колебания проникают в наружный слуховой канал, где вызывают колебания барабанной перепонки. Благодаря своей форме наружное ухо и наружный слуховой проход, действуя как естественные резонаторы, способствуют усилению звука до 17 дБ. Это усиление наблюдается в диапазоне 2-3 кГц, что важно для разнообразия согласных.

2. Затем вибрации барабанной перепонки передаются цепи слуховых косточек в среднем ухе.

3. Последняя кость этой цепи передает вибрации во внутреннее ухо через мембрану овального окна. Звукопередаточный механизм среднего уха также обеспечивает усиление звука. Это очень важно, так как звук из среднего уха, заполненного воздухом, передается во внутреннее ухо, заполненное жидкостью. В жидкой среде звук ослабляется на 30 дБ. Система среднего уха усиливает звук, компенсируя эти потери до 33 дБ в зависимости от частоты сигналов. Усиление звука обеспечивается рычажной системой косточек среднего уха, а также благодаря разности площадей барабанной перепонки и мембраны овального окна.

4. Колебания мембраны овального окна приводят в движение жидкость в улитке, что вызывает колебания базилярной мембраны, на которой расположены волосковые клетки. Это вызывает смещение и раздражение волосков волосковых клеток покровной мембраной, в которую они погружены.

5. В результате волосковые клетки создают слабые электрические потенциалы. Благодаря этим потенциалам выделяется химический медиатор.

6. Химический медиатор раздражает окончания волокон слухового нерва, что приводит к возникновению электрических импульсов.

7. Электрические импульсы передаются по слуховому нерву к подкорковым слуховым центрам мозга.

8. В подкорковых центрах производится анализ информации о частоте звука, локализации источника звука, формируются произвольные реакции

на звуки и др. Подкорковые центры передают информацию в корковые центры мозга.

9. В слуховых корковых центрах осуществляется основная обработка речи, узнавание, запоминание, интерпретация речевых и неречевых сигналов, формируются различные слуховые ощущения (музыки, речи и пр.) [8, 9].

Таким образом, можно утверждать, что человек слышит не ушами, а мозгом [1].

3.1.3 Влияние акустических условий на восприятие речи

Восприятие речи человеком в естественных условиях часто происходит на расстоянии, на фоне различных шумов (речь окружающих людей, шум различных устройств, ветер, шаги, уличный шум и многое другое), в помещениях с реверберацией.

Все эти факторы ухудшают восприятие речи при нормальном слухе человека. У детей и взрослых с нарушениями слуха, в том числе у пациентов со слуховыми аппаратами или кохлеарными имплантатами, центральными нарушениями слуха этот эффект более значителен. В таких условиях даже дети, которые хорошо понимают речь, могут воспринимать речь только на слух и визуально.

Особенно резко нарушается восприятие речи в таких условиях у детей. Взрослый, когда он не слышит часть слова или предложения, может угадать (восстановить) недостающие части на основе информации, хранящейся в его памяти. Он может даже не заметить, что это слово не было слышно отчетливо.

Ребенок не может этого сделать, если этого слова нет в его памяти.

Существуют множество факторов, оказывающие значительное влияние на идентификацию и понимание речи.

1. Расстояние. Уровень речи, воспринимаемой слушающим, падает на 6 дБ с удвоением расстояния от говорящего. Например, на расстоянии одного метра от слушающего уровень речи 60 дБ, а на расстоянии двух метров эта же речь будет иметь уровень 54 дБ. На уровень воспринимаемой речи влияет не только расстояние между говорящим и слушателем, но и их взаимное расположение. Когда один из говорящих находится за спиной, уровень речи снижается и соответствует увеличению расстояния в 2 раза. При таком уровне речи многие согласные звучат неразборчиво.

2. Дифракция – огибание звуками предметов (в том числе и голова человека), встречающихся на пути звуковой волны, которая распространяется от источника звука во все стороны. Низкие звуки лучше огибают препятствия, чем высокие. Поэтому при восприятии речи в таких условиях лучше

воспринимаются низкочастотные гласные по сравнению с более высокочастотными согласными, которые еще и более тихие.

3. Отражение и поглощение звука (ревербация). Реверберация – продолжение звучания звука после его окончания вследствие наложения отраженного звука. В помещении звук успевает отразиться от стен за сотые доли секунды, поскольку имеет высокую скорость распространения, и поэтому отражается от стен несколько раз. При этом отраженный звук накладывается на исходящий звук, маскируя и искажая его. При таких условиях восприятие звука нарушается и сильно влияет на восприятие речи. У слова из трех слогов последний слог слышится одновременно с отражением второго слога и более слабым отражением первого слога, что делает речь почти неразборчивой. Даже нормально слышащий ребенок хуже понимает речь других людей в помещении с сильной реверберацией.

4. Маскировка. С точки зрения восприятия в конкретной ситуации звуковые сигналы можно разделить на полезный сигнал и мешающие шумы. Полезный сигнал – это звук, который в данный момент представляет интерес для человека, при этом другие звуки являются мешающими и маскируют полезный сигнал. Если человек слушает, что ему говорит другой человек, то шум улицы, работающего телевизора, речи других людей маскируют речь говорящего и мешают ее восприятию. Высокочастотные звуки сильнее маскируются другими звуками, чем низкочастотные. На восприятие речи сильнее влияет речь других людей, чем окружающий неречевой шум, поскольку речь имеет тот же частотный состав.

5. Резонанс. Если в звуковом поле одного источника звука есть объект, способный звучать, то под воздействием звуковой волны он начинает колебаться, становясь вторичным излучателем звука – резонатором. Резонанс особенно выражен, когда резонатор и основной источник звука имеют одинаковую частоту собственных колебаний.

Особенности восприятия звука складываются из множества составляющих и не исчерпываются только физическими параметрами, как громкостью звука, четкостью речи и всеми факторами, что описаны выше. Даже такие психологические факторы, как стрессоустойчивость играют роль в восприятии речи [1].

3.2 Аудиометрия и аудиограмма

Аудиометрия – это измерение остроты слуха. Этот метод субъективен и требует обратной связи с пациентом. Исследователь посылает сигнал с помощью аудиометра, и испытуемый, чей слух исследуется, дает знать,

слышит ли он этот звук или нет. Чаще всего он нажимает для этого кнопку, реже – поднимает руку или кивает, дети складывают игрушки в корзину.

Существуют различные типы аудиометрии: тональная пороговая, надпороговая и речевая. На практике чаще всего используется тональная пороговая аудиометрия, которая определяет минимальный порог слуха (самый тихий звук, который слышит человек в децибелах (дБ)) на различных частотах (обычно в диапазоне 125 Гц — 8000 Гц, реже до 12500 и даже до 20 000 Гц). Эти данные помечаются на специальной форме.

Аудиограмма – это график слуховых ощущений пациента. Эти ощущения могут зависеть от человека, его общего состояния, артериального и внутричерепного давления, настроения и т. д., а также от внешних факторов, атмосферного воздействия шума в помещении, отвлекающих факторов и т. д. Пример аудиограммы для потери слуха второй степени приведён на рисунке 3.2.

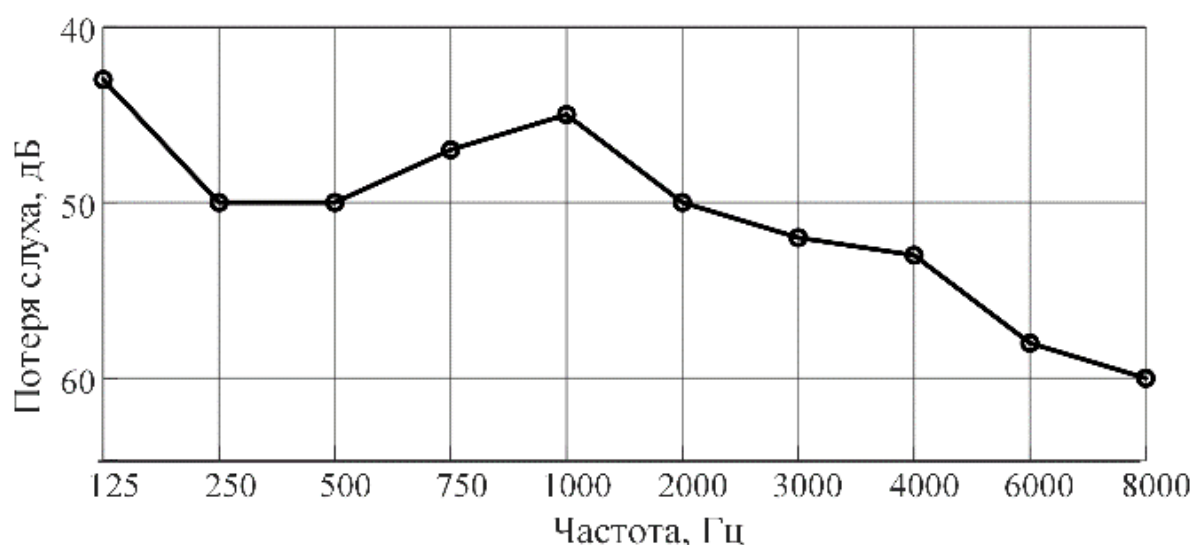


Рисунок 3.2 – Пример аудиограммы

Аудиограмма может быть восходящей (чаще при кондуктивной тугоухости), нисходящей (чаще при сенсоневральной тугоухости), горизонтальной, а также иметь ряд других конфигураций [8].

Настройка разработанной системы выполняется в соответствии с параметрами слуха конкретного тугоухого человека, поэтому понятие «аудиограмма» играет важную роль в этом проекте.

3.3 Принцип работы системы повышения разборчивости речи

Принцип работы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала представляет собой следующий алгоритм: входной сигнал $x[n]$ разделяется на субполосы (компоненты) при помощи 20-канального банка гамматон-фильтров. Каждая компонента проходит через отдельный компрессор динамического диапазона (КДД). Основная задача компрессоров – расчёт коэффициентов субполосного усиления сигнала. Система настраивается по аудиограмме слабослышащего человека. Выходной звуковой сигнал $y[n]$ синтезируется при помощи суммирования обработанных субполос.

Структура системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала представлена на рисунке 3.3.

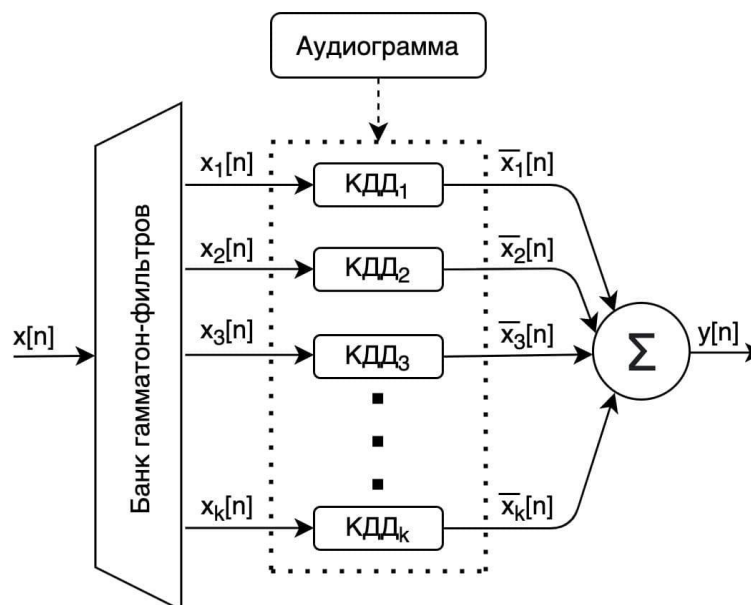


Рисунок 3.3 – Структура системы повышения разборчивости речи

Более детальная структура системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала представлена в виде графического документа ГУИР 460818.001 ПЛ к данной работе.

3.3.1 Банк гамматон-фильтров

В качестве модуля, разделяющего сигнал на субполосы, используется банк гамматон-фильтров, имитирующий механизм частотной декомпозиции

звука в улитке. Фильтры получили свое название от того, что их огибающая формируется как произведение степенной и экспоненциальной функций, которое является функцией гамма распределения, а косинусная несущая является тоном, когда её частота находится в слышимом диапазоне.

Импульсная характеристика гамматон-фильтра описывается следующим выражением:

$$g(t) = at^{l-1} e^{-2\pi b \text{ERB}(f_c)t} \cos(2\pi f_c t + \phi), \text{ при } t \geq 0, \quad (3.1)$$

где a – амплитуда, l – порядок фильтра, ERB – эквивалентная прямоугольная полоса пропускания, b – параметр, определяющий ширину полосы фильтра, f_c – центральная частота, ϕ – фаза [10].

При умеренных уровнях мощности сигнала $\text{ERB}(f_c)$ описывается следующей формулой [11]:

$$\text{ERB}(f_c) = 24.7 + 0.108f_c, \quad (3.2)$$

В данной работе порядок фильтров l выбран равным 4 ($l = 4$), фаза ϕ равна нулю ($\phi = 0$), параметр b равен 1.019 ($b = 1.019$) [12].

На практике гамматон-фильтры моделируются в виде цифровых КИХ-фильтров. Однако импульсные характеристики фильтров имеют большую длительность. На рисунке 3.4 приведены импульсная характеристика гамматон-фильтра для частоты 500 Гц, а также огибающие импульсных характеристик для центральных частот 500 и 80 Гц [13].

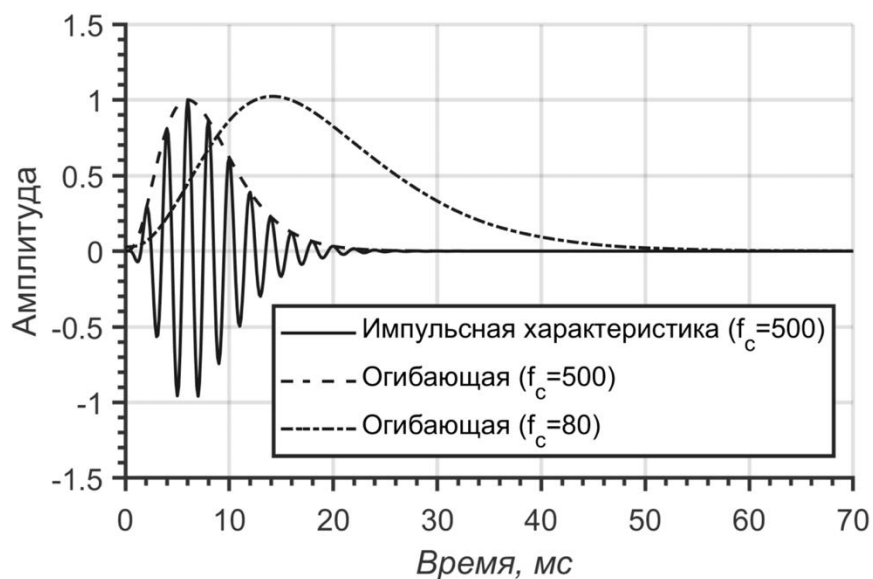


Рисунок 3.4 – Импульсная характеристика и огибающие импульсных характеристик гамматон-фильтров для различных центральных частот [13]

Можно заметить, что для центральной частоты 500 Гц длина импульсной характеристики фильтра будет составлять порядка 240 отсчетов (при частоте дискретизации 8000 Гц). Для более низких центральных частот, длина импульсной характеристики возрастает ещё больше (например, для центральной частоты равной 80 Гц, она будет содержать 600 отсчетов). Таким образом, получающиеся КИХ-фильтры в вычислительном плане являются довольно затратными [13].

Центральные частоты банка фильтров (f_c) выбираются в соответствии с психоакустической шкалой Барков [14]. Шкала барков колеблется от 1 до 24 и соответствует первым 24 критическим полосам слуха. Поскольку ширина этих полос (в герцах) неравномерна и увеличивается с возрастанием частоты звуковых колебаний, то шкала также является нелинейной. График, отражающий зависимость критических полос от частоты, представлен на рисунке 3.4 [15].

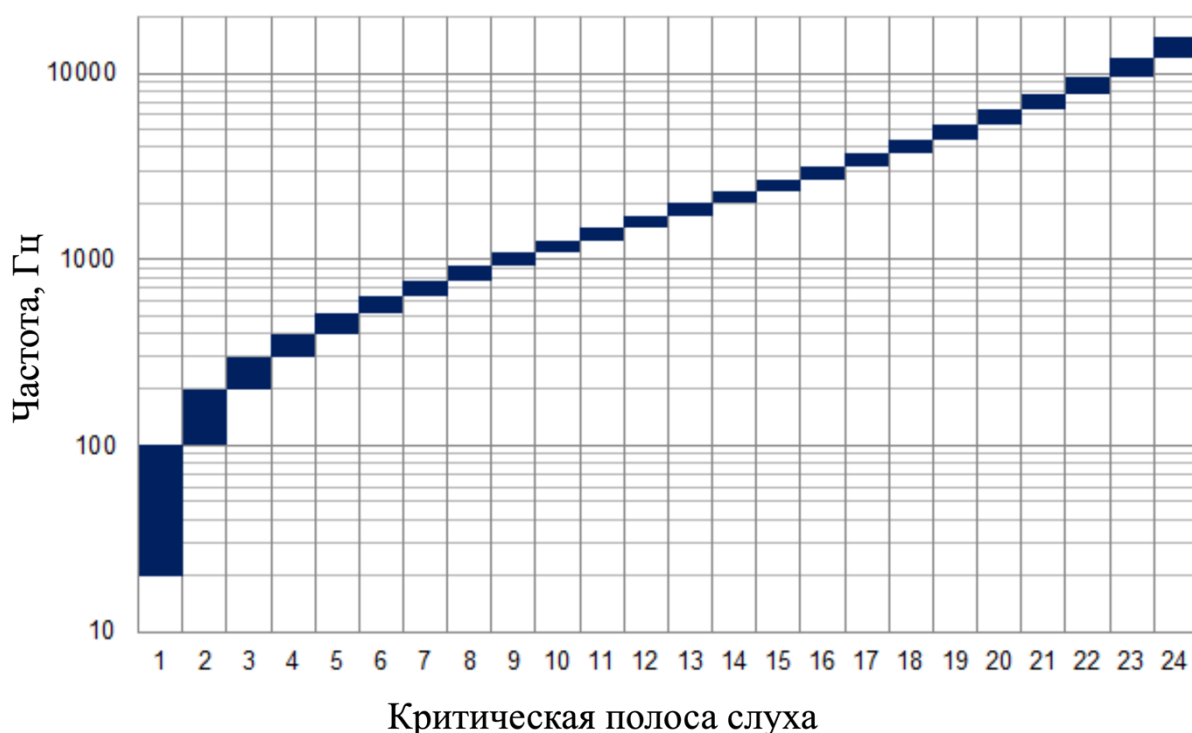


Рисунок 3.4 – Зависимость критических полос от частоты [15]

В данной работе используются значения центральных частот для 20 критических субполос, начиная со второй. Соответствие значений f_c и номеров критических полос представлены в таблице 3.1.

Таблица 3.1– Значения центральных частот банка фильтров

Номер критической полосы	Центральная частота, Гц
2.	153
3.	250
4.	347
5.	453
6.	569
7.	697
8.	838
9.	995
10.	1172
11.	1370
12.	1595
13.	1853
14.	2152
15.	2501
16.	2915
17.	3413
18.	4025
19.	4794
20.	5790
21.	7019

АЧХ 20-канального банка гамматон-фильтров приведена на рисунке 3.5.

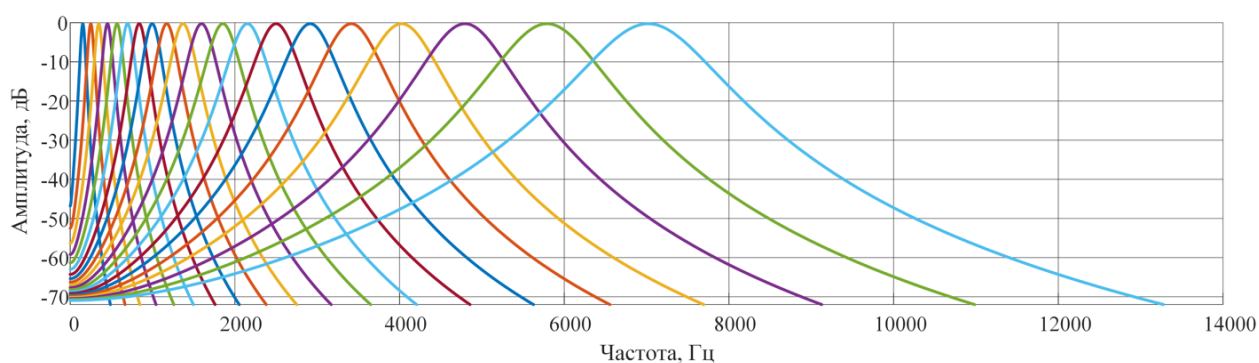


Рисунок 3.5 – АЧХ банка гамматон-фильтров

Рисунок 3.5 демонстрирует нелинейность восприятия звука человеческим ухом.

Схема программы реализации банка фильтров для системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала представлена в виде графического документа ГУИР 460818.002 ПД к данной работе.

3.3.2 Настройка компрессора динамического диапазона (КДД)

Потеря слуха характеризуется сужением слышимого динамического диапазона, снижением порогов слышимости, ухудшением частотного разрешения уха. Следствие этого – разборчивость речи у тугоухого человека падает. Для её повышения необходимо решить задачу отображения широкого динамического диапазона нормально слышащего человека в остаточный диапазон тугоухого человека, т.е. выполнить компрессию динамического диапазона. Помимо того, необходимо выполнить частотно-зависимое усиление сигнала. Основная цель КДД в разработанной системе – автоматический выбор коэффициентов усиления субполос в зависимости от уровня мощности сигнала. Для выполнения задачи частотно-зависимого усиления сигнала для каждой субполосе строится отдельный компрессор, настраиваемый согласно характеристикам слуха (аудиограмме) тугоухого человека. Использование такого механизма настройки системы позволяет ослаблять субполосы с высокой мощностью, а с малой – усиливать. Благодаря этому, тихие звуки становятся слышимыми, а громкие звуки не вызывают дискомфортных ощущений [2].

Пример характеристики вход/выход КДД для частоты 153 Гц, построенный по аудиограмме из рисунка 3.2, приведён на рисунке 3.6.

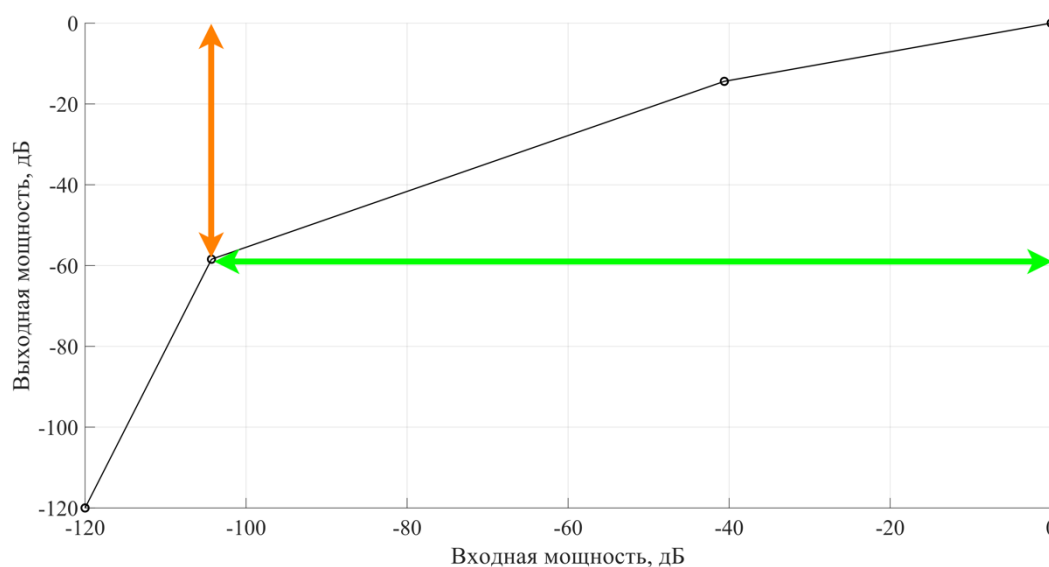


Рисунок 3.6 – Пример характеристики вход/выход КДД

Порог компрессии – параметр КДД, определяющий точку изгиба характеристики вход/выход компрессора, после которой алгоритм компрессии становится активным. Если уровень входного сигнала ниже порога компрессии, то выходной сигнал усиливается линейно, а если выше – коэффициент усиления уменьшается.

Настройка системы под характеристики слуха конкретного тугоухого человека осуществляется на центральных частотах банка фильтров (f_c) с использованием аудиограммы [14].

Помимо характеристики вход/выход КДД, рисунок 3.6 демонстрирует изменение слышимого диапазона при наличии потери слуха. Так, например, слышимый диапазон для слуха в норме на рисунке 3.6 составляет 0...-105 дБ (зелёная стрелка), а для тугоухого человека: 0...-60 дБ (оранжевая стрелка), что соответствует эффекту уменьшения слышимого динамического диапазона при ухудшении слуха.

Для настройки системы необходимо определить две точки перегиба характеристики вход/выход КДД для каждой субполосы: порог слышимости и порог комфорта. Порог слышимости – это наименьший уровень звукового давления (УЗД), который может быть воспринят человеческим ухом. Порог комфорта – это уровень звукового давления, который позволяет человеческому уху воспринимать более естественный и более комфортный звук [16].

Для определения первой точки необходимо знать пороги слышимости в норме и при тугоухости на частотах f_c . Из аудиограммы извлекается порог слышимости тугоухого человека (HT_{HI}). Порог слышимости нормально слышащего человека (HT_{NORM}) определяется исходя из стандарта ISO226, описывающего кривые равной громкости [17]. Таким образом, зная пороги слышимости в норме и при наличии патологии можно найти первую точку перегиба характеристики вход/выход КДД. Для этого, на первом этапе выполняется интерполяция порогов слышимости (в норме и с патологией) на сетку центральных частот банка фильтров. Далее, необходимо выполнить перевод значений порогов слышимости из УЗД (уровень звукового давления) в «цифровые» дБ. Следующим этапом определяется точка перегиба (порог слышимости) характеристики вход/выход КДД, где в качестве входного используется значение порога слышимости в норме (HT_{NORM}), а выходного – значение порога слышимости тугоухого человека (HT_{HI}) [2, 18].

В связи с эпидемиологической ситуацией в мире и невозможностью проведения нужных экспериментов в Республиканском научно-практическом центре оториноларингологии при участии людей, имеющих патологии слуха, было принято решение пороги комфорта на центральных частотах банка

фильтров при нормальном слухе, при второй степени потери слуха и при третьей степени потери слуха аппроксимировать значениями в дБ для каждой субполосы. Аппроксимированные значения порогов комфорта представлены в таблице 3.2

Таблица 3.2– Значения для порогов комфорта

№	Порог комфорта при нормальном слухе	Порог комфорта при второй степени потери слуха	Порог комфорта при третьей степени потери слуха
1.	-40.2	-17.6	-7.6
2.	-36.1	-16.3	-5.9
3.	-39.1	-22.1	-10.5
4.	-37.5	-19.5	-5.2
5.	-36.5	-16.9	-7.8
6.	-37.2	-21.2	-11.9
7.	-42.4	-17.4	-9.3
8.	-39	-22.8	-5.5
9.	-40.6	-14.4	-11.7
10.	-41.8	-18.1	-11.7
11.	-41.4	-16.5	-10.9
12.	-37	-18.9	-9.7
13.	-34.4	-15.1	-4.9
14.	-42.6	-18.5	-9.6
15.	-41.4	-22.5	-11.4
16.	-41.4	-15.6	-8
17.	-41.6	-15.9	-7.9
18.	-40.5	-24.7	-11.8
19.	-41.5	-22.1	-11
20.	-39.3	-17.8	-9.7

Блок-схема алгоритма настройки компрессора динамического диапазона сигнала представлена на рисунке 3.7.

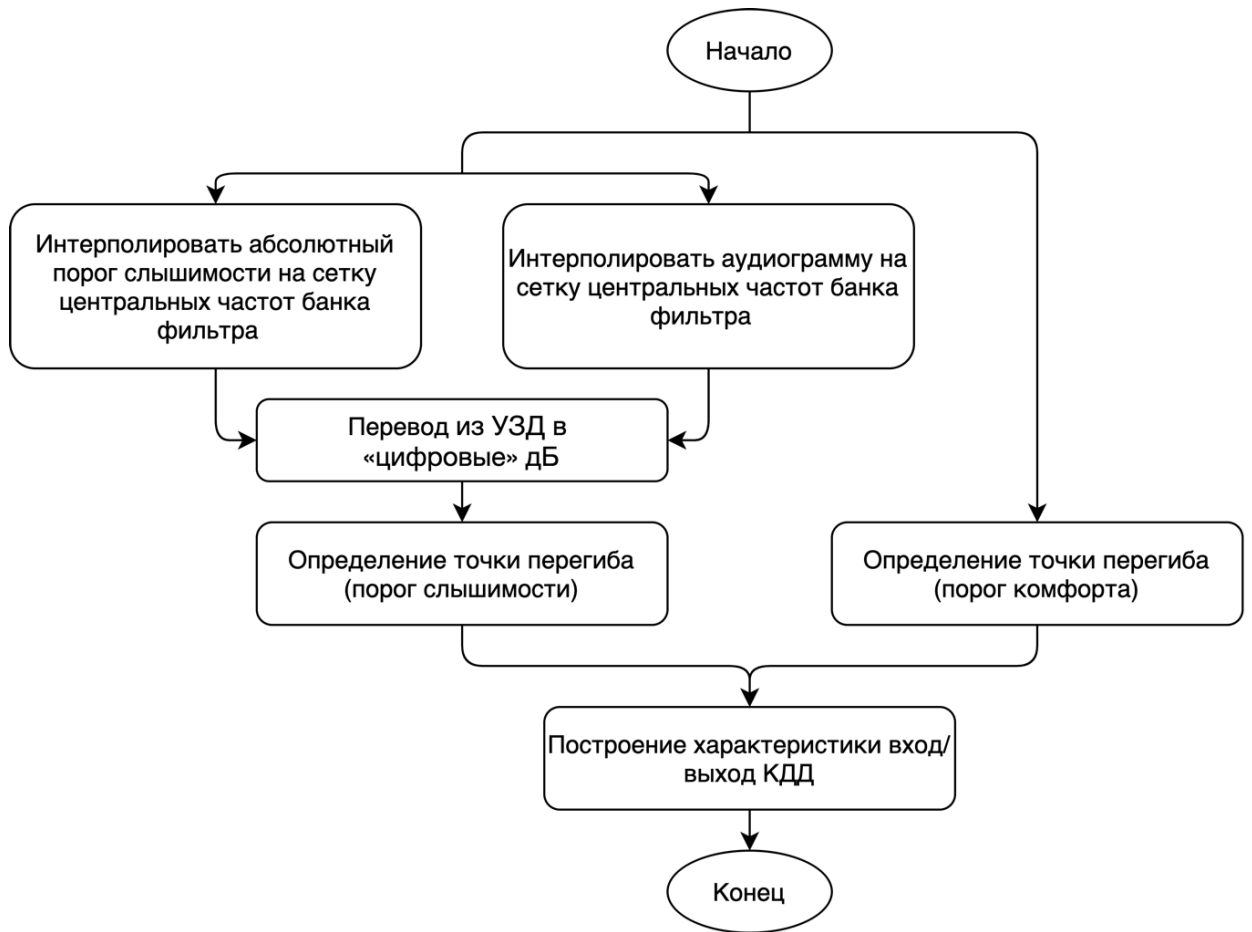


Рисунок 3.7 – Блок-схема алгоритма настройки КДД

Настройка системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала по аудиограмме представлена на плакате ГУИР 460818.004 ПЛ к данной работе. На плакате отображены абсолютный порог слышимости и порог слышимости в норме, характеристика КДД для центральной частоты равной 1172 Гц, аудиограмма человека со второй степенью потери слуха, амплитудно-частотная характеристика для банка гамматон-фильтров и их взаимосвязь для определения точки порога слышимости для настройки системы.

3.3.3 Реализация КДД

Для реализации КДД в каждой субполосе на выходе банка фильтров, на частотах f_c производится оценка мощности сигнала:

$$P_k^{in}[n] = \begin{cases} P_k[n-1]\alpha + x_k^2[n](1-\alpha), & \text{если } x_k^2[n] > P_k[n-1] \\ P_k[n-1]\beta + x_k^2[n](1-\beta), & \text{если } x_k^2[n] \leq P_k[n-1] \end{cases}, \quad (3.1)$$

$$k = 1, 2 \dots K,$$

где $P_k[n]$ – мощность сигнала на выходе k -го канала банка фильтров, K – число каналов банка фильтров, $x_k[n]$ – сигнал на выходе k -го канала банка фильтров, α – коэффициент атаки, β – коэффициент восстановления, которые вычисляются по следующим формулам:

$$\alpha = 1 - \exp\left(\frac{-2.2T_S}{t_a/1000}\right) \quad (3.2)$$

$$\beta = 1 - \exp\left(\frac{-2.2T_S}{t_r/1000}\right) \quad (3.3)$$

где T_S – период дискретизации, t_a – время атаки в секундах, t_r – время восстановления в секундах [2].

В данной работе параметры α и β приняты равными 0.99 и 0.9992 ($\alpha = 0.99$, $\beta = 0.9992$).

Далее, рассчитываются субполосные коэффициенты усиления g_k :

1) Полученный уровень мощности P_k^{in} переводится в дБ:

$$P_{k,dB}^{in} = 10 \log_{10} P_k^{in} \quad (3.4)$$

2) С использованием линейной интерполяции рассчитывается выходная мощность $P_{k,dB}^{out}$ исходя из характеристики вход/выход КДД;

3) вычисляются коэффициенты усиления в каждой субполосе (в дБ):

$$G_{k,dB} = P_{k,dB}^{out} - P_{k,dB}^{in} \quad (3.5)$$

4) коэффициенты $G_{k,dB}$ переводятся из дБ в линейный масштаб:

$$g_k = 10^{G_{k,dB} / 20} \quad (3.6)$$

После нахождения коэффициентов g_k они умножаются на соответствующие субполосы $x_k[n]$ [19]:

$$y_k[n] = x_k[n]g_k[n] \quad (3.7)$$

Выходной сигнал синтезируется путём суммирования субполосных компонент сигнала.

3.4 Принцип работы системы моделирования потери слуха

Система моделирования потери слуха может быть использована в качестве альтернативного подхода к экспериментальным исследованиям эффективности метода коррекции слуха: вместо проверки на тугоухих она выполняется на нормально слышащих людях, а сама система имитирует наличие слуховой патологии.

Принцип работы системы моделирования потери слуха представляет собой схожий алгоритм с тем, что описан в пункте 3.3 для системы повышения разборчивости речи. Отличие заключается в том, для каких целей используется КДД и, соответственно, в настройке КДД. Основная задача компрессоров для этой системы – расчёт коэффициентов субполосного ослабления сигнала. Система также настраивается по аудиограмме слабослышащего человека. Выходной звуковой сигнал $y[n]$ синтезируется при помощи суммирования обработанных субполос.

Для настройки системы необходимо определить одну точку перегиба (порог слышимости) характеристики вход/выход КДД для каждой субполосы.

Блок-схема алгоритма настройки компрессора динамического диапазона для системы моделирования потери слуха представлена на рисунке 3.8.

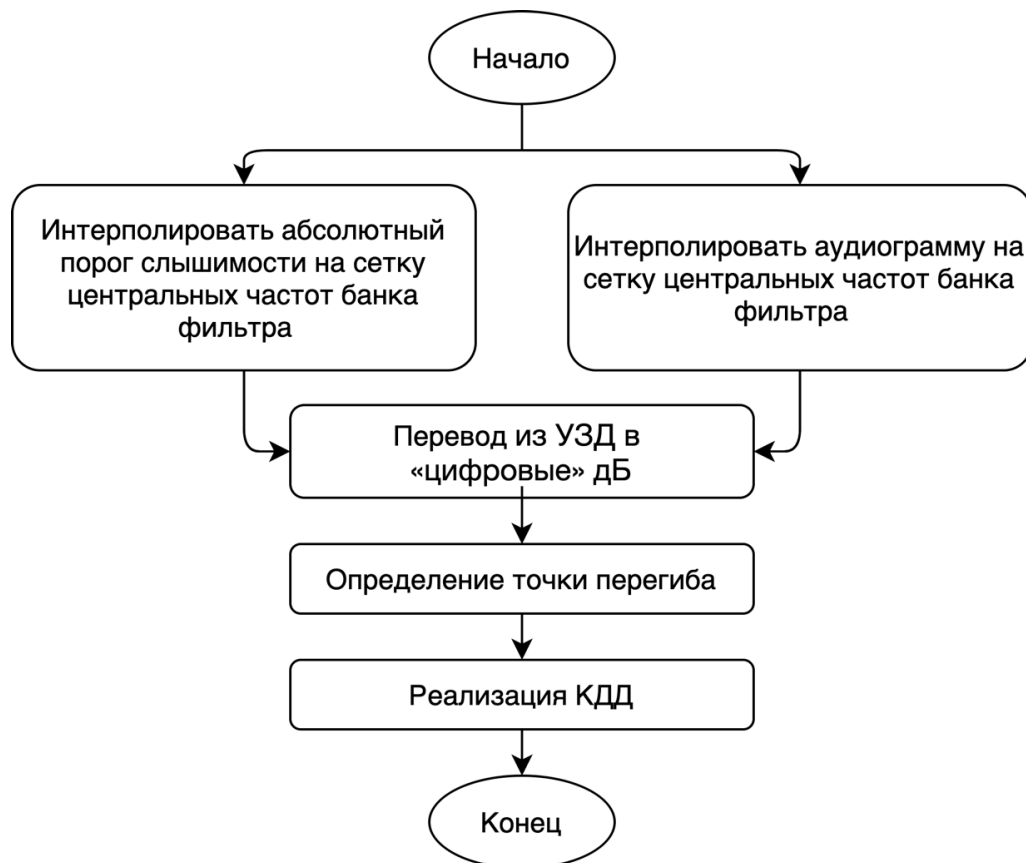


Рисунок 3.8 – Блок-схема алгоритма настройки КДД

Пример характеристики вход/выход КДД, построенный по аудиограмме из рисунка 3.2, приведён на рисунке 3.9.

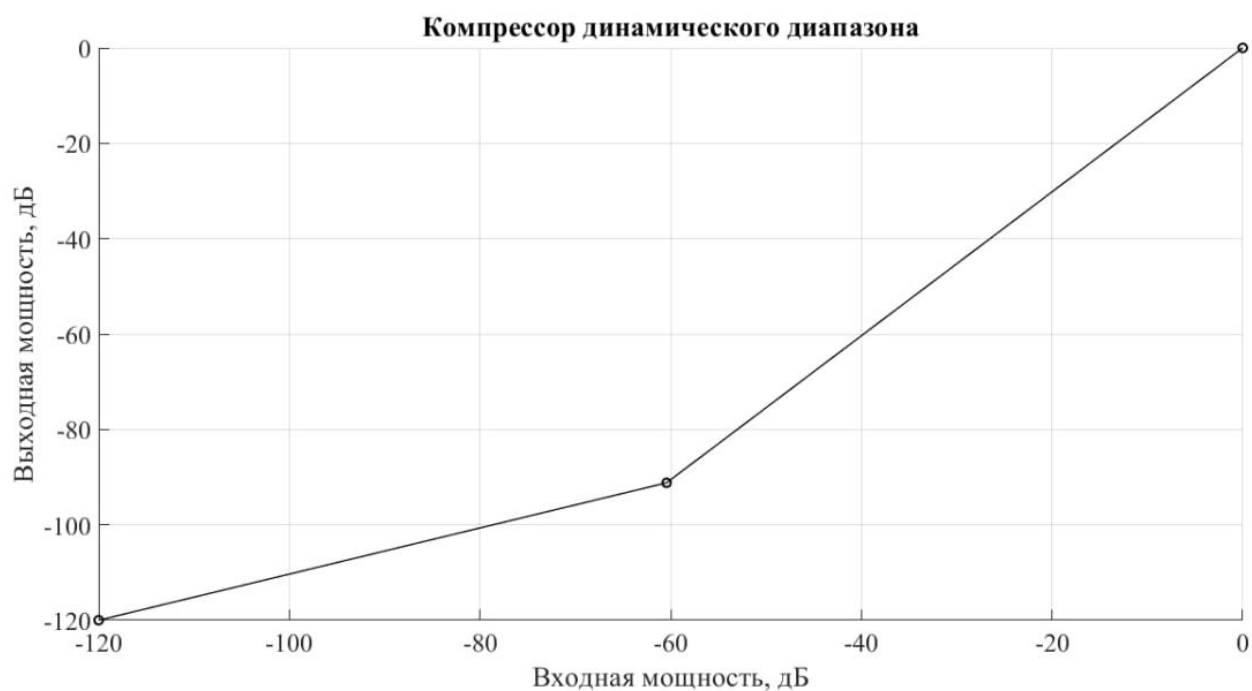


Рисунок 3.9 – Пример характеристики вход/выход КДД для настройки модели потери слуха

4 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ

Программная реализация системы повышения разборчивости речи на основе слухового восприятия и компрессии динамического диапазона сигнала осуществлялась с помощью вычислительной платформы MATLAB или Matrix Laboratory.

4.1 Описание пакета MATLAB

MATLAB – это программная и числовая вычислительная платформа, используемая миллионами инженеров и ученых для анализа данных, разработки алгоритмов и создания моделей. MATLAB сочетает в себе настольную среду, настроенную для итеративного анализа и проектирования процессов, с языком программирования, который непосредственно выражает математику матриц и массивов. Он включает в себя живой редактор для создания сценариев, которые объединяют код, выходные данные и форматированный текст в исполняемом блокноте.

MATLAB можно использовать для следующих целей:

- проектирование, тестирование и внедрение систем управления;
- обучение моделей;
- анализ сигналов и данных временных рядов. Проектирование и моделирование системы обработки сигналов;
- подготовка данных, проектирование, моделирование и развертывание глубоких нейронных сетей;
- разработка и развертывание программного обеспечения для мониторинга состояния и прогнозирования технического обслуживания;
- сбор, анализ и изучение данных, автоматизация тестов;
- получение, обработка и анализ изображений и видео для разработки алгоритмов и проектирования систем;
- преобразование идеи и концепции робототехники в автономные системы, которые легко работают в реальных условиях;
- создание, проектирование, тестирование, проверка систем беспроводной связи.

MATLAB – это высокоуровневый интерпретируемый язык программирования, который включает в себя матричные структуры данных, широкий спектр функций, интегрированную среду разработки, объектно-ориентированные функции и интерфейсы к программам, написанным на других языках программирования.

Программы, написанные на языке MATLAB, бывают двух типов – функции и скрипты. Функции имеют входные и выходные аргументы, а также собственное рабочее пространство для хранения промежуточных результатов вычислений и переменных. Сценарии имеют общую рабочую область. Как скрипты, так и функции хранятся в виде текстовых файлов и динамически компилируются в машинный код. Существует также возможность сохранения так называемых предварительно проанализированных программ-функций и скриптов, обработанных в удобной для машинного исполнения форме. Как правило, такие программы работают быстрее, чем обычные программы, особенно если функция содержит команды построения графиков. Главной особенностью языка MATLAB являются его обширные возможности для работы с матрицами [20].

4.2 Программная реализация системы повышения разборчивости речи

Основываясь на алгоритме, приведенном в подразделе 3.3, алгоритм программной реализации содержит следующие этапы:

1. Определение центральных частот для банка фильтров с помощью функции *bark2frq*;
2. Определение порога слышимости в УЗД для нормального слуха с помощью стандарта ISO226;
3. Интерполяция порога слышимости в норму и аудиограммы на сетку центральных частот;
4. Расчёт порога слышимости при наличии патологии;
5. Считывание входного сигнала;
6. Разложение сигнала на субполосы с помощью функции *gammatoneFast*;
7. Коррекция амплитуд;
8. Производится усиление для каждой субполосы;
 - 8.1. Извлечение параметров для настройки КДД;
 - 8.2. Для оптимизации вычислительной мощности каждый сигнал разбивается на фреймы ($N_{frames} = 256$);
 - 8.3. Для каждого фрейма рассчитывается уровень мощности;
 - 8.4. Полученный уровень мощности переводится в дБ;
 - 8.5. Расчет коэффициентов усиления/ослабления для каждого фрейма;
 - 8.6. Коэффициенты переводятся из дБ в линейный масштаб;
9. Синтез выходного сигнала;

10. Запись в файл выходного сигнала для дальнейшего использования в тестировании системы.

Исходный код программной реализации алгоритма, приведенного в пункте 3.3, отражен в приложении А.1.

Схема программы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала приведена в графической документации к данной работе ГУИР 460818.003 ПД.

4.3 Описание вспомогательных функций

4.3.1 Функция *gammatoneFast*

Алгоритм работы функции заключается в следующем: *gammatoneFast* декомпозирует входной сигнал x используя банк гамматон-фильтров четвертого порядка с центральными частотами, которые передаются в параметре *cfs*. Фильтрация исходного сигнала выполняется в частотной области при помощи функции *fftfilt*. Функция *gammatoneFast* возвращает матрицу, каждая строка/столбец которой соответствует выходу фильтра с центральной частотой, определяемой соответствующим элементом в *cfs*. Ориентация выходных данных определяется ориентацией входных данных: если x – вектор строк, то выходные данные будут содержать по одной строке для каждого выхода фильтра, и наоборот.

Центральные частоты могут иметь любое значение ниже частоты Найквиста (определяется частотой дискретизации fs). Параметр *align* позволяет применять выравнивание фаз. По умолчанию данный параметр отключен и не применяется в данной работе.

Описание параметров данной функции представлено в таблице 4.1.

Таблица 4.1– Описание параметров функции *gammatoneFast*

Параметры	Возможные значения	Описание
x	Вектор-строка, вектор-столбец	Входной сигнал
<i>cfs</i>	Вектор-строка	Центральные частоты
fs	Число (не обязательный)	Частота дискретизация
<i>align</i>	true/false (не обязательный)	Параметр выравнивания

Описание возвращаемых значений функцией *gammatoneFast* представлено в таблице 4.2.

Таблица 4.2– Описание возвращаемых значений функции *gammatoneFast*

Возвращаемые значения	Возможные значения	Описание
<i>bm</i>	Матрица, каждая строка/столбец которой соответствует выходу фильтра с центральной частотой, определяемой соответствующим элементом в <i>cfs</i>	Сигнал, разбитый на субполосы
<i>env</i>	Вектор-строка	Огибающая
<i>delay</i>	Вектор того же размера, что вектор центральных частот	Задержка

Данная функция возвращает сигнал, разбитый на субполосы, мгновенные огибающие *env* для каждого фильтра и задержку *delay* (вектор того же размера, что и *cfs*).

4.3.2 Функция *bark2frq*

Для определения центральных частот фильтров, которые выбираются в соответствии с психоакустической шкалой Барков, удобно воспользоваться функцией перевода из шкалы Гц в шкалу Барков – *bark2frq*.

Входным параметром является массив, который содержит массив значений в Барках. Шкала Барков находится в диапазоне от 1 до 24 Барк, что соответствует первым 24 критическим полосам слуха [21]. Выходное значение представляет собой массив значений в Гц той же размерности, что входной параметр.

4.3.3 Функция *ttest*

После проведения экспериментального исследования эффективности работы разработанной системы необходимо провести проверку значимости полученных результатов для исключения возможности получения случайных результатов из-за небольшой выборки. В данной работе проверка выполнена при помощи MATLAB-функции *ttest*, использующая парный t-критерий.

Алгоритм такой проверки заключается в том, что для каждого участника эксперимента вычисляется разница между разборчивостью речи с применением и без применения метода коррекции слуха. При этом нулевая гипотеза подразумевает то, что истинная разность равна нулю. Для проверки этой гипотезы в среде MATLAB использовался t-критерий Стьюдента для вычисления р-значения с установленным пороговым уровнем значимости α который в данной работе был выбран $\alpha = 0.05$. Нулевая гипотеза отвергается, если р-значение меньше уровня значимости α [22].

Описание параметров данной функции представлено в таблице 4.3., а описание возвращаемых значений представлено в таблице 4.4.

Таблица 4.3– Описание параметров функции *ttest*

Параметры	Возможные значения	Описание
x	Массив чисел	Первая выборка значений
y	Массив чисел	Вторая выборка значений
<i>name</i>	Строка	Название параметра, который задается следующим параметром <i>value</i>
<i>value</i>	Число	Пороговый уровень значимости

Таблица 4.4– Описание возвращаемых значений функции *ttest*

Возвращаемые значения	Возможные значения	Описание
h	Число	Результат для нулевой гипотезы
p	Число	Уровень значимости
ci	Массив из двух чисел	Доверительный интервал
<i>stats</i>	Структура со следующими полями: <i>tstat</i> (число), <i>df</i> : (число), <i>sd</i> (число)	Информацию о статистике теста

Исходный код программной реализации проверки приведен в приложении А.2.

4.3.4 Функция *iso226_rev*

Эта функция возвращает 29-точечный кривых равной громкости для желаемого уровня звука. Частоты, оцениваемые в этой функции, лежат в диапазоне от 20 Гц до 12,5 кГц, и покрывают только 29 избирательных частот. Это ограничение стандарта ISO.

Кроме того, допустимый диапазон входного параметра *phon* должен составлять 0-90 дБ в линейном масштабе (1 *phon* = 1 дБ при 1 кГц). Значения за пределами этого диапазона не имеют экспериментальных значений, и их кривые следует рассматривать как неточные.

Описание возвращаемых значений данной функции представлено в таблице 4.5.

Таблица 4.5– Описание возвращаемых значений функции *iso226_rev*

Возвращаемые значения	Возможные значения	Описание
<i>spl</i>	Число	Амплитуда уровня звукового давления, возвращаемая для каждой из 29 частот, оцененных по стандарту ISO226
<i>freq</i>	Вектор чисел	Вектор частот для создания кривой равной громкости

5 ТЕХНИКО-ЭКОНОМИЧЕСКОЕ ОБОСНОВАНИЕ ЦЕЛЕСООБРАЗНОСТИ ПРОЕКТИРОВАНИЯ И ВНЕДРЕНИЯ СИСТЕМЫ ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ

5.1 Характеристика системы повышения разборчивости речи

Проводимая научно-исследовательская работа заключается в проектировании системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала.

По состоянию на 2021 год более 430 миллионов человек нуждаются в реабилитации для решения проблемы потери слуха. По прогнозам, к 2050 г. по меньшей мере 700 миллионов человек будут нуждаться в реабилитационных услугах в связи с потерей слуха и около 2,5 миллиарда человек в той или иной степени будут страдать от проблем со слухом. В связи с этим существует задача оперативной коррекции слуховых патологий.

Главная область применения разработанной системы – слуховые аппараты. Основная цель системы повышения разборчивости речи заключается в непосредственном усилении входного сигнала и повышения разборчивости в соответствии с характеристикой слуха тугоухого человека. Важной особенностью системы является применение в ней банка гамматон-фильтров, имитирующего механизм частотной декомпозиции звука в ухе. Другой особенностью системы является её автоматическая настройка по аудиограмме слабослышащего человека. Эти особенности помогают разработать систему, которая справляется с задачей повышения разборчивости лучше, чем ряд существующих методов.

5.2 Составление плана и определение трудоемкости выполнения научно-исследовательской работы

Данная работа проводится в рамках внутреннего исследования университета с дальнейшим развитием в научных статьях и журналах. План на проведение научно-исследовательской работы представлен в таблице 5.1. В данном плане отражены этапы проектирования системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала, а также количество и состав исполнителей: научный руководитель и ответственный исполнитель.

Основными методами определения трудоёмкости являются система аналогов, метод прямого счёта и метод экспертных оценок. В данной работе при расчете трудоемкости использовался метод прямого счета. При прямом

счёте трудоёмкость обосновывается руководителем совместно с дипломником.

Таблица 5.1– План на проведение научно-исследовательской работы

№	Наименование этапов и видов работ	Исполнитель (должность, квалификация)	Количество исполнителей	Трудоемкость, чел.-дн.
1	Сопоставление и утверждение ТЗ на НИР	научный руководитель, ответственный исполнитель	2	0.5
2	Сбор и изучение научно-технической литературы и других материалов, относящихся к теме исследования	ответственный исполнитель	1	5
3	Проведение патентных исследований	ответственный исполнитель	1	5
4	Формулирование возможных направлений решения задач, поставленных в ТЗ на НИР и их сравнительная оценка	ответственный исполнитель	1	4
5	Составление аналитического обзора состояния вопросов по теме	ответственный исполнитель	1	5
6	Выбор и обоснование принятого направления проведения исследований способов решения поставленных задач	научный руководитель, ответственный исполнитель	2	2.5

Продолжение таблицы 5.1

№	Наименование этапов и видов работ	Исполнитель (должность, квалификация)	Количество исполнителей	Трудоемкость, чел.-дн.
7	Разработка общей методики проведения исследований	научный руководитель, ответственный исполнитель	2	1
8	Проектирование интерфейса системы повышения разборчивости речи	ответственный исполнитель	1	4
9	Проектирование алгоритма работы системы повышения разборчивости речи	ответственный исполнитель	1	6
10	Программная реализация интерфейса и алгоритма работы системы повышения разборчивости речи	ответственный исполнитель	1	10
11	Составление промежуточного отчёта и его рассмотрение	научный руководитель, ответственный исполнитель	2	3
12	Разработка частных методик проведения экспериментальных исследований.	ответственный исполнитель	1	5
13	Подготовка моделей, а также испытательного оборудования, необходимых для проведения экспериментальных исследований	научный руководитель, ответственный исполнитель	2	1.5

Продолжение таблицы 5.1

№	Наименование этапов и видов работ	Исполнитель (должность, квалификация)	Количество исполнителей	Трудоемкость, чел.-дн.
14	Проведение исследований	ответственный исполнитель	1	1
15	Сопоставление результатов экспериментов с теоретическими исследованиями	научный руководитель, ответственный исполнитель	2	1
16	Обобщение результатов предыдущих этапов работы	научный руководитель, ответственный исполнитель	2	2.5
17	Оценка полноты решения поставленных задач	ответственный исполнитель	1	5
18	Составление и оформление отчета	ответственный исполнитель	1	18
19	Рассмотрение результатов проведенной НИР	ответственный исполнитель	1	2
Итого				99

5.3 Расчёт цены научно-технической продукции

5.3.1 Расчёт затрат по статье «Топливо-энергетические ресурсы для научно-экспериментальных целей»

В статью «Топливо-энергетические ресурсы для научно-экспериментальных целей» включаются затраты на топливо-энергетические ресурсы, потребляемые оборудованием, используемые для научно-экспериментальных и технологических целей.

Таблица 5.2– Расчёт затрат на топливно-энергетические ресурсы

№	Наименование оборудования	Установочная мощность, кВт	Время использования, ч.	Тариф за кВт/ч	Сумма, руб.
1	Ноутбук Dell Vostro 14 (5490-279601)	0.065	792	0.32181	16.57
Итого					16.57

5.3.2 Расчёт затрат по статье «Спецоборудование для научных (экспериментальных) работ»

В данную статью включаются затраты на приобретение специальных инструментов, приспособлений, приборов, стендов, устройств и другого специального оборудования.

Информация о ценах на программное обеспечение взято с официального сайта поставщика ПО.

Таблица 5.3 – Расчёт затрат на спецоборудование

№	Наименование оборудования	Количество единиц оборудования, шт.	Цена за единицу, руб.	Сумма, руб.
1	Ноутбук Dell Vostro 14 (5490-279601)	1	1 639,33	1 639,33
2	Программное обеспечение Matlab (Annual license) для научного руководителя	1	704	704
3	Программное обеспечение Matlab (Student license) для исполнителя	1	74,24	74,24
Итого				2417.57

5.3.3 Расчёт затрат по статье «Основная заработная плата научно-технического персонала»

Величина затрат исчисляется исходя из численности различных категорий исполнителей и трудоёмкости выполнения отдельных видов работ (см. табл. 5.1), тарифных ставок за один день или месячных должностных окладов, используемых премиальных систем оплаты труда исполнителей. Среднемесячная заработная плата работников была взята по данным организации, где проводились исследования.

Расчёт основной заработной платы представлен в таблице 5.4.

Таблица 5.4 – Расчёт затрат на основную заработную плату

Категории исполнителей	Количество исполнителей	Трудозатраты, чел.-дн.	Ставка за 1 день работы, руб.	Сумма тарифной зарплаты, руб..
Научный руководитель	1	12	17.61	221.32
Ответственный исполнитель	1	87	10.36	901.32
Итого	-			1122.64

5.3.4 Дополнительная заработная плата научно-технического персонала

Дополнительная заработная плата научно-технического персонала определяется в процентах от основной заработной платы:

$$З_{д} = \frac{З_{о} \cdot Н_{д}}{100} = \frac{1571.7 \cdot 25}{100} = 392.93 \text{ руб.} \quad (5.1)$$

где $H_{д}$ – норматив дополнительной заработной платы ($H_{д} = 10 - 25\%$).

5.3.5 Отчисления на социальные нужды (отчисления в фонд социальной защиты населения и обязательное страхование)

Отчисления в фонд социальной защиты и по обязательному страхованию от несчастных случаев на производстве определяется от фонда оплаты по труду (основной и дополнительной зарплаты научно-технического персонала) рассчитываются по следующей формуле:

$$P_{\text{соц}} = \frac{(Z_o + Z_d) \cdot H_{\text{соц}}}{100} = \frac{(1122.64 + 392.93) \cdot 35}{100} = 530.45 \text{ руб.} \quad (5.2)$$

где $H_{\text{соц}}$ – установленные законодательно нормативы отчисления на социальные нужды ($H_{\text{соц}} = 35\%$).

5.3.6 Накладные расходы

В статью «Накладные расходы» включаются расходы по управлению и хозяйственному обслуживанию. Они определяются в процентах к основной заработной плате:

$$P_n = \frac{Z_o \cdot H_n}{100} = \frac{1122.64 \cdot 100}{100} = 1122.64 \text{ руб.} \quad (5.3)$$

где H_n – норматив на накладных расходов ($H_n = 50 - 100\%$).

5.3.7 Полная себестоимость

Полная себестоимость научно-исследовательской работы рассчитывается по следующей формуле:

$$C_n = P_z + P_{\text{об}} + Z_o + Z_d + P_{\text{соц}} + P_n = 16.57 + 2417.57 + 1122.64 + 392.93 + 530.45 + 1122.64 = 5602.8 \text{ руб.} \quad (5.4)$$

Данное исследование проводилось в течение 99 дней исследовательской командой в составе двух специалистов и затраты на проведение исследования составили 5602.8 руб.

5.4 Оценка уровня качества научно-технического результата

При оценке научной (научно-технической) результативности НИОКР используют различные критерии. Важнейшими из них являются новизна, значимость для науки и практики, объективность, доказательность, точность. Признаки-критерии выражаются с помощью показателей, отражающих способы измерения степени проявления используемых признаков-критериев при оценке результатов научной деятельности. Характеристика и количественное выражение в баллах по каждому критерию представлены в таблице 5.5.

Таблица 5.5 – Характеристика и количественное выражение в баллах

№	Критерий	Характеристика	Уровень	Оценка	$K_{нzi}$
1	Новизна результата	Сделан элементарный анализ связей и взаимозависимостей между фактами. Дана классификация фактов.	2	0.2	0,0625
2	Значимость для науки и практики	Результат окажет положительное влияние на развитие отдельного научного Направления	2	0.2	0,0625
3	Объективность на основе учета квалификации и компетентности разработчиков и экспертов	Результат получен без участия научных работников высшей квалификации. Рассмотрен и оценен секцией Ученого совета	3	0.9	0,2812
4	Объективность на основе форм его признания	Результат одобрен Ученым (научно-техническим) советом. Опубликован в виде доклада, сделанного на отечественной конференции, и научной статьи в отечественном издании	3	0.9	0,2812
5	Доказательность	Результат получен на основе экспериментальных и теоретических исследований, лабораторных испытаний с использованием математических методов обработки данных	3	1	0,3125
Итого				3.2	1

Показатели каждого признака-критерия классифицируются по пяти уровням качества. По каждому показателю даётся развёрнутая характеристика и количественное выражение в баллах. Чем больше балл, тем выше результат научных исследований по признаку-критерию. Балл, равный 5, соответствует высшей степени проявления признака-критерия.

Количественная оценка уровня научного (научно-технического) результата, полученного в дипломном проекте, рассчитывается по формуле:

$$K_k = \sum_{i=1}^n K_{nzi} \cdot B_{di}, \quad (5.5)$$

где K_k – комплексный показатель достигнутого уровня (качества) результата выполненных исследований; K_{nzi} – нормированный коэффициент значимости i -го критерия, используемого для оценки (должно выполняться условие $\sum_{i=1}^n K_{nzi} = 1$); B_{di} – достигнутый уровень по i -му критерию; n – количество критериев (признаков) научной (научно-технической) прогрессивности и полезности результатов, полученных в дипломном проекте.

На основе приведённых в таблице 1.6 показателей можно рассчитать комплексный показатель достигнутого уровня (качества):

$$K_k = 0.0625 \cdot 2 + 0.0625 \cdot 2 + 0.28125 \cdot 3 + 0.28125 \cdot 3 + 0.31252 \cdot 3 = 2.875 \quad (5.6)$$

Максимально возможное значение комплексного коэффициента качества, взвешенного по всем учитываемым критериям, равно 5, а в общем случае равно $K_k \leq 5$. Чем ближе значение K_k к 5, тем выше научная (научно-техническая) результативность дипломного проекта. Если $K_k \geq 3$, то можно считать, что полученные в дипломном проекте результаты соответствуют современным требованиям. Если $K_k < 3$, можно считать, что дипломный проект не удовлетворяет по уровню качества современным требованиям.

На основании полученного значения комплексного коэффициента качества равного 2.875 можно сделать вывод, что полученные в дипломном проекте результаты немного не соответствуют современным требованиям. Данная работа будет иметь дальнейшее развитие в научных журналах и статьях. Проведение более масштабных исследований для получения доказательных результатов, внедрение данного алгоритма в устройство, а также совершенствование кодовой базы, позволит значительно увеличить комплексный показатель достигнутого уровня.

6 АНАЛИЗ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ СИСТЕМЫ

6.1 Подготовка данных для проведения эксперимента

Для проведения эксперимента необходимо подготовить для каждого испытуемого набор записей, состоящий из четырёх заранее обработанных аудиосигналов:

- 1) сигнал, обработанные системой моделирования потери слуха, настроенной по аудиограмме второй степени;
- 2) сигнал, обработанные системой моделирования потери слуха, настроенной по аудиограмме третьей степени;
- 3) сигналы, пропущенные сначала через разработанную систему повышения разборчивости речи с настройкой по второй степени потери слуха, а после через систему моделирования потери слуха по той же степени;
- 4) сигналы, пропущенные сначала через систему повышения разборчивости речи с настройкой по третьей степени потери слуха, а после через систему моделирования потери слуха по той же степени.

Содержание каждой записи представляет собой слова, произнесенные с секундным интервалом, чтобы испытуемый успел повторить услышанное им слово. Все пять записей для каждого испытуемого содержат разные наборы слов, чтобы избежать фактора запоминания, что может повлиять на конечный результат. В эксперименте было использовано пять наборов по 50 слов в каждом, которые представлены в виде таблиц 6.1-6.5. [23]

Таблица 6.1 – Первый набор

Таблица 1				
бал	бур	док	ром	ваш
вол	май	душ	моль	ноль
гром	грамм	краб	шик	рак
бор	дар	ком	пар	сам
корм	март	торф	плач	брань
бот	рот	мол	хор	ток
морс	пуд	муж	фон	мечь
срам	брак	флюс	мал	хан
бук	вид	путь	пуст	туп
бант	мост	цен	сруб	млад

Таблица 6.2 – Второй набор

Таблица 2				
год	мол	дуй	мель	зад
дай	миг	лоб	марь	куб
путь	сень	сер	сук	лис
рожь	сок	тир	сыпь	лом
кросс	слой	брошь	скот	строй
ноль	бок	был	мой	лад
бить	пей	рак	сак	жар
жест	факт	соль	тик	май
мат	ром	тиф	ныл	пил
жнёт	крут	трап	гнут	драл

Таблица 6.3 – Третий набор

Таблица 3				
таз	груз	лес	снох	встань
нар	век	сталь	знак	пот
быль	вам	рак	шток	трал
воз	час	гроз	чур	сон
дыр	как	газ	кран	толь
лет	жал	лорд	мак	шум
хрен	парт	бот	лень	чех
рот	ныть	сечь	щель	фас
бук	вид	путь	пуст	туп
бант	мост	цен	сруб	млад

Таблица 6.4 – Четвёртый набор

Таблица 4				
бак	босс	тушь	жал	чей
гуд	дал	граф	лад	шаль
драк	конь	пар	мусс	сын
вшит	вдень	сдай	спал	сдуй
враг	злим	плен	стол	сталь
дань	трут	пол	рыть	куй
век	кон	туз	поп	сушь
сей	пир	руль	риск	сук

Таблица 6.5 – Пятый набор

Таблица 5				
чей	жал	тут	бок	бал
шах	ларь	грач	дал	гуж
сыпь	муть	пат	кок	драп
сдут	спарь	сдам	вдеть	вшить
стан	стон	плешь	злишь	враг
суй	ныть	дол	трут	рань
куш	лоб	кус	гон	хек
сук	писк	пуль	нир	сей

После обработки аудиосигналов из пяти наборов слов получилось 20 записей:

- 5 записей, обработанных системой моделирования потери слуха, настроенной по аудиограмме второй степени (далее такой тип записи будет иметь обозначение HL2);

- 5 записей, обработанных системой моделирования потери слуха, настроенной по аудиограмме третьей степени (далее такой тип записи будет иметь обозначение HL3);

- 5 записей, обработанных разработанной системой повышения разборчивости речи, и после системой моделирования потери слуха, настроенной по аудиограмме второй степени (далее такой тип записи будет иметь обозначение corr+HL2);

- 5 записей, обработанных разработанной системой повышения разборчивости речи, и после системой моделирования потери слуха, настроенной по аудиограмме третьей степени (далее такой тип записи будет иметь обозначение corr+HL3).

Схема данных для экспериментального исследования эффективности работы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала по аудиограмме представлена в графическом материале ГУИР 460818.006 ДП.

6.2 Описание эксперимента

Для определения качества и эффективности разработанной системы был проведён эксперимент, в котором приняло участие 15 человек в возрасте от 18 до 22 лет с нормальным слухом. Каждому участнику предлагалось прослушать 5 подготовленных записей. После того, как участник эксперимента слышал

слово, ему необходимо было повторить это слово. Автор работы помечал правильно и неправильно услышанные слова для последующих расчетов.

Номера записей определялись случайным образом для каждого испытуемого. Информация о соответствии номеров таблиц определённому участнику представлена в таблице 6.6.

Таблица 6.6 – Таблица соответствий условий эксперимента

Участник №	HL2	HL3	corr+HL2	corr+HL3
1.	5	1	2	4
2.	2	1	5	4
3.	4	1	3	2
4.	2	1	5	4
5.	3	2	4	5
6.	5	4	2	1
7.	4	5	3	2
8.	2	3	1	4
9.	5	3	2	1
10.	1	5	3	2
11.	1	2	5	3
12.	4	3	1	2
13.	2	5	4	1
14.	3	5	2	4
15.	5	1	2	4

После получения необходимой информации были проведены расчеты для установления эффективности разработанной системы. Результаты расчетов отражены в пункте 6.3.

6.3 Результаты эксперимента

По результатам эксперимента была составлена таблица 6.7, отражающая процентное соотношение правильно услышанных слов ко всем словам из тестовых звуковых файлов.

Таблица 6.7 – Результаты эксперимента

Участник №	HL2	HL3	corr+HL2	corr+HL3
1.	26	16	90	88
2.	28	10	98	84
3.	30	14	92	74
4.	42	8	88	82
5.	24	14	84	90
6.	28	10	94	84
7.	26	16	92	80
8.	22	12	94	72
9.	24	10	96	88
10.	18	14	90	84
11.	22	8	88	84
12.	18	8	86	78
13.	24	12	92	86
14.	32	10	90	78
15.	20	14	90	84

Из представленных значений в таблице видно, что разработанная система повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала действительно помогает решить проблему потери слуха. Система почти полностью восстанавливает сигнал при разных степенях потери слуха, что говорит о гибкости системы.

После проведения экспериментального исследования эффективности работы разработанной системы необходимо провести проверку значимости полученных результатов для исключения возможности получения случайных результатов из-за небольшой выборки. В данной работе проверка выполнена при помощи *ttest*, использующая парный t-критерий. В результате проверки значимости результатов эксперимента нулевая гипотеза была отвергнута. Итоговый уровень значимости составил $p < 0.01$, что говорит о том, что вероятность того, что результаты эксперимента не являются значимыми очень мала ($< 1\%$).

ЗАКЛЮЧЕНИЕ

В настоящее время задача коррекции слуха является актуальной в связи с ростом количества людей с проблемами слуха.

В работе рассмотрена система повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала.

Проект системы соответствует требованиям технического задания. В процессе проектирования учтены:

- независимость от степени потери слуха;
- декомпозиция сигнала на основе «слухового» банка фильтров;
- выбор менее ресурсоёмких алгоритмов цифровой обработки сигналов.

Требование независимости от степени потери слуха для системы повышения разборчивости речи было удовлетворено путем применения системы, которая автоматически настраивается согласно характеристикам слуха человека (аудиограмме).

Декомпозиция входного производилась с помощью «слухового» банка гамматон-фильтров. Использованный банк согласован с работой человеческого уха и имитирует механизм частотной декомпозиции звука в ухе.

Выбор менее ресурсоёмких алгоритмов цифровой обработки сигналов обеспечило достаточно быструю работу программной реализации. Данное требование было выполнено благодаря внедрению алгоритмических оптимизаций в код: расчет коэффициентов субполосного усиления производился не для каждого отсчета, а для каждого фрейма ($N_frames = 256$), что значительно сократило время обработки сигнала.

Выполненная программная реализация проекта системы удовлетворяет следующим требованиям технического задания:

- минимальная частота дискретизации: 12 кГц;
- число каналов во входном звуковом сигнале: 1;
- минимальное количество каналов банка фильтров: 20.

Проведение тестирования работы системы повышения разборчивости речи, что результаты эксперимента являются значимыми.

Основными преимуществами разработанной системы является то, что моделирование производится неравнополосным банком гамматон-фильтров, имитирующего механизм частотной декомпозиции звука в ухе. Кроме того, компрессор динамического диапазона сигнала настраивается автоматически по аудиограмме слабослышащего человека. Недостатком является то, что на

низких частотах использованный банк фильтров имеет длинную импульсную характеристику, что увеличивает количество операций с сигналом.

Настройка КДД выполнена согласно аудиограммам для второй и третьей степени потери слуха.

Для проверки работоспособности системы выполнено её MATLAB-моделирование. В качестве входного сигнала использовался речевой сигнал, спектрограммы и временное представление которого представлены в графическом документе ГУИР 460818.005 ПЛ к данной работе. Кроме того, на плакате отражены спектрограммы сигнала, обработанного методом модерирования потери слуха и методом повышения разборчивости речи для второй степени потери слуха. Этот сигнал является пятнадцатисекундным отрезком аудиодорожки (таблица 1), которую давали прослушать участникам эксперимента. Данные спектрограммы объясняют результаты эксперимента, которые также представлены на данном плакате в виде гистограмм.

Из результатов, отраженных на плакате видно, что используемая система КДД выполняет частотно-зависимое усиление, что, в свою очередь, приводит к увеличению уровня исходного звукового сигнала и положительно влияет на разборчивость речи. Искажения, появляющиеся в результате обработки, не влияют на общий уровень разборчивости речи и возможность идентифицировать говорящего человека.

Стоит отметить, что существует возможность проведения усовершенствования полученной системы в результате проектирования, например путем добавления дополнительной точки перегиба (болевого порога) для характеристики вход/выход КДД. Это позволит избежать такой формы искажения звука, как клиппинг.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- [1] Королева И. В., Введение в аудиологию и слухопротезирование. – СПб.:КАРО, 2012. – 400 с. – (Серия «Специальная педагогика»).
- [2] Вашкевич М.И., Азаров И.С., Петровский А.А., Косинусно-модулированные банки фильтров с фазовым преобразованием: реализация и применение в слуховых аппаратах. – Москва: Горячая линия «Телеком», 2014 – 219 с.
- [3] Huan-Yu Dong, Chang-Myung Lee, Speech intelligibility improvement in noisy reverberant environments based on speech enhancement and inverse filtering. – Journal on Audio, Speech, and Music Processing, 2018.
- [4] Йен-Дэ Лю, Рональд Й. Чан, Юй Цао, Йи-Пин Чанг, Новая методика уменьшения частоты для пользователей слуховых аппаратов, говорящих на северокитайском языке. – Глобальная конференция IEEE по обработке сигналов и информации (GlobalSIP), 2015.
- [5] Арпита Нагеш К, Кавья П, Кавьяшри Б К, Крутишри К С., Удаление шума с помощью метода адаптивной фильтрации для людей страдающих нейросенсорной тугоухостью. – Международная конференция по современным тенденциям в области компьютеров, электротехники, электроники и связи (ICSTCEEC), 2017.
- [6] Хунмэй Ху, Цзиньцю Сан, Марк Э Лутман, Стефан Блик, Моделирование потери слуха с использованием компрессионных слуховых gammachirp-фильтров. – Международная конференция IEEE по акустике, речи и обработке сигналов (ICASSP), 2011.
- [7] Anna Warzybok, Jan Rennies, Thomas Brand, Birger Kollmeier, Prediction of binaural speech intelligibility in normal-hearing and hearing-impaired listeners: a psychoacoustically motivated extension. – Annual German Congress on Acoustics, 2014.
- [8] Таварткиладзе Г.А., Гвелесиани Т.Г., Клиническая аудиология. – М.:Святигор Пресс, 2003. – 74 с.
- [9] Альтман Я.А. Слуховая система. – Наука. 1990. – 620 с. – (Основы современной физиологии).
- [10] Patterson R. D., Nimmo-Smith I., Holdsworth J., Rice P., An efficient auditory filterbank based on the gammatone function / A meeting of the IOC Speech Group on Auditory Modelling at RSRE. 2 (7) – 1987.
- [11] Ambikairajahm E., Epps J., Lin L., Wideband speech and audio coding using gammatone filter banks / School of Electrical Engineering and Telecommunications, – 2021.

[12] Jiang Y., Zu Y., Liu L., Wang Q., Ren P. and Zhou H., Gammatone filterbank based energy masking algorithm for active hearing protection system / 5th International Conference on BioMedical Engineering and Informatics (BMEI 2012), Chongqing, China, – 2012. – P. 537–540.

[13] Вашкевич М. И., Азаров И. С., Биоинспирированный и психоакустически мотивированные методы анализа звуковых сигналов. – Кафедра электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники.

[14] Zoelzer U., Digital audio signal processing / John Wiley & Sons Software, – 2008. – 334 p.

[15] Wikipedia [Электронный ресурс]. – Электронные данные. – Режим доступа: https://en.wikipedia.org/wiki/Bark_scale

[16] Long-tao Cao, Ru-wei Li, Yong-qiang Shi, Shuai Wang, Loudness Compensation Method Based On Human Auditory For Digital Hearing Aids. – 7th International Conference on BioMedical Engineering and Informatics (BMEI), 2014.

[17] ГОСТ Р ИСО 226—2009, Стандартные кривые равной громкости, Москва. – 2010. – 15 с.

[18] Вашкевич, М. И., Биоинспирированный и психоакустически мотивированный анализ звуковых сигналов / Вашкевич М. И., Азаров И. С. // Информационные технологии и системы 2020 (ИТС 2020) = Information Technologies and Systems 2020 (ITS 2020) : материалы международной научной конференции, Минск, 18 ноября 2020 г. / Белорусский государственный университет информатики и радиоэлектроники ; редкол. : Л. Ю. Шилин [и др.]. – Минск, 2020. – С. 9–12.

[19] Вашкевич М. И., Петровский А. А., Компрессор речевых сигналов для слуховых аппаратов на основе банка кохлеарных фильтров / Цифровая обработка сигналов и ее применение: труды 14-й междунар. конф., Россия, Москва. – 2012. – Т. 1. – С. 87–91.

[20] MATLAB [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://www.mathworks.com/products/matlab.html>

[21] Zwicker E., Fastl H., Psychoacoustics, Facts and Models, Berlin: Springer Verlag, 1990.

[22] Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2015. – 400 с.: ил.

[23] ГОСТ 16600—72, Требования к разборчивости речи методы артикуляционных измерений, Москва. – 1972. – 92 с.

ПРИЛОЖЕНИЕ А

(обязательное)

Код программы

А.1 Программная реализация алгоритм системы повышения разборчивости речи

```
close all; clc;

% Инициализация констант
K = 21;
frame_size = 256;
coefficient_a = 0.9900; % attack time coefficient
coefficient_b = 0.9992; % release time coefficient
freq_aud = [125 250 500 750 1000 2000 3000 4000 6000 8000]; % сетка частот
% HT = [24 30 36 35 34 40 42 44 41 40]; % 1я степень тугоухости
HT = [43 50 50 47 45 50 52 53 58 60]; % 2я степень тугоухости
% HT = [53 66 60 62 64 54 60 67 69 70]; % 3я степень тугоухости
% HT = [70 75 76 78 80 77 79 80 82 84]; % 4я степень тугоухости
input_signal = zeros(1,2^12);
input_signal(1,1) = 1;
bark_sc=1.5:1:K-0.5;
CL_in_2 = [-40.2,-36.1,-39.1,-37.5,-36.7,-37.2,-42.4,-39,-40.6,-41.8,-41.4,-37,-34.4,-42.6,-41.4,-41.4,-41.6,-40.5,-41.5,-39.3];

CL_out_2 = [-17.6,-16.3,-22.1,-19.5,-16.9,-21.2,-17.4,-22.8,-14.4,-18.1,-16.5,-18.9,-15.1,-18.5,-22.5,-15.6,-15.9,-24.7,-22.1,-17.8];

CL_in_3 = [-40.2,-36.1,-39.1,-37.5,-36.7,-37.2,-42.4,-39,-40.6,-41.8,-41.4,-37,-34.4,-42.6,-41.4,-41.4,-41.6,-40.5,-41.5,-39.3];

CL_out_3 = [-7.6,-5.9,-10.5,-5.2,-7.8,-11.9,-9.3,-5.5,-11.7,-11.7,-10.9,-9.7,-4.9,-9.6,-11.4,-8,-7.9,-11.8,-11,-9.7];

CL_in = CL_in_2;
CL_out = CL_out_2;

% Выбираются центральные частоты банка фильтров в соответствии с
% психоакустической шкалой Барков
fc = bark2frq(bark_sc);
[spl, freq_iso] = iso226_rev(0);

% Считывается входного звукового сигнала
addpath('input_signals');
filename = 'черный ворон 25с';
cond = '_corr_HL3';
[audio, fs] = audioread([filename '.mp3']);

audio_input_signal = audio(:, 1)';

output_signal = gammatoneFast(input_signal, fc);
create_gammaton_bank_filter_figure(output_signal, K, fs, freq_aud);

from_digdb_to_spl = 105;

% Нахождение порога слышимости
interped_HT = interp1(freq_aud, HT, fc, 'linear');
interped_spl = interp1(freq_iso, spl-from_digdb_to_spl, fc, 'linear');
hi = interped_HT+interped_spl;
```

```

% Построение графиков аудиограммы, порогов слышимости в норме и при наличии
% патологии
create_hearing_thresholds_figure(freq_aud, HT, hi, freq_iso, spl, fc);

% Происходит разложение входного сигнала на субполосы
audio_output_signal = gammatoneFast(audio_input_signal, fc);

% Производится коррекция амплитуд
corrected_audio_output_signal = audio_output_signal.*1.8;

% Инициализация переменных
N_frames=floor(length(corrected_audio_output_signal(1,:))/frame_size);
p_start = repmat(2^(-10),length(1:K-1),1);
g_begin = ones(length(1:K-1));
synthesized_audio_output_signal = zeros(1, N_frames * frame_size);
processed_audio_output_signal = zeros(K, N_frames * frame_size);
y = zeros(1, N_frames * frame_size);

% Конфигурация графика компрессора динамического диапазона
figure;
configure_figure_settings('Компрессор динамического диапазона', 'Входная
мощность, дБ', 'Выходная мощность, дБ');
hold on;
grid on;

for M=1:K-1
    % Определение точек перегиба компрессора динамического диапазона
    drc_in = [-120 interp_spl(1,M) CL_in(M) 0];
    drc_out = [-120 hi(1,M) CL_out(M) 0];
    % Построение графика компрессора динамического диапазона
    plot (drc_in, drc_out, 'k', 'LineWidth', 1);
    plot (drc_in, drc_out, 'ok', 'LineWidth', 1.5);

    % Реализация алгоритма нахождения коэффициентов усиления для каждой
    % субполосы
    for N=1:N_frames
        i_beg = (N-1)*frame_size+1;
        i_end = N*frame_size;
        x_frame = corrected_audio_output_signal(M, i_beg : i_end);

        % Разбиение каждой субполосы на фреймы для оптимизации алгоритма
        for j=1:frame_size
            % Производится оценка уровня мощности
            p_cur = x_frame(j)^2;

            if p_cur > p_start(M)
                p_start(M) = p_start(M) * coefficient_a + p_cur * (1 -
coefficient_a);
            else
                p_start(M) = p_start(M) * coefficient_b + p_cur * (1 -
coefficient_b);
            end
        end

        p_end = p_start(M);
        % Полученный уровень мощности переводится в дБ
        p_end_db = 10*log10(p_end);
        % Вычисляется коэффициент усиления в каждой субполосе
        p_out_db = interp1(drc_in, drc_out, [p_end_db], 'linear');
        g_end_db = p_out_db - p_end_db;
        % Коэффициент усиления переводится из дБ в линейный масштаб
        g_end = 10^(g_end_db / 20);
        g_line = interp1([0 frame_size], [g_begin(M) g_end],1:frame_size,
'linear');
    end
end

```

```

        g_begin(M) = g_end;
        % Производится усиления
        y(i_beg:i_end) = x_frame.*g_line;
    end

    % Синтез выходного сигнала
    synthesized_audio_output_signal = synthesized_audio_output_signal + y;
    processed_audio_output_signal(M, :) = y;
end

create_characteristic_drc_input_output_figure(drc_in, drc_out);
create_input_output_subband_signals_figure(corrected_audio_output_signal,
processed_audio_output_signal);

% Построение графиков временного представления и спектрограмм входного и
выходного сигнала
create_input_output_signal_figure(audio_input_signal,
synthesized_audio_output_signal, fs, freq_aud);

% Запись выходного звукового сигнала
audiowrite([filename cond '.wav'], synthesized_audio_output_signal, fs);

function create_gammaton_bank_filter_figure(corrected_output_signal, K, fs,
freq_aud)
    figure;
    hold on;
    set(gca, 'YLim', [-90 0]);

    configure_figure_settings('', 'Частота, Гц', 'Амплитуда, дБ');

    for N=1:K-1
        [h, w] = freqz(corrected_output_signal(N,:),1,4096);
        hDb = 20*log10(h);
        wHz = w/pi*(fs/2);
        plot(wHz, hDb, 'LineWidth',2.5);
    end
    grid on;
    ylim ([-72 0]);
end

function create_hearing_thresholds_figure(freq_aud, HT, hi, freq_iso, spl,
fc)
    figure;
    array = 1:length(freq_aud);
    hold on;
    plot (array, -1*HT, 'k', 'LineWidth', 2);
    ylim ([-70 10]);
    xlim ([array(1) array(end)]);
    plot (array,-1*HT, 'ok', 'LineWidth', 2);
    grid on;
    set(gca, 'YTickLabel', [70:-10:-10]); %90,80,70,60,50,40,30,20,10,0,-10
    set(gca, 'XTickLabel',freq_aud);
    set(gca, 'XTick',array(1):array(end));
    set(gca, 'Clim', [-65 15]);
    configure_figure_settings('Аудиограмма при 2-ой степени тугоухости',
'Частота, Гц', 'Потеря слуха, дБ');

    figure;
    plot(freq_iso, spl-80, 'k', 'LineWidth', 2);
    configure_figure_settings('Абсолютный порог слышимости', 'Частота, Гц',
'Амплитуда, дБ');
    xlim ([freq_iso(1) freq_iso(end)]);
    ylim ([-90 0]);
    grid on;

```

```

figure;
plot(fc, hi, 'k', 'LineWidth', 2);
configure_figure_settings('Порог слышимости при патологии', 'Частота,
Гц', 'Амплитуда, дБ');
xlim ([fc(1) fc(end)]);
ylim ([-90 0]);
grid on;
end

function configure_figure_settings(titleSpecgram, xlabelSpecgram,
ylabelSpecgram)
set(gca, 'Clim', [-65 15]);
title(titleSpecgram);
xlabel(xlabelSpecgram);
ylabel(ylabelSpecgram);
set(gca, 'FontName', 'Times New Roman');
set(gca, 'FontSize', 18);
end

function create_characteristic_drc_input_output_figure(drc_in, drc_out)
% Передаточная характеристика устройства
figure;
plot(drc_in, drc_out);
grid on;
configure_figure_settings('DRC Characteristic', 'Input magnitude, dB',
'Output magnitude, dB');
xlim([-90 0]);
ylim([-90 0]);
end

function create_input_output_signal_figure(audio_input_signal,
synthesized_audio_output_signal, fs, freq_aud)
figure;

subplot(2, 2, 1);
specgram(audio_input_signal, 512, fs, kaiser(512,7),475);
ylim ([0 freq_aud(end)]);
configure_figure_settings('Спектрограмма входного сигнала', 'Время, с',
'Частота, Гц');

subplot(2, 2, 2);
specgram(synthesized_audio_output_signal, 512, fs, kaiser(512,7),475);
ylim ([0 freq_aud(end)]);
configure_figure_settings('Спектрограмма выходного сигнала', 'Время, с',
'Частота, Гц');

subplot(2, 2, 3);
plot((1:length(audio_input_signal)) / fs, audio_input_signal);
ylim ([-2 2]);
xlim ([0 25]);
configure_figure_settings('Временное представление сигнала', 'Время, с',
'Амплитуда, дБ');

subplot(2, 2, 4);
plot((1:length(synthesized_audio_output_signal)) / fs,
synthesized_audio_output_signal);
ylim ([-2 2]);
xlim ([0 25]);
configure_figure_settings('Временное представление сигнала', 'Время, с',
'Амплитуда, дБ');
end

```



```

function
create_input_output_subband_signals_figure(corrected_audio_output_signal,
processed_audio_output_signal)
    figure;
    subplot(2,1,1);
    plot(corrected_audio_output_signal);
    configure_figure_settings('Input signal', 'Time, s', 'Magnitude, dB');
    ylim([-1 1]);

    subplot(2,1,2);
    plot(processed_audio_output_signal);
    configure_figure_settings('Output signal', 'Time, s', 'Magnitude, dB');
    ylim([-1 1]);
end

function create_gammaton_bank_filter_figure(corrected_output_signal, K, fs,
freq_aud)
    figure;
    hold on;
    set(gca, 'YLim', [-90 0]);

    configure_figure_settings('', 'Частота, Гц', 'Амплитуда, дБ');

    for N=1:K-1
        [h, w] = freqz(corrected_output_signal(N,:),1,4096);
        hDb = 20*log10(h);
        wHz = w/pi*(fs/2);
        plot(wHz, hDb, 'LineWidth',2.5);
    end
    grid on;
    ylim ([-72 0]);
end

function f = bark2frq(b)
f=52547.6*(26.28-b-(22.11-1.1*b).*(b>20.1)/6.1-(3*b-6).*(b<2)/17).^(-1)-1960;

function [bm,env,delay] = gammatoneFast(x,cfs,fs,align)
    if nargin < 3
        fs = 16000; % default sampling frequency
    end
    if nargin < 4
        align = false; % default phase alignment
    end

    % check inputs
    assert(isvector(x) & isnumeric(x), 'iosr:gammatoneFast:invalidX', 'x must
be a vector')
    assert(isvector(cfs) & isnumeric(cfs), 'iosr:gammatoneFast:invalidCfs',
'cfs must be a vector')
    assert(isscalar(fs), 'iosr:gammatoneFast:invalidFs', 'fs must be a
scalar')
    assert(islogical(align) & numel(align)==1,
'iosr:gammatoneFast:invalidAlign', 'align must be logical')

    % number of frequency channels
    numchans = length(cfs);

    filterOrder = 4; % filter order
    gL = 2^nextpow2(0.128*fs); % gammatone filter length at least 128 ms
    b = 1.019.*24.7.*(4.37.*cfs./1000+1); % rate of decay or bandwidth

    gt = zeros(gL,numchans); % Initialise IR
    tc = zeros(size(cfs)); % Initialise time lead
    phase = 0;

```

```

tpt=(2*pi)/fs;
gain=((1.019.*b.*tpt).^filterOrder)./6; % based on integral of impulse

tmp_t = (0:gL-1)/fs;

% calculate impulse response
for i = 1:numchans
    if align
        tc(i) = (filterOrder-1)./(2*pi*b(i));
        phase = -2*pi*cfs(i)*tc(i);
    end
    gt(:,i) = gain(i)*fs^3*tmp_t.^(filterOrder-1).*exp(-
2*pi*b(i)*tmp_t).*cos(2*pi*cfs(i)*tmp_t+phase);
end

% if input is row vector, transpose to column vector
rot = false;
if size(x,1)==1
    x = x';
    rot = true;
end

% gammatone filtering using FFTFILT
bm = fftfilt(gt, repmat(x,1,numchans));

% Hilbert envelope
env = abs(hilbert(bm));

% delay due to time lead
delay = round(tc.*fs);

% remove time lead
for i = 1:numchans
    bm(:,i) = [bm(delay(i)+1:end,i); zeros(delay(i),1)];
    env(:,i) = [env(delay(i)+1:end,i); zeros(delay(i),1)];
end

% transpose output if necessary
if rot
    bm = bm';
    env = env';
end

end

```

А.2 Программная реализация алгоритм определения значимости результатов

```
function check_results()
% Значения для сигнала 3-ей степени потери слуха
% p_1 = [16 88];
% p_2 = [10 84];
% p_3 = [14 74];
% p_4 = [8 82];
% p_5 = [14 90];
% p_6 = [10 84];
% p_7 = [16 80];
% p_8 = [12 72];
% p_9 = [10 88];
% p_10 = [14 84];
% p_11 = [8 84];
% p_12 = [8 78];
% p_13 = [12 86];
% p_14 = [10 78];
% p_15 = [14 84];

% Значения для сигнала 2-ой степени потери слуха
p_1 = [26 90];
p_2 = [28 98];
p_3 = [30 92];
p_4 = [42 88];
p_5 = [24 84];
p_6 = [28 94];
p_7 = [26 92];
p_8 = [22 94];
p_9 = [24 96];
p_10 = [18 90];
p_11 = [22 88];
p_12 = [18 86];
p_13 = [24 92];
p_14 = [32 90];
p_15 = [20 90];

res =
[p_1;p_2;p_3;p_4;p_5;p_6;p_7;p_8;p_9;p_10;p_11;p_12;p_13;p_14;p_15];
dif_1 = p_1(2) - p_1(1);
dif_2 = p_2(2) - p_2(1);
dif_3 = p_3(2) - p_3(1);
dif_4 = p_4(2) - p_4(1);
dif_5 = p_5(2) - p_5(1);
dif_6 = p_6(2) - p_6(1);
dif_7 = p_7(2) - p_7(1);
dif_8 = p_8(2) - p_8(1);
dif_9 = p_9(2) - p_9(1);
dif_10 = p_10(2) - p_10(1);
dif_11 = p_11(2) - p_11(1);
dif_12 = p_12(2) - p_12(1);
dif_13 = p_13(2) - p_13(1);
dif_14 = p_14(2) - p_14(1);
```

```

dif_15 = p_15(2) - p_15(1);

dif = dif_1 + dif_2 + dif_3 + dif_4 + dif_5 + dif_6 + dif_7 + dif_8 +
dif_9 + dif_10 + dif_11 + dif_12 + dif_13 + dif_14 + dif_15;
dif = dif/15;
min_dif = min ([dif_1 dif_2 dif_3 dif_4 dif_5 dif_6 dif_7 dif_8 dif_9
dif_10 dif_11 dif_12 dif_13 dif_14 dif_15]);
max_dif = max ([dif_1 dif_2 dif_3 dif_4 dif_5 dif_6 dif_7 dif_8 dif_9
dif_10 dif_11 dif_12 dif_13 dif_14 dif_15]);

disp ('Средний прирост разборчивости = ' + string(dif) + ' %');
disp ('Минимальный процент разборчивости = ' + string(min_dif) + ' %');
disp ('Максимальный процент разборчивости = ' + string(max_dif) + ' %');


figure;
bar(res);
xlabel ('Участник');
ylim ([0 112]);
xlim ([0.5 15.5]);
ylabel('Процент разборчивости');
legend ('С моделированием 2-ой степени потери слуха', 'С коррекцией
слуха');


figure;
bar(res);
xlabel ('Участник');
ylim ([0 112]);
xlim ([0.5 15.5]);
ylabel('Процент разборчивости');
legend ('С моделированием 3-ей степени потери слуха', 'С коррекцией
слуха');


alpha = 0.05;
X = [p_1(1) p_2(1) p_3(1) p_4(1) p_5(1) p_6(1) p_7(1) p_8(1) p_9(1)
p_10(1) p_11(1) p_12(1) p_13(1) p_14(1) p_15(1)];
Y = [p_1(2) p_2(2) p_3(2) p_4(2) p_5(2) p_6(2) p_7(2) p_8(2) p_9(2)
p_10(2) p_11(2) p_12(2) p_13(2) p_14(2) p_15(2)];
[H,p] = ttest(X, Y, 'Alpha', alpha)
end

```

ПРИЛОЖЕНИЕ Б

(обязательное)

Отчёт о проверке на заимствование



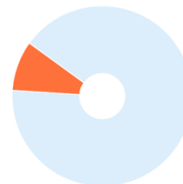
Автор: Semenuk Natasha
Проверяющий: Semenuk Natasha (nata.semenuck@gmail.com) / ID: 9028012)
Отчет предоставлен сервисом «Антиплагиат» - users.antiplagiat.ru

ИНФОРМАЦИЯ О ДОКУМЕНТЕ

№ документа: 4
Начало загрузки: 26.05.2021 09:47:19
Длительность загрузки: 00:00:03
Имя исходного файла: Пояснительная записка-конвертирован.pdf
Название документа: Пояснительная записка
Размер текста: 98 кБ
Символов в тексте: 100732
Слов в тексте: 11507
Число предложений: 520

ИНФОРМАЦИЯ ОБ ОТЧЕТЕ

Начало проверки: 26.05.2021 09:47:22
Длительность проверки: 00:00:14
Комментарии: не указано
Модули поиска: Модуль поиска inet



ЗАИМСТВОВАНИЯ
8,53%

САМОЦИТИРОВАНИЯ
0%

ЦИТИРОВАНИЯ
0%

ОРИГИНАЛЬНОСТЬ
91,47%

Рисунок Б.1 – Отчёт о проверке на заимствование

_____ Н.А. Семенюк

Перв. примен.		Зона	Обозначение	Наименование	Дополнительные сведения					
				Текстовые документы						
			БГУИР ДП 1-40-02 02 01 016 ПЗ	Пояснительная записка	72 с.					
Справочный №				Отзыв руководителя	1 с.					
				Рецензия	1 с.					
				Графические документы						
			ГУИР 460818.001 ПЛ	Структура системы	Формат А1					
				разборчивости речи на основе						
				модели слухового восприятия						
				и компрессии динамического						
				диапазона сигнала						
			ГУИР 460818.002 ПД	Схема программы реализации						
				банка фильтров для системы	Формат А1					
				повышения разборчивости						
				речи на основе модели						
				слухового восприятия и						
				компрессии динамического						
				диапазона сигнала						
Подпись и дата										
Инв. № дубл.										
Взам. Инв. №										
Подпись и дата										
Инв. № подл.		Изм.	Лист	№ докум.	Подп.	Дата	БГУИР ДП 1-40 02 02 01 016 Д1			
		Разраб.	Семенюк Н.А.				Система повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала	Лит.	Лист	Листов
		Пров.	Порхун М.И.					Т	70	72
		Т.контр.	Порхун М.И.					Кафедра ЭВС гр. 750701		
		Н.контр.	Лихачёв Д.С.							
Утв.	Азаров И.С.									

Перв. примен.		Зона	Обозначение	Наименование	Дополнительные сведения
				Графические документы	
			ГУИР 460818.003 ПД	Схема программы системы повышения разборчивости речи	Формат А1
Справочный №				на основе модели слухового восприятия и компрессии динамического диапазона сигнала	
			ГУИР 460818.004 ПЛ	Настройка системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала	Формат А1
			ГУИР 460818.005 ПЛ	Экспериментальное исследование эффективности работы системы повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала по аудиограмме	Формат А1
Подпись и дата					
Инв. № дубл.					
Взам. Инв. №					
Подпись и дата					
<div> <div>Изм.</div> <div>Лист</div> <div>№ докум.</div> <div>Подп.</div> <div>Дата</div> </div> <div>БГУИР ДП 1-40 02 02 01 016 Д1</div> <div> <div>Лист</div> <div>71</div> </div>					

Перв. примен.		Зона	Обозначение	Наименование		Дополнительные сведения			
				<u>Графические документы</u>					
			ГУИР 460818.006 ПД	Схема данных для экспериментального		Формат А1			
Справочный №				исследования эффективности					
				работы системы повышения					
				разборчивости речи на основе					
				модели слухового восприятия					
				и компрессии динамического					
				диапазона сигнала по					
				аудиограмме					
Подпись и дата									
Инв. № дубл.									
Взам. Инв. №									
Подпись и дата									
		Изм.	Лист	№ докум.	Подп.	Дата	БГУИР ДП 1-40 02 02 01 016 Д1		Лист
									72