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 ДМНАМИЧЕСКОГО ДИАПАЗОНА СИГНАЛА

БГУИР ДП 1-40 02 02 021 ПЗ

Студент

Руководитель

Консультанты:

*от кафедры ЭВС*

*по экономической части*

Нормоконтролер

Рецензент

Н.А. Семенюк

М.И. Порхун

М.И. Порхун

Т.А. Рыковская

Д.С. Лихачев

Д.С. Лихачев

Минск 2021

CОДЕРЖАНИЕ

[1.1 Метод 1 4](#_Toc70454144)

[2.1 Анализ требований к алгоритму 5](#_Toc70454145)

[2.2 Анализ требований к программной реализации 5](#_Toc70454146)

[2.3 Выбор и обоснование метода решения задачи 5](#_Toc70454147)

[3.1 Слуховое восприятие 7](#_Toc70454148)

[3.1.1 Строение слуховой системы 7](#_Toc70454149)

[3.1.2 Управление светодиодами с помощью ATmega328 9](#_Toc70454150)

[3.1.3 Управление тактовыми кнопками с помощью ATmega328 9](#_Toc70454151)

[3.2 Описание работы системы по принципиальной схеме 9](#_Toc70454152)

[4.1 Разработка схемы алгоритма работы системы и программы 10](#_Toc70454153)

[4.2 Описание алгоритма работы системы и программы 10](#_Toc70454154)

[5.1 Выбор системы моделирования 12](#_Toc70454155)

[5.2 Описание процесса моделирования 12](#_Toc70454156)

Введение

Звук возникает в результате вибрации любых объектов с определенной частотой. Пространство, в котором распространяется звук, называется звуковым полем. Звуки многих музыкальных инструментов являются результатом колебаний натянутых струн (скрипка, гитара, фортепиано) или натянутой кожи (барабан). 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. Отсюда следует, что человек слышит мозгом, а не ухом [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. Таким образом, 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].

# Обзор существующих методов повышения разборчивости речи на основе модели слухового восприятия

## Метод 1

Что хорошо, что плохо, что за метод и структурная схема

# Анализ технического задания

## Анализ требований к алгоритму

Основным требованием к алгоритму работы системы повышения разборчивости речи на основе модели слухового восприятия является использование оптимальных затрат вычислительных ресурсов, чтобы достичь малой алгоритмической задержки, для работы в режиме реального времени.

Оптимальность затрат вычислительных ресурсов можно получить путем применения менее ресурсоемких алгоритмов цифровой обработки сигналов. Поэтому при осуществлении анализа и разработки алгоритма работы системы необходимо обратить внимание на оценку производительности.

## 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.

## 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 систем 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.

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.

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 системы анализа/синтез на основе неравнополосного банка гамматон-фильтров [2].

# Разработка структуры и алгоритма работы системы повышения разборчивости речи.

## 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.

### 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.

Строение уха представлено на рисунке 3.1 [1].

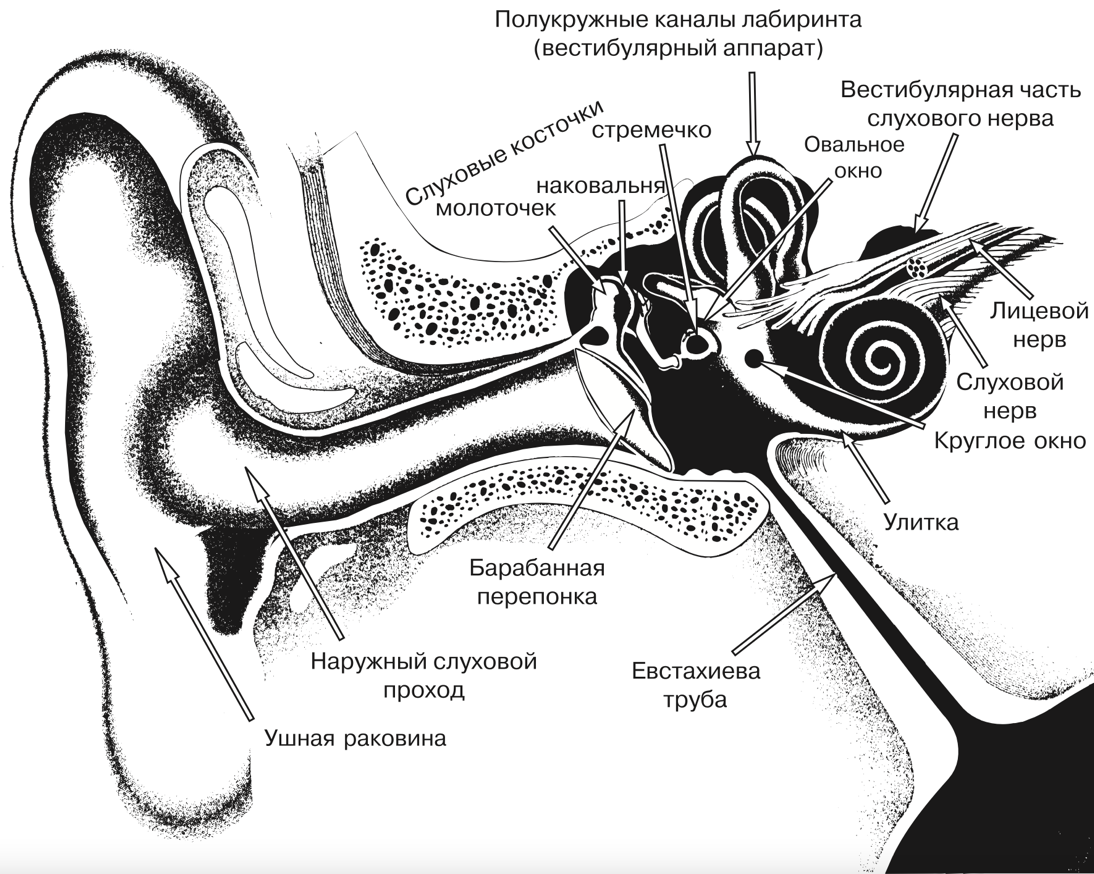


Рисунок 3.1 – Строение уха

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.

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 звуковое 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, 3].

### Возникновение слуховых ощущений

Существуют 2 способа передачи звуковых колебаний волосковым клеткам в улитке:

* 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 непроизвольные 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.) [3, 4].

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

### 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.

* 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.
* 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.
* 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 условия 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.
* 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.
* 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 восприятие речи [1].

## 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: тональный 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.

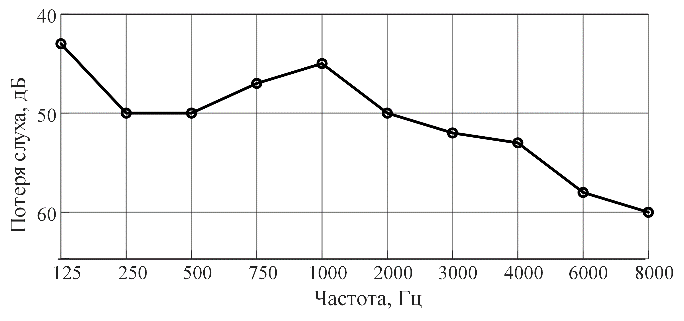


Рисунок 3.2 – Пример аудиограммы

Аудиограмма может быть восходящей (чаще при кондуктивной тугоухости), нисходящей (чаще при сенсоневральной тугоухости), горизонтальной, а также других конфигураций [3].

Настройка компрессора динамического диапазона сигнала проводилась с использованием аудиограммы, поэтому понятие "аудиограмма" играет важную роль в этом проекте.

## Разработка структуры системы повышения разборчивости речи

В данной работе рассмотрена система повышения разборчивости речи на основе модели слухового восприятия и компрессии динамического диапазона сигнала. Структура системы представлена на рисунке 4.1.

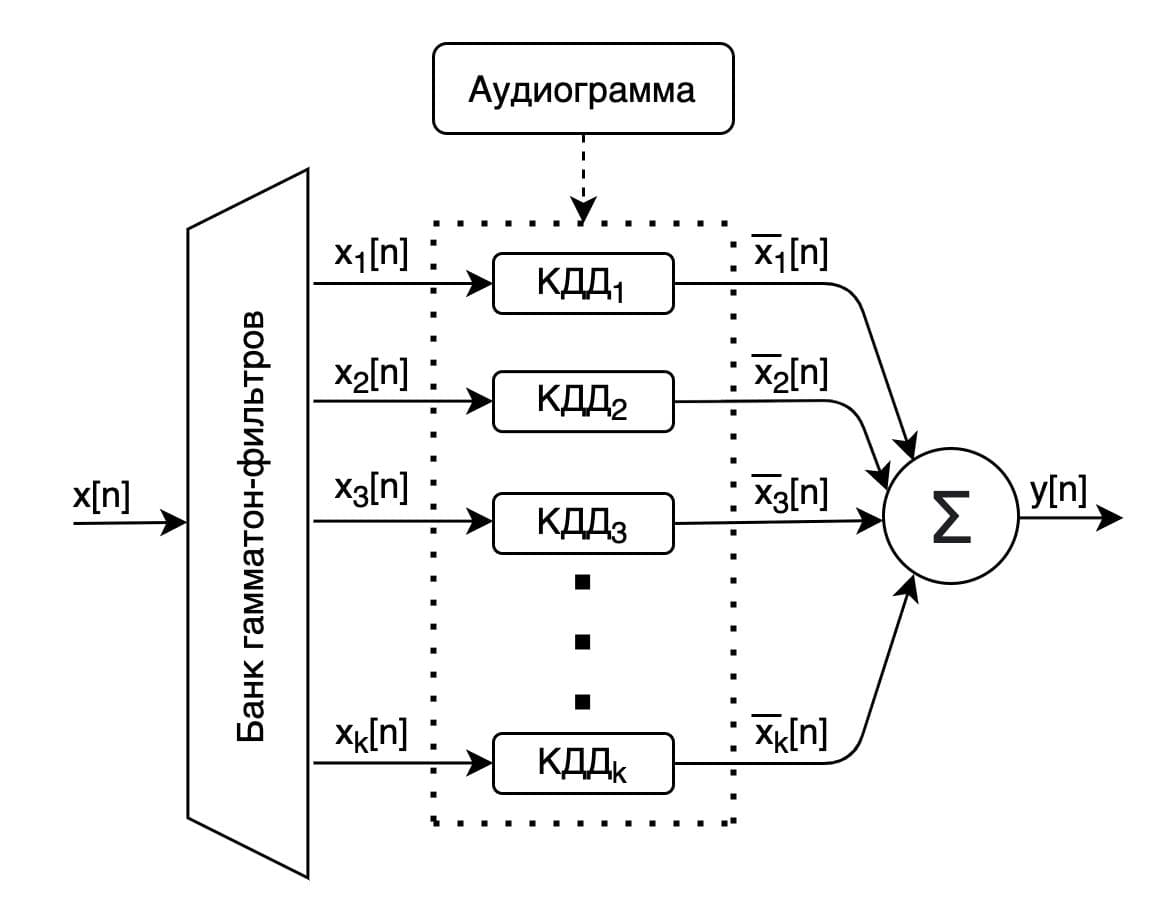


Рисунок 3.3 – Структура системы повышения разборчивости речи

# Проектирование программного обеспечения

## Разработка схемы алгоритма работы системы и программы

## Описание алгоритма работы системы и программы

# Моделирование работы системы

## Выбор системы моделирования

## Описание процесса моделирования

Заключение

Список использованных источников

1. Рабинер Л., Шафер Р., Цифровая обработка речевых сигналов. – М.: Радио и связь, 2006, – 496 С.
2. Texas Instruments [Электронный ресурс]. – Электронные данные. – Режим доступа : <http://www.ti.com/>

ПРИЛОЖЕНИЕ А  
(Обязательное)  
Схема электрическая структурная

ПРИЛОЖЕНИЕ Б  
(Обязательное)  
Схема электрическая принципиальная

ПРИЛОЖЕНИЕ В  
(Обязательное)  
Блок-схема алгоритма

ПРИЛОЖЕНИЕ Г  
(Обязательное)  
Код программы

void main () {

}

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Обозначение* | | | | | *Наименование* | | *Дополнительные сведения* | | |
|  | | | | | Текстовые документы | |  | | |
|  | | | | |  | |  | | |
| БГУИР КР 1-40 02 02 00X ПЗ | | | | | Пояснительная записка | | 33 с. | | |
|  | | | | |  | |  | | |
|  | | | | | *Графические документы* | |  | | |
|  | | | | |  | |  | | |
| ГУИР XXXX.001 Э1 | | | | | Схема электрическая структурная | | Формат А4 | | |
|  | | | | |  | |  | | |
| ГУИР XXXX.002 Э3 | | | | | Схема электрическая принципиальная | | Формат А4 | | |
|  | | | | |  | |  | | |
| ГУИР XXXX.004 ПД | | | | | Схема алгоритма работы | | Формат А4 | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  | | | | |  | |  | | |
|  |  |  |  |  | *БГУИР КП 1-40 02 02 007 ПЗ* | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| *Изм.* | *Л.* | *№ докум.* | *Подп.* | *Дата* | *Название*  *Ведомость курсового проекта* | *Лит* | | *Лист* | *Листов* |
| *Разраб.* | | *Иванов И.И.* |  |  | *T* | | 33 | 33 |
| *Пров.* | | *Иванов И.И.* |  |  | *Кафедра ЭВС,  гр.X5070X* | | | |
| *Т.контр.* | |  |  |  |
| *Н. Контр.* | |  |  |  |
| *Утв.* | |  |  |  |