

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 ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

БГУИР ДП 1-40 02 02 01 035 ПЗ

Студент	Я. А. Рогова
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 2022

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

« 31 » марта 2022 г.

ЗАДАНИЕ
по дипломному проекту студента

Роговой Яны Александровны
(фамилия, имя, отчество)

1. Тема проекта: Система шумоподавления для речевых сигналов

утверждена приказом по университету от « 31 » марта 2022 г. № 878-с

2. Срок сдачи студентом законченного проекта 15 июня 2022 года

3. Исходные данные к проекту

3.1 Назначение разработки: система предназначена снижения уровня шума в речевом сигнале.

3.2 Технические характеристики:

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

4. Содержание пояснительной записки (перечень подлежащих разработке вопросов)

4.1 Введение. 4.2 Обзор существующих методов шумоподавления для речевых сигналов. 4.3 Анализ технического задания. 4.4 Разработка структуры системы шумоподавления для речевых сигналов. 4.5 Разработка алгоритма работы системы шумоподавления для речевых сигналов. 4.6 Программная реализация системы шумоподавления для речевых сигналов. 4.7 Технико-экономическое обоснование разработки системы шумоподавления для речевых сигналов. 4.8 Анализ результатов тестирования системы шумоподавления для речевых сигналов. 4.9 Заключение. 4.10 Список используемых источников.

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

5.1 Структура системы шумоподавления для речевых сигналов – 1 лист формата А1 (плакат).

5.2 Схема программы системы шумоподавления для речевых сигналов – 1 лист формата А1.

5.3 Схема программы обработки фрейма в системе шумоподавления для речевых сигналов – 1 лист формата А1.

5.4 Результаты разработки системы шумоподавления для речевых сигналов – 1 лист формата А1 (плакат).

5.5 Схема данных для экспериментального исследования эффективности работы системы шумоподавления для речевых сигналов – 1 лист формата А1.

5.6 Экспериментальное исследование эффективности работы системы шумоподавления для речевых сигналов – 1 лист формата А1 (плакат).

6. Содержание задания по технико-экономическому обоснованию

6.1 Расчет затрат на проведение исследования системы шумоподавления для речевых сигналов.

6.2 Оценка уровня качества разработанной системы шумоподавления для речевых сигналов.

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	60	20.04.22	
II этап – п.4.5, п.5.4	20	27.04.22	
III этап – п.4.6 – 4.8., п.5.5, 5.6.	20	12.05.22	
Нормоконтроль		16.05.22 – 20.05.22	
Рабочая комиссия		23.05.22 – 27.05.22	
Рецензирование		30.05.22 – 10.06.22	
Защита		15.06.22 – 30.06.22 (в соответствии с графиком заседаний ГЭК)	

Дата выдачи задания «31» марта 2022 г. Руководитель _____ / М. И. Порхун /

Задание принял к исполнению _____ / Я. А. Рогова /

РЕФЕРАТ

СИСТЕМА ШУМОПОДАВЛЕНИЯ В РЕЧЕВЫХ СИГНАЛАХ :
дипломный проект / Я. А. Рогова. – Минск : БГУИР, 2022, – п.з. – 73 с.,
чертежей (плакатов) – 6 л. формата А1

Основными задачами данного дипломного проекта являются исследование методов шумоподавления и разработка системы шумоподавления для речевых сигналов.

Важной особенностью разработанной системы является использование метода субполосного спектрального вычитания для выполнения шумоподавления. Недостатком системы является возникновение музыкальных шумов, которые снижают качество обработанного сигнала.

Результаты тестирования системы, а также анализ существующих методов, позволили заключить, что данная система обладает достаточным уровнем конкурентоспособности в своей области применения.

Ключевые слова: шумоподавление, субполосное спектральное вычитание, цифровая обработка сигналов.

СОДЕРЖАНИЕ

Введение.....	8
1 Обзор существующих методов шумоподавления для речевых сигналов....	10
1.1 Существующие методы шумоподавления в речевых сигналах	10
1.2 Методы, основанные на спектральном вычитании	10
1.2.1 Базовый алгоритм спектрального вычитания	10
1.2.2 Спектральное сверхвычитание	13
1.2.3 Субполосное спектральное вычитание.....	14
1.3 Адаптивные фильтры.....	15
1.3.1 Адаптивный алгоритм LMS.....	16
1.3.2 Адаптивный алгоритм RLS.....	16
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 системы шумоподавления для речевых сигналов ...	19
3.1 Слуховое восприятие	19
3.1.1 Влияние шума на восприятие речи	19
3.1.2 Влияние шума на качество слуха.....	20
3.2 Классификация шумов.....	21
3.2.1 Классификация по характеру спектра	21
3.2.2 Классификация по временным характеристикам.....	22
3.2.3 Цветные шумы	22
3.3 Принцип работы системы шумоподавления для речевых сигналов.....	24
4 Разработка алгоритма работы системы шумоподавления для речевых сигналов.....	25
4.1 Основные этапы работы системы	25
4.2 Алгоритм анализа сигнала.....	25
4.3 Алгоритм оценки начального уровня мощности шума.....	27
4.4 Алгоритм обработки фрейма.....	27
4.4.1 Определение наличия речи во фрейме	29
4.4.2 Вычисление ОСШ.....	30
4.4.3 Субполосное спектральное вычитание.....	31
4.4.4 Новая оценка мощности шума	31
4.5 Синтез сигнала.....	32

5 Программная реализация системы шумоподавления для речевых сигналов.....	33
5.1 Описание пакета MATLAB	33
5.2 Программная реализация системы шумоподавления для речевых сигналов	34
5.3 Описание вспомогательных функций пакета MATLAB	35
5.3.1 Функция <i>audioread</i>	35
5.3.2 Функция <i>hamming</i>	36
5.3.3 Функция <i>fft</i>	36
5.3.4 Функция <i>flipud</i>	37
5.3.5 Функция <i>conj</i>	37
5.3.6 Функция <i>ifft</i>	37
5.3.7 Функция <i>audiowrite</i>	38
5.4 Описание пользовательских функций.....	39
5.4.1 Функция <i>signal_to_frames</i>	39
5.4.2 Функция <i>start_noise</i>	40
5.4.3 Функция <i>processing_frames</i>	40
5.4.4 Функция <i>voice_activity_detection</i>	40
5.4.5 Функция <i>processing_frame</i>	41
5.4.6 Функция <i>snr</i>	41
5.4.7 Функция <i>spectral_subtraction</i>	42
5.4.8 Функция <i>band_borders</i>	42
5.4.9 Функция <i>new_noise</i>	42
5.4.10 Функция <i>synthesis_signal</i>	42
6 Техничко-экономическое обоснование разработки системы шумоподавления для речевых сигналов	43
6.1 Расчет затрат на проведение исследования системы шумоподавления для речевых сигналов	43
6.1.1 Разработка плана проведения научно-исследовательской работы	43
6.1.2 Расчет затрат на топливно-энергетические ресурсы.....	46
6.1.3 Расчет затрат на спецоборудование, приобретаемое для научно-исследовательской работы.....	47
6.1.4 Расчет основной заработной платы научно-технического персонала	47
6.1.5 Расчет дополнительной заработной платы научно-технического персонала	48
6.1.6 Расчет отчислений на социальные нужды	48

6.1.7 Расчет накладных затрат	49
6.1.8 Расчет полной себестоимости	49
6.1.9 Расчет плановой прибыли	49
6.1.10 Расчет сметной стоимости научно-технической продукции	49
6.2 Оценка уровня качества разработанной системы шумоподавления для речевых сигналов	50
7 Анализ результатов тестирования системы шумоподавления для речевых сигналов.....	52
7.1 Тестирование разработанной системы шумоподавления	52
7.2 Экспериментальное исследование системы шумоподавления для речевых сигналов	54
7.2.1 Описание проводимого эксперимента.....	54
7.2.2 Подготовка данных для проведения эксперимента	55
7.2.3 Результаты проведенного эксперимента	57
7.3 Анализ результатов тестирования	59
Заключение	60
Список используемых источников.....	61
Приложение А (обязательное) Отчет о проверке на заимствования	63
Приложение Б (обязательное) Листинг	64
Б.1 Листинг программы шумоподавления для речевых сигналов	64
Б.2 Листинг программы для тестирования системы шумоподавления	71

ВВЕДЕНИЕ

Речь является одной из наиболее часто используемых форм обмена информацией между людьми. В настоящее время существует множество технологий, обеспечивающих передачу речевых сигналов на расстояние.

Зачастую во время записи речевые сигналы подвергаются воздействию тех или иных шумов, возникающих из-за несовершенства записывающей аппаратуры, собственных шумов микрофонов и шума окружающей среды [1]. В результате появления аддитивного шума в записанном речевом сигнале ухудшается его разборчивость при воспроизведении. В связи с этим создание методов обработки речевых сигналов, позволяющих понизить шум для повышения качества звучания, является актуальной задачей.

В настоящее время существует множество методов цифровой обработки сигналов, очищающих речевые сигналы от шумов: адаптивные методы шумоподавления; методы, использующие спектральные характеристики шума; методы, основанные на использовании моделей нейронных сетей; методы, основанные на моделях восприятия речи человеком.

Целью данного дипломного проекта является исследование современных методов шумоподавления, применяемых для повышения качества речевых сигналов, а также проектирование системы шумоподавления с использованием одного из существующих методов.

Проектирование системы шумоподавления в речевых сигналах производилось в несколько этапов, отраженных в структуре пояснительной записки данного дипломного проекта.

Первоначально необходимо ознакомиться с существующими методами шумоподавления для речевых сигналов. Для выполнения этого этапа был осуществлен поиск теоретического материала и его последующий анализ. В первом разделе пояснительной записки приведены результаты анализа наиболее часто используемых методов шумоподавления для речевых сигналов.

Следующим этапом выполнения дипломного проекта является анализ технического задания дипломного проекта, в результате которого был выбран метод, составляющий основу алгоритма работы системы. Результаты данного этапа приведены во втором разделе пояснительной записки.

На основе выбранного метода шумоподавления было проведено проектирование системы. Этот этап описан в третьем и четвертом разделах

пояснительной записки, которые содержат в себе описание структуры и алгоритма работы системы соответственно.

Для проверки работоспособности алгоритма работы системы была осуществлена его программная реализация. Особенности этого этапа приведены в пятом разделе.

Следующим этапом является технико-экономическое обоснование разработки системы шумоподавления для речевых сигналов, результаты выполнения которого приведены в шестом разделе.

На следующем этапе осуществляется тестирование разработанной системы. Тестирование производилось в среде MATLAB на различных типах сигналов с различным уровнем шума. Также для тестирования разработанной системы был проведен эксперимент, определяющий изменение разборчивости речи для сигналов с низким значением отношения сигнал-шум до и после шумоподавления. Результаты тестирования системы отражены в седьмом разделе.

В завершение был проведен анализ результатов проектирования системы, результаты которого представлены в заключении пояснительной записки.

Отчет о проверке на заимствования представлен в приложении А.

1 ОБЗОР СУЩЕСТВУЮЩИХ МЕТОДОВ ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

1.1 Существующие методы шумоподавления в речевых сигналах

На данный момент существует множество методов для решения задачи шумоподавления в речевых сигналах. Главными направлениями в решении задачи шумоподавления в речевых сигналах являются алгоритмы, основанные на использовании спектральных характеристик шума, и адаптивная фильтрация.

Спектральные характеристики шума чаще всего используется в алгоритмах спектрального вычитания, в которых входной сигнал представляется как сумма речевого сигнала и шума. На участках, где отсутствует речь, оценивается спектральная плотность мощности шума. Восстановленный сигнал формируется посредством вычитания спектральной плотности мощности шума из спектральной плотности мощности речевого сигнала [2].

Другой подход к решению задачи шумоподавления в речевых сигналах – это использование адаптивных фильтров. Адаптивный фильтр состоит из цифрового фильтра с изменяемыми коэффициентами и адаптивного алгоритма, который используется для изменения коэффициентов фильтра. Коэффициенты цифрового фильтра регулируются некоторым адаптивным алгоритмом, в котором сигнал ошибки минимизируется согласно выбранному критерию [3].

1.2 Методы, основанные на спектральном вычитании

Алгоритм используются в случаях, когда входящий зашумленный сигнал $x(n)$ можно представить как сумму речевого сигнала $s(n)$ и аддитивного шума $d(n)$:

$$x(n) = s(n) + d(n). \quad (1.1)$$

1.2.1 Базовый алгоритм спектрального вычитания

Базовый алгоритм спектрального вычитания был разработан для очистки сигналов от белого шума. Входящий зашумленный сигнал разбивается на перекрывающиеся фреймы. Предполагается, что k первых фреймов содержат только шум. Начальное значение спектра мощности шума

рассчитывается как среднее значение для каждого отсчета в первых k фреймах.

Последующая обработка осуществляется последовательно для каждого фрейма. Фрейм умножается на оконную функцию, после чего применяется преобразование Фурье для перехода в частотную область. Далее осуществляется переход к полярным координатам, чтобы фаза выходного сигнала была такой же, как у входного. Для определения наличия речевого сигнала в текущем фрейме высчитывается значение сегментного отношения сигнал/шум (ОСШ).

После выполнения шагов, перечисленных выше, выполняется спектральное вычитание, формула которого в общем виде выглядит следующим образом:

$$|\hat{S}(\omega)|^p = |X(\omega)|^p - |\hat{D}(\omega)|^p, \quad (1.2)$$

где $|\hat{S}(\omega)|^p$, $|X(\omega)|^p$ – амплитудные спектры очищенного и зашумленного сигнала соответственно; $|\hat{D}(\omega)|^p$ – оценка амплитудного спектра шума, определенная как усредненное значение его амплитуд во время пауз в речи; p – показатель степени.

На практике зачастую используют показатель $p = 2$:

$$|\hat{S}(\omega)|^2 = |X(\omega)|^2 - |\hat{D}(\omega)|^2. \quad (1.3)$$

В правой части формулы (1.3) может получиться значение ниже нуля при неточной оценке амплитудного спектра шума. Однако значение амплитуды не может быть отрицательным числом, поэтому необходимо ограничить отрицательные компоненты:

$$|\hat{S}(\omega)|^2 = \begin{cases} |X(\omega)|^2 - |\hat{D}(\omega)|^2, & |X(\omega)|^2 - |\hat{D}(\omega)|^2 > 0 \\ 0, & \text{иначе.} \end{cases} \quad (1.4)$$

После выполнения спектрального вычитания необходимо определить, присутствует ли во фрейме речь. Для этого ранее полученное значение сегментного ОСШ сравнивается с его пороговым значением, определенным экспериментально. Если полученное значение больше порогового, значит во фрейме присутствует речь, иначе фрейм содержит только шум. Если

определено, что во фрейме отсутствует речь, то происходит новая оценка спектра мощности шума методом экспоненциального усреднения:

$$|\hat{D}(\omega)|^2 = \gamma |\hat{D}(\omega)|^2 + (1 - \gamma) |X(\omega)|^2, \quad (1.5)$$

где γ – коэффициент усреднения ($0,9 < \gamma < 1$).

Далее осуществляется переход из полярных координат к декартовым и выполняется обратное преобразование Фурье, в результате которого получается итоговый выходной сигнал.

Структура описанного алгоритма приведена на рисунке 1.1:

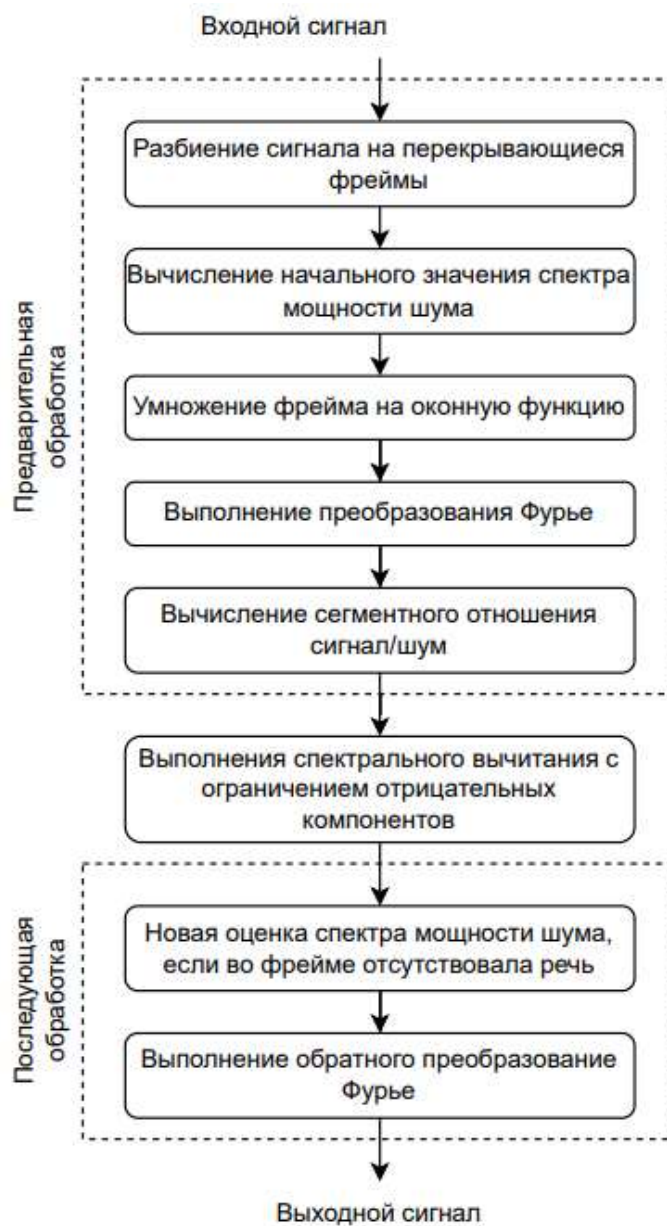


Рисунок 1.1 – Структура базового алгоритма спектрального вычитания

В случае вычисления спектра шума на коротком интервале времени в нем присутствуют локальные минимумы и максимумы. Когда из спектра зашумленного сигнала вычитается оценка среднего значения шума, в окрестности локального минимума спектр принимает нулевое значение, а максимумы лишь уменьшают амплитуду. В результате в спектре шума остаются локальные максимумы, которые продолжают восприниматься как шум. При этом наиболее широкие максимумы воспринимаются как широкополосный шум, а узкие – как изменяющиеся тона, которые называют музыкальным шумом.

Плюсами данного алгоритма являются его простота и вычислительная скорость. К минусам можно отнести неполную очистку речевого сигнала от шума и появление музыкального шума в очищенном сигнале.

1.2.2 Спектральное сверхвычитание

Для уменьшения музыкального шума, возникающего в базовом алгоритме спектрального вычитания используется данная модификация алгоритма. В ней вводятся два дополнительных параметра: коэффициент сверхвычитания α и коэффициент спектрального минимума шума β [4]. Амплитудный спектр очищенного сигнала вычисляется следующим образом:

$$|\hat{S}(\omega)|^2 = \begin{cases} |X(\omega)|^2 - \alpha |\hat{D}(\omega)|^2, & |X(\omega)|^2 > (\alpha + \beta) |\hat{D}(\omega)|^2 \\ \beta |\hat{D}(\omega)|^2 & \text{иначе,} \end{cases} \quad (1.6)$$

при этом $\alpha \geq 1$, $0 < \beta \ll 1$.

Коэффициент сверхвычитания α позволяет увеличить вычитаемый амплитудный спектр шума, что приводит к лучшему очищению сигнала от шума, в то время как коэффициент спектрального минимума шума β уменьшает воспринимаемый музыкальный шум. При больших значениях β остаточный шум достаточно ощутим, а при низких возникает высокий уровень музыкального шума. Таким образом, использование коэффициентов α и β позволяет добиться компромисса между количеством остаточного шума и уровнем воспринимаемого музыкального шума.

Преимуществом данного подхода является лучшая очистка сигнала от шума, чем при использовании базового алгоритма, и лучшее подавление музыкального шума. Однако после применения присутствует остаточный шум, влияющий на качество сигнала. Так же недостатком алгоритма является допущение, что шум одинаково влияет на весь речевой спектр, из-за чего

используется один коэффициент α для всего спектра речи, в результате чего речь искажается.

1.2.3 Субполосное спектральное вычитание

Реальный шум влияет на речевой сигнал по-разному на разных участках спектра. Сегментное ОСШ в низкочастотных диапазонах значительно выше, чем в высокочастотных. Идея данной модификации состоит в том, чтобы использовать большие значения коэффициента сверхвычитания α на частотах с низким значением ОСШ и малые – на частотах с высоким [4].

В этой модификации алгоритма спектр сигнала делится на k неперекрывающихся полос и в каждой полосе происходит независимое спектральное вычитание [5]:

$$|\hat{S}_i(\omega)|^2 = |X_i(\omega)|^2 - \alpha_i \delta_i |\hat{D}_i(\omega)|^2 \quad b_i \leq \omega \leq e_i, \quad (1.7)$$

где i – номер полосы; α_i – коэффициент сверхвычитания, зависящий от значения ОСШ в i -й полосе; δ_i – дополнительный коэффициент вычитания в i -й полосе, который устанавливается отдельно для каждой полосы; b_i и e_i – минимальное и максимальное значение частоты i -й полосы соответственно.

В случае получения отрицательного результата в правой части выражения (1.7) необходимо провести ограничение отрицательных компонентов:

$$|\hat{S}_i(\omega)|^2 = \begin{cases} |\hat{S}_i(\omega)|^2, & |\hat{S}_i(\omega)|^2 > 0 \\ \beta |X_i(\omega)|^2 & \text{иначе.} \end{cases} \quad (1.8)$$

В статье [5] говорится, что для получения хорошего качества речи после обработки данным методом достаточно деления спектра на три полосы.

На рисунке 1.2 приведено описание последовательности выполнения модификации алгоритма, при этом предварительная и последующая обработки совпадают с базовым алгоритмом спектрального вычитания.

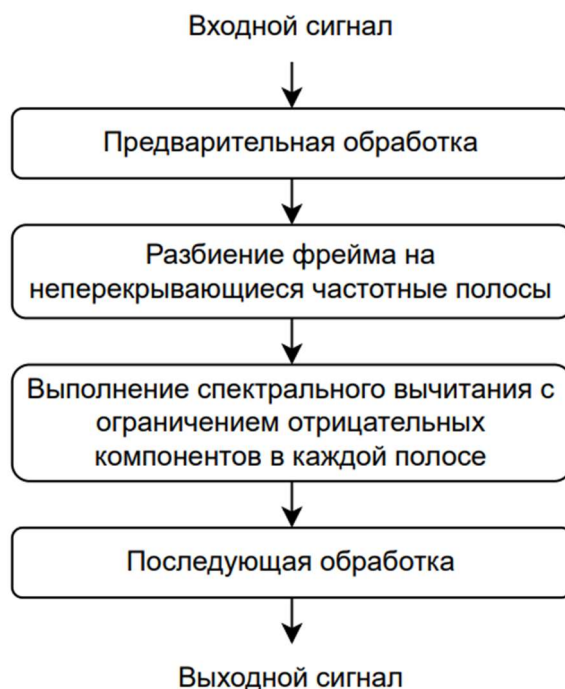


Рисунок 1.2 – Структура алгоритма субполосного спектрального вычитания

Преимуществом данного метода над базовым алгоритмом спектрального вычитания является лучшее качество очистки речи от различного рода шумов, а также снижение воспринимаемого музыкального шума, при этом дополнительная вычислительная сложность алгоритма минимальна.

1.3 Адаптивные фильтры

На адаптивный фильтр поступает два сигнала: зашумленный речевой сигнал $d(k)$, в котором речь и шум не коррелируют, и мера зашумленного сигнала $x(k)$, коррелирующая с шумом из первого сигнала. Цифровой фильтр обрабатывает шумовой сигнал для получения оценки шума $\hat{y}(k)$, после чего очищенный сигнал $\hat{s}(k)$ оценивается как разность зашумленного сигнала и полученной оценки шума. Выходной сигнал является одновременно результирующим очищенным сигналом и сигналом ошибки $e(k)$, применяем для корректировки коэффициентов цифрового фильтра. Зачастую критерием минимизации сигнала ошибки e_k выступает среднее квадратическое отклонение (СКО).

На рисунке 1.3 изображена схема работы адаптивного фильтра:

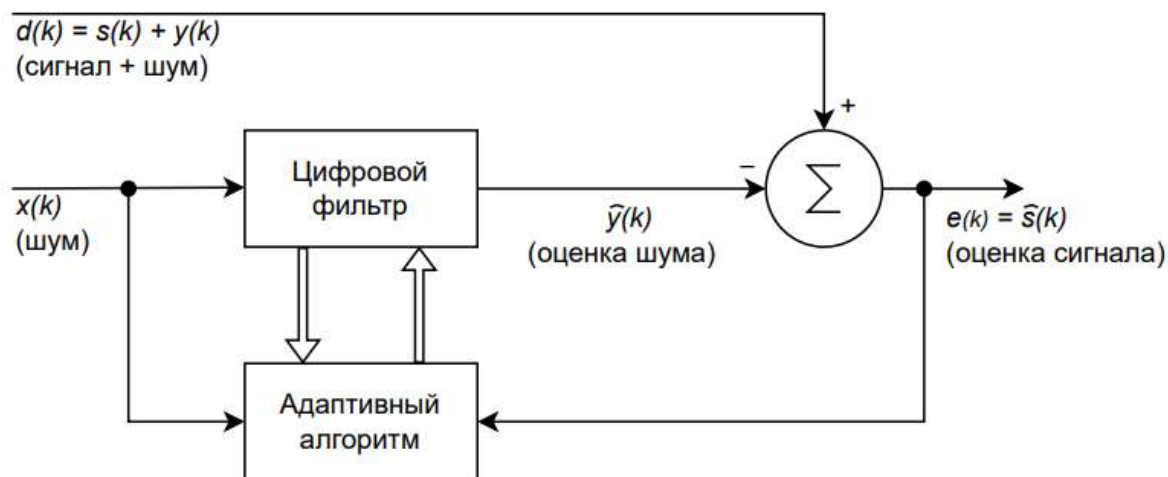


Рисунок 1.3 – Структурная схема адаптивного фильтра

Наиболее известными адаптивными алгоритмами являются алгоритм LMS (метод наименьшего квадрата) и алгоритм RLS (рекурсивный метод наименьших квадратов).

1.3.1 Адаптивный алгоритм LMS

Главным достоинством алгоритма LMS является низкая вычислительная сложность алгоритма – для подстройки коэффициентов фильтра на каждом шаге необходимо выполнить $N + 1$ пар «умножение-сложение». Минусами являются медленная сходимость и повышенная по сравнению с минимально достижимым значением дисперсия ошибки в установившемся режиме [6].

1.3.2 Адаптивный алгоритм RLS

Алгоритм RLS требует значительно большего числа операций по сравнению с алгоритмом LMS, что является его минусом. Его вычислительная сложность вырастает квадратично с увеличением порядка фильтра. В то же время RLS-алгоритм сходится значительно быстрее LMS-алгоритма.

2 АНАЛИЗ ТЕХНИЧЕСКОГО ЗАДАНИЯ

2.1 Анализ требований к алгоритму

Основными требованиями к алгоритму работы системы шумоподавления для речевых сигналов являются:

- независимость от типа аддитивного шума;
- высокая скорость обработки сигнала от шума.

Для удовлетворения данных требований были определены способы их решения.

Независимость от типа аддитивного шума, присутствующего в начальном сигнале, позволяет использовать систему для очистки сигналов от различных типов шумов, а не от одного конкретного. Указанное требование удовлетворяется путем выбора алгоритма шумоподавления, нацеленного на подавление различных типов шумов.

Высокая скорость обработки речевого сигнала от шума достигается путем выбора менее ресурсоемких алгоритмов цифровой обработки сигналов.

2.2 Анализ требований к программной реализации

В техническом задании по данному дипломному проекту указаны следующие требования к программной реализации системы шумоподавления для речевых сигналов:

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

Перечисленные требования были учтены при выполнении программной реализации системы.

2.3 Выбор и обоснование метода решения задачи

Задачей разрабатываемой системы является снижение уровня шума в записанном речевом сигнале.

Исходя из информации, приведенной в разделе 1, в настоящее время широко используются следующие методы шумоподавления для речевых сигналов:

- спектральное вычитание;
- адаптивная фильтрация.

Перед осуществлением выбора основного метода было необходимо произвести анализ указанных методов.

Методы спектрального вычитания обладают низкой вычислительной сложностью алгоритма, однако в результате их применения возникают музыкальные шумы.

При использовании адаптивной фильтрации происходит лучшая очистка сигнала от шума посредством того, что коэффициенты фильтра пересчитываются. Однако при использовании данного типа фильтрации, алгоритм обладает либо низкой вычислительной сложностью и медленной сходимостью, либо быстрой сходимостью и высокой вычислительной сложностью [6].

Основываясь на информации, изложенной выше, а также на требованиях, предъявленных к алгоритму, был выбран метод шумоподавления, основанный на методе субполосного спектрального вычитания. Выбранный метод шумоподавления позволяет добиться быстрой обработки речевого сигнала при обеспечении оптимального потребления вычислительных ресурсов. Данный метод производит обработку сигнала фрейм за фреймом, что позволяет, при необходимости, реализовать систему реального времени. При разработке системы использовался детектор речи, оценивающий наличие речи по трем параметрам: нормированной краткосрочной энергии сигнала, доминантной частоте и тональному коэффициенту. Новая оценка уровня мощности шума производится во фреймах, в которых отсутствует речь, и зависит от уровня, определенного при предыдущей оценке.

3 РАЗРАБОТКА СТРУКТУРЫ СИСТЕМЫ ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

3.1 Слуховое восприятие

Диапазон звуков по частоте, воспринимаемых человеком, в норме составляет от 20 до 20000 Гц. Порог слухового ощущения, абсолютный порог слуха – минимальный уровень звука, который воспринимает человек. Порог слуха определяет слуховую чувствительность. Наибольшая чувствительность слуха человека находится в диапазоне от 1000 до 4000 Гц, так как в этом диапазоне находится значительная часть информации о речевых сигналах. Так пороги слуха на частоте 200 Гц выше на 35 дБ, а на 100 Гц – на 60 дБ, чем пороги слуха на частоте 1000 Гц. Благодаря этому человек не слышит низкочастотных шумов своего тела, вызываемых движением крови по сосудам, сокращением мышц, дыханием.

Слуховая система, как и другие сенсорные системы, обладает свойством адаптации. Адаптация проявляется во временном повышении порогов слуха, возникающем в результате предшествующей звуковой стимуляции. Это защищает слуховую систему, и прежде всего волосковые клетки, от повреждения длительными звуками. Однако регулярное длительное воздействие интенсивных звуков (например, шумные производства или постоянное слушание громкой музыки) приводит к необратимому повреждению волосковых клеток и других структур улитки, вызывает атрофические изменения в слуховом нерве [7].

3.1.1 Влияние шума на восприятие речи

С точки зрения восприятия в конкретной ситуации звуковые сигналы можно разделить на полезный сигнал и мешающие шумы.

Полезный сигнал – это звуки, которые в данный момент представляют интерес для человека. По отношению к ним остальные звуки являются мешающими, при этом они маскируют полезный сигнал. Если человек слушает, что ему говорит другой человек, то шум улицы или речь 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 окружающие неречевые шумы, так как речь имеет одинаковый частотный состав.

Для оценки влияния шума на слуховое восприятие используется ОСШ. Критический уровень ОСШ возникает тогда, когда человек узнает 50% услышанных слов, произносимых изолированно, что позволяет ему в слитной речи понять большую часть сказанного за счет контекста предложения и ситуации. Критический уровень ОСШ у нормально слышащего взрослого составляет минус 6 дБ. У человека с сенсоневральной тугоухостью этот уровень составляет 15 – 20 дБ. У человека с нарушенным слухом снижение критического уровня речь/шум только на 1 дБ снижает разборчивость речи на 20%. При снижении этого показателя на 3 дБ происходит полная потеря разборчивости речи.

Дети даже с нормальным слухом хуже воспринимают речь в шуме, чем взрослые. Оптимальное ОСШ, которое позволяет ребенку понимать речь и осваивать новые слова, составляет:

- для нормально слышащих детей 20 дБ;
- для слабослышащих детей 30 дБ [7].

3.1.2 Влияние шума на качество слуха

В долгой перспективе влияние шума на организм человека имеет негативные последствия. Существует такое профессиональное заболевание, как профессиональная тугоухость. Под этим термином подразумевают постепенное снижение слуха, обусловленное длительным (многолетним) воздействием производственного шума. К шумоопасным производствам относятся добывающая, дерево-, металло-, камнеобрабатывающая промышленность, ткацкое производство, машино-, авиа- и судостроение, транспорт и др. К числу шумоопасных профессий также относится работа инструкторов фитнес-клубов. Показано, что сочетанное воздействие громкой музыки с физическими упражнениями усиливают повреждающее воздействие громких звуков на слух.

В результате постоянной микротравматизации шумом формируются сосудистые и дистрофические изменения в улитке и спиральном ганглии. Слух снижается сначала на высоких частотах (4000 – 8000 Гц), которые в основном составляют спектр частот промышленного шума. Постепенно слух снижается в средне- и низкочастотном диапазонах, больной начинает плохо различать слова, затем плохо слышать разговорную речь, начинает предъявлять жалобы на снижение слуха.

Нормы допустимых воздействий шума представлены в таблице 3.1.

Таблица 3.1 – Допустимые воздействия шума

Длительность в течение дня	Уровень шума, дБ
8 часов	90
4 часа	95
2 часа	100
15 минут	105
4 минуты	110
2 минуты	114
30 секунд	120

Шум влияет не только на слух. Он является стрессовым фактором, воздействует на весь организм в целом:

- угнетает центральную нервную систему;
- изменяет скорость дыхания и пульса;
- приводит к нарушению обмена веществ;
- повышает риск возникновения сердечно-сосудистых заболеваний, гипертонической болезни и др.;
- снижает работоспособность и производительность труда.

Действие шума во многих случаях сочетается с воздействием вибрации, пыли, неблагоприятных факторов микро- и макроклимата, с вынужденным неудобным рабочим положением тела, физическим перенапряжением, повышенным вниманием. Эти факторы ускоряют развитие патологии и обуславливают ее разнообразные проявления.

3.2 Классификация шумов

Шумы можно классифицировать по нескольким параметрам. Далее приведена классификация характеру спектра, временным характеристикам, а также представлены отдельные категории шумов.

3.2.1 Классификация по характеру спектра

По характеру спектра выделяют следующие типы шумов:

- 1 Широкополосный шум.
- 2 Тональный шум, в спектре которого имеются выраженные тона. Тон считается выраженным, если одна из полос частот превышает остальные не менее, чем на 10 дБ [8]. Таким образом, музыкальный шум относится к тональным шумам.

3.2.2 Классификация по временным характеристикам

1 Стационарный шум – шум, который характеризуется постоянством средних параметров, таких как мощность, спектральная плотность (распределение интенсивности по спектру) и автокорреляционная функция. Классической моделью стационарного шума является белый шум [9].

2 Нестационарный шум – шум, меняющийся через короткие промежутки времени.

Алгоритм шумоподавления, основанный на методе субполосного спектрального вычитания, подходит для решения задачи очистки сигнала от стационарных шумов. В свою очередь, для очистки сигнала от нестационарного шума лучшего результата достигли бы адаптивные алгоритмы. Таким образом, разрабатываемая система нацелена на очистку сигнала от стационарных шумов.

3.2.3 Цветные шумы

Цветовые соответствия различных типов шумов определяются с помощью гистограмм спектральной плотности.

Белый шум – это сигнал с равномерной спектральной плотностью на всех частотах. Белый шум является стационарным случайным процессом. Белый шум, имеющий одинаковую спектральную мощность на всех частотах, не встречается ввиду того, что такой сигнал имел бы бесконечную мощность, однако под категорию белых шумов попадают любые шумы, спектральная плотность которых одинакова (или слабо отличается) в рассматриваемом диапазоне частот [10]. Спектр белого шума приведен на рисунке 3.1

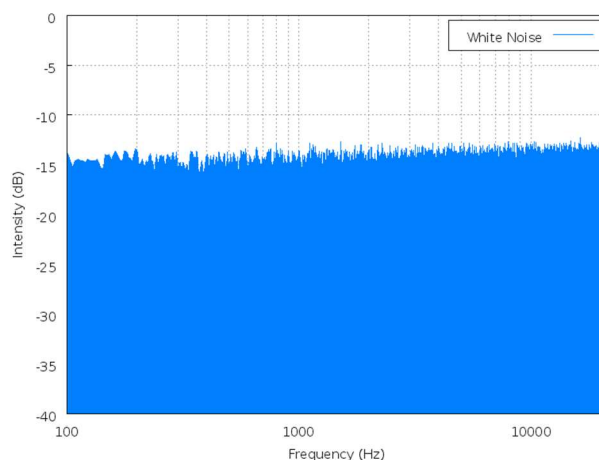


Рисунок 3.1 – Спектр белого шума

Розовый шум – шум, спектральная плотность мощности которого обратно пропорциональна частоте. В логарифмическом шкале такой шум является равномерно убывающим. Розовый шум обнаруживается, в сердечных ритмах, в графиках электрической активности мозга, а также практически в любых электронных и механических устройствах. Спектр розового шума приведен на рисунке 3.2.

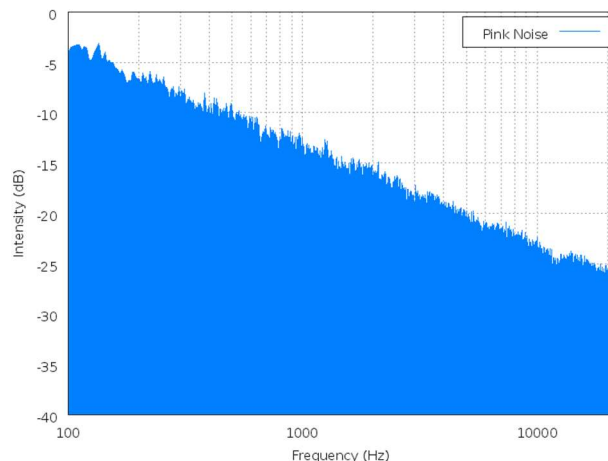


Рисунок 3.2 – Спектр розового шума

Серый шум – шумовой сигнал, имеющий одинаковую субъективную громкость для человека на всём диапазоне воспринимаемых частот. Спектр серого шума значительно ниже на средних частотах, чем на низких и высоких, однако человеческий слух субъективно воспринимает серый шум как равномерный по спектральной плотности. Спектр серого шума приведен на рисунке 3.3.

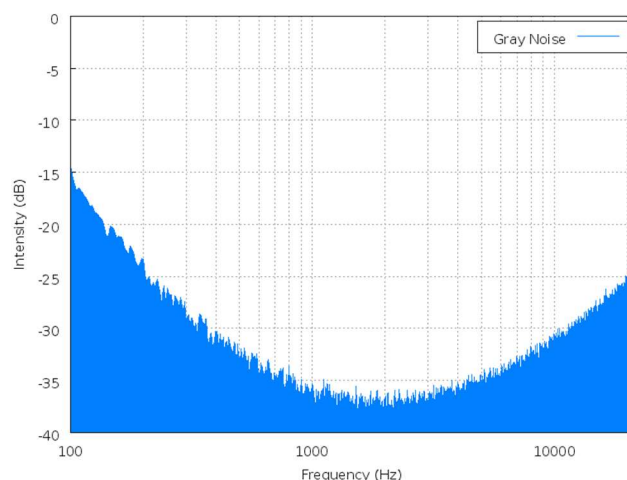


Рисунок 3.3 – Спектр серого шума

3.3 Принцип работы системы шумоподавления для речевых сигналов

Принцип работы системы шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания заключается в следующем: входной сигнал $x(n)$ разбивается на фреймы с перекрытием 50%, каждый фрейм умножается на окно Хэмминга. Осуществляется оценка начального уровня шума с допущением, что первые k фреймов не содержат речи. Далее осуществляется обработка фрейм за фреймом. Для каждого фрейма выполняется субполосное спектральное вычитание, далее, в случае отсутствия речи в текущем фрейме, производится новая оценка уровня шума. После обработки последнего фрейма синтезируется сигнал с помощью метода перекрытия с суммированием.

Структура системы шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания представлена на рисунке 3.4.

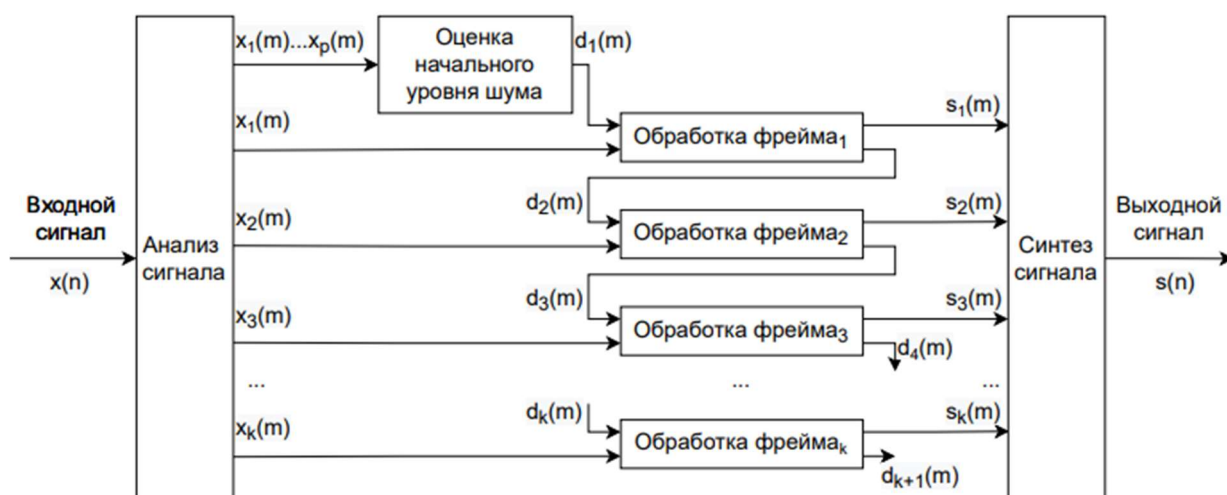


Рисунок 3.4 – Структура системы шумоподавления для речевых сигналов

Более детальная структура системы шумоподавления для речевых сигналов представлена на плакате ГУИР.460818.001 ПЛ.

4 РАЗРАБОТКА АЛГОРИТМА РАБОТЫ СИСТЕМЫ ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

4.1 Основные этапы работы системы

Обобщённый алгоритм системы шумоподавления для речевых сигналов с использованием метода субполосного спектрального вычитания включает в себя следующие этапы:

- 1 Считывание входного сигнала.
- 2 Анализ сигнала.
- 3 Оценка начального значения мощности шума.
- 4 Обработка сигнала фрейм за фреймом:
 - 4.1 Определение наличия речи во фрейме;
 - 4.2 ДПФ;
 - 4.3 Вычисление ОСШ;
 - 4.4 Выполнение субполосного спектрального вычитания;
 - 4.5 ОДПФ;
 - 4.6 Новая оценка мощности шума, в случае если во фрейме отсутствует речь;
 - 4.7 Передача нового значения мощности шума в следующий фрейм.
- 5 Синтез выходного сигнала.
- 6 Запись выходного сигнала в файл.

Схема программы системы шумоподавления для речевых сигналов представлена на чертеже ГУИР.460818.001 ПД.

4.2 Алгоритм анализа сигнала

Под анализом сигнала в разрабатываемом алгоритме подразумевается разбиение сигнала на фреймы с перекрытием 50 % и умножение каждого фрейма на оконную функцию Хэмминга.

Длина одного фрейма равна $frame_size$, а

$$half_size = 0.5 \cdot frame_size, \quad (4.1)$$

тогда для формирования буфера фреймов с перекрытием 50% необходимо двигаться по сигналу с шагом $half_size$.

На рисунке 4.1 показано как формируются фреймы при использовании оконной функции и перекрытия 50%.

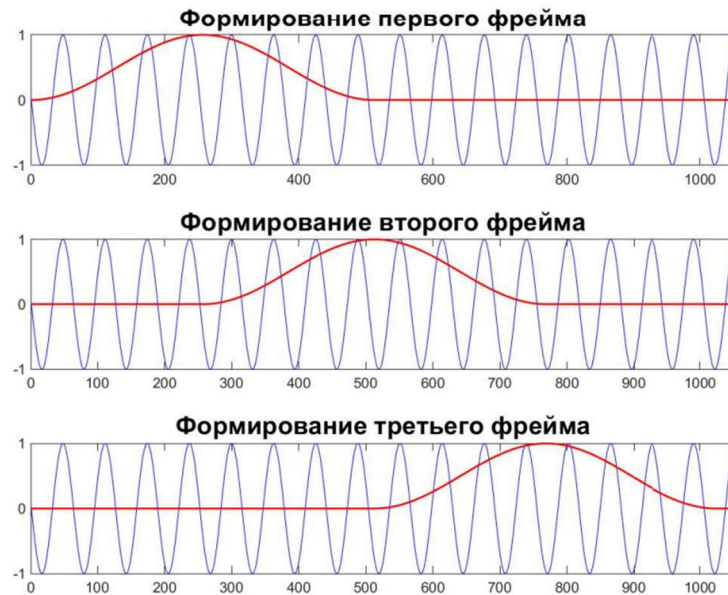


Рисунок 4.1 – Формирование фреймов с перекрытием 50 %

После того, как очередной фрейм сформирован, он умножается на оконную функцию Хэмминга длиной $frame_size$ [12]:

$$w(k) = 0.54 - 0.46 \cos\left(\frac{2\pi k}{N}\right), 0 \leq k \leq frame_size - 1. \quad (4.2)$$

На рисунке 4.2 показаны представления окна Хэмминга: а – во временной области; б – в частотной области.

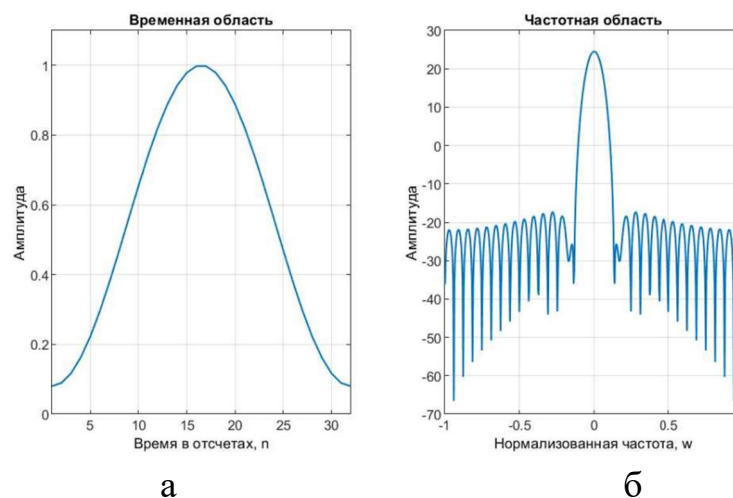


Рисунок 4.2 – Окно Хэмминга

4.3 Алгоритм оценки начального уровня мощности шума

Предполагается, что на первых k фреймах отсутствует речь. Исходя из этого предположения начальная оценка мощности шума вычисляется как среднее значение мощности на первых k фреймах для каждого отсчета. Оценка производится как для фрейма в целом, так и для каждой частотной полосы в отдельности.

Схема алгоритма представлена на рисунке 4.3.

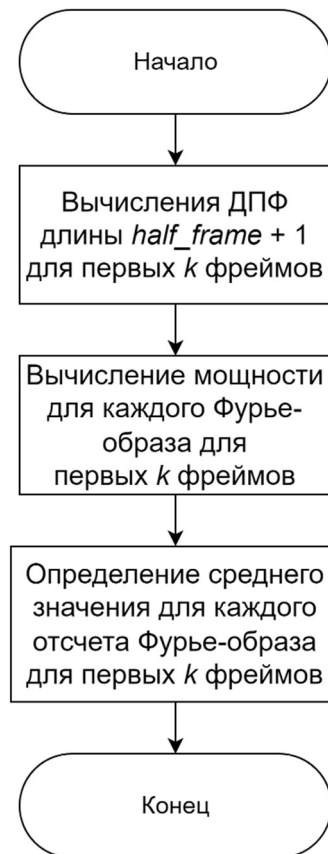


Рисунок 4.3 – Схема алгоритма начальной оценки шума

4.4 Алгоритм обработки фрейма

В разрабатываемой системе используется обработка фрейм за фреймом. При обработке первого фрейма используется начальная оценка уровня мощности шума. Далее оценка производится вновь, если было определено, что во фрейме отсутствует речь, и при обработке следующего фрейма используется уже обновленная оценка мощности шума.

Алгоритм обработки фрейма представлен на рисунке 4.4.

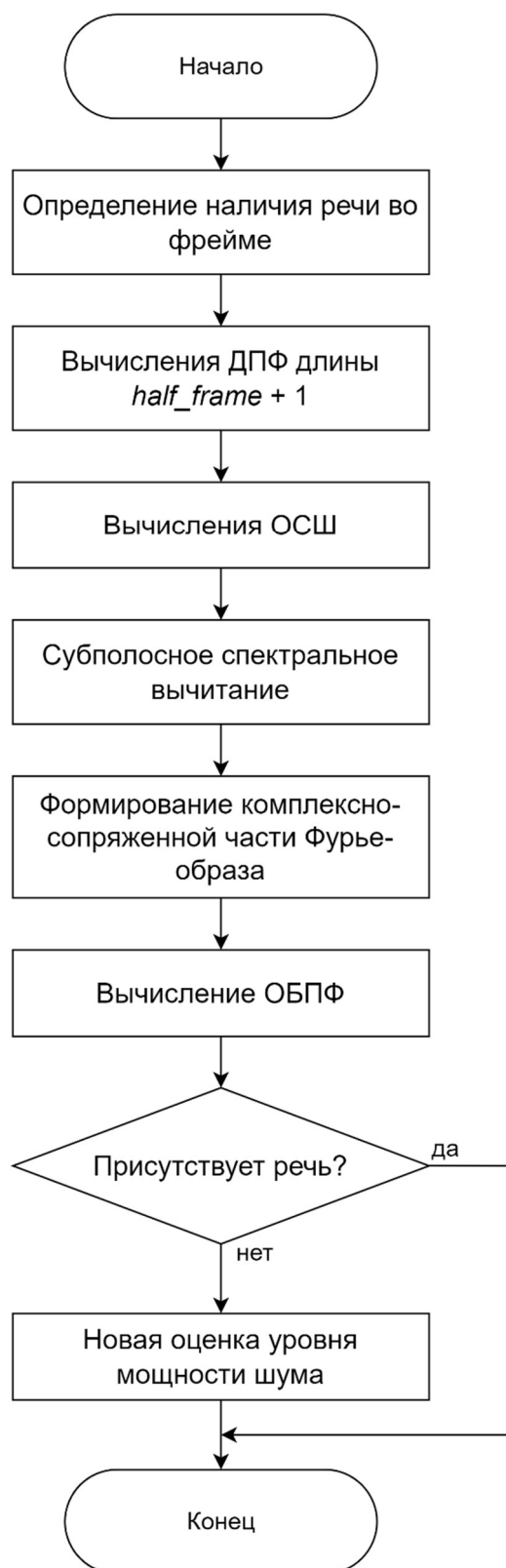


Рисунок 4.4 – Схема алгоритма обработки фрейма

Более детальная схема программы обработки фрейма системы шумоподавления для речевых сигналов представлена в виде чертежа ГУИР.460818.002 ПД к данному проекту.

4.4.1 Определение наличия речи во фрейме

Определение наличия речи во фрейме осуществляется при помощи алгоритма, описанного в статье [11]. Оценка производится по трем параметрам: нормированной краткосрочной энергии сигнала, доминантной частоте и тональному коэффициенту.

Нормированная краткосрочная энергия E вычисляется по формуле:

$$E = \frac{1}{N} \sum_{k=1}^N x_n, \quad (4.3)$$

где k – номер отсчета фрейма; N – длина фрейма; x_n – значение k -го отсчета фрейма.

Доминантная, то есть наиболее часто встречающаяся, частота F определяется по формуле:

$$F = \max(|X(\omega)|), \quad (4.4)$$

где $|X(\omega)|$ – амплитуда Фурье-образа фрейма.

Для вычисления тонального коэффициента SFM используется формула:

$$\text{SFM} = 10 \lg \frac{\frac{1}{N} \sum_{n=0}^{N-1} \ln |X(\omega)|}{\frac{1}{N} \sum_{n=0}^{N-1} |X(\omega)|}. \quad (4.5)$$

На первых k фреймах определяются минимальные значения нормированной краткосрочной энергии E_{min} , доминантной частоты F_{min} и тонального коэффициента SFM_{min} . Также предварительно задаются пороговые значения E_{thr} , F_{thr} и SFM_{thr} .

Промежутки шума длиной менее десяти последовательных кадров и речи менее пяти кадров игнорируются, это осуществляется для того, чтобы предотвратить ложные срабатывания детектора.

Описание алгоритма детектора речи представлено на рисунке 4.5.

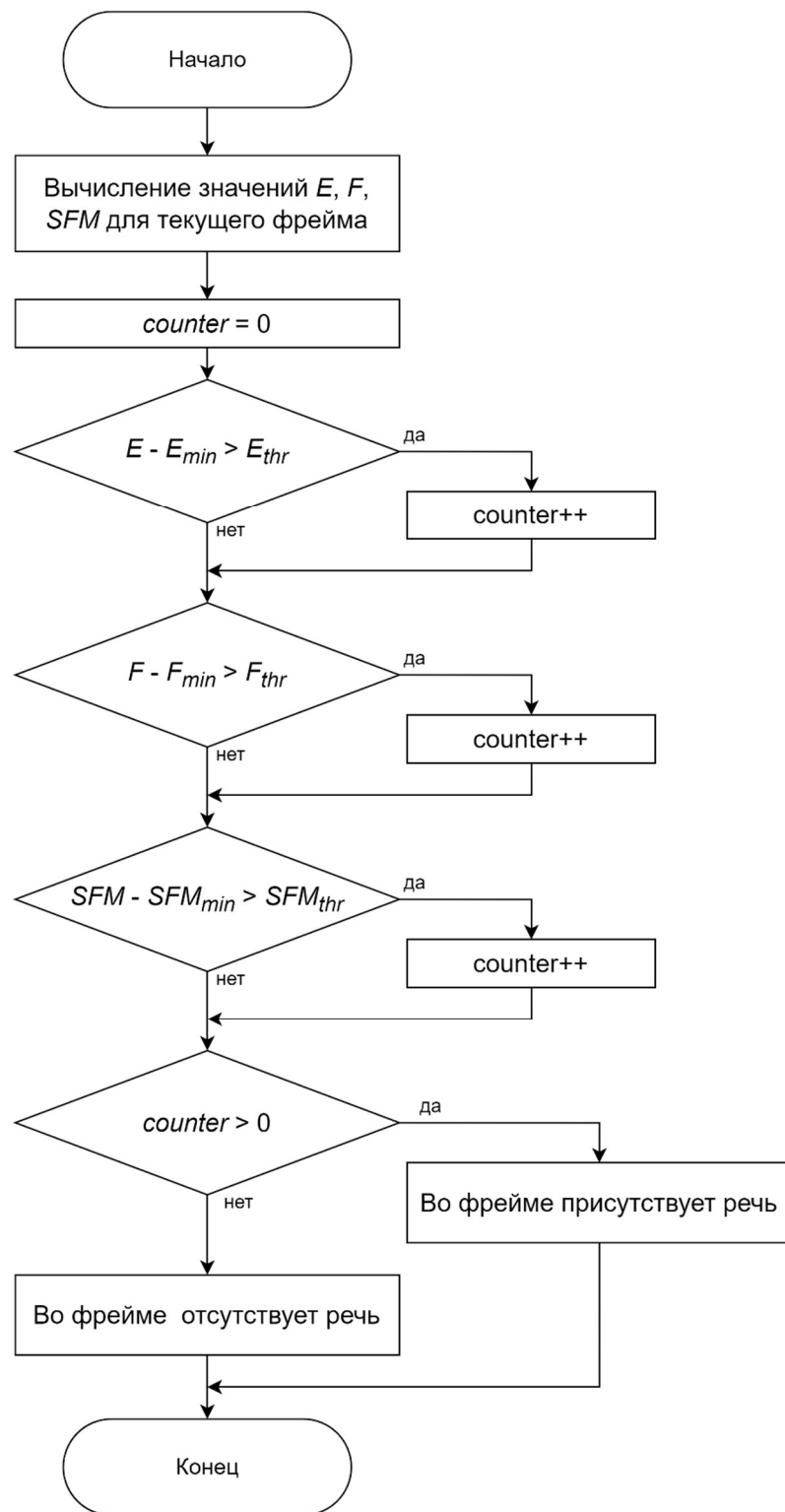


Рисунок 4.5 – Алгоритм оценки наличия речи

4.4.2 Вычисление ОСШ

Вычисление отношения сигнал/шум $segSNR$ (в дБ) определяется по формуле 4.6 [12]:

$$segSNR = 10 \lg \frac{\sum_1^{half_frame+1} |X(\omega)|^2}{\sum_1^{half_frame+1} |D(\omega)|^2}, \quad (4.6)$$

где $|X(\omega)|^2$ – уровень мощности звукового сигнала, $|D(\omega)|^2$ – уровень мощности шума.

4.4.3 Субполосное спектральное вычитание

Алгоритм субполосного спектрального вычитания заключается в следующих этапах:

1 Вычисляется ОСШ $subSNR$ для каждой частотной полосы по формуле 4.6.

2 В зависимости от $subSNR$ вычисляется коэффициент α для каждой полосы по формуле 4.7 [5]:

$$\alpha = \begin{cases} 10, & subSNR < -5 \\ 8 - \frac{3}{20} subSNR, & -5 \leq subSNR \leq 20 \\ 1, & \text{иначе.} \end{cases} \quad (4.7)$$

3 Выполняется субполосное спектральное вычитание и ограничение отрицательных компонентов по формулам 1.7, 1.8 соответственно.

4.4.4 Новая оценка мощности шума

Оценка мощности шума производится во фреймах, не содержащих речи. Для оценки мощности шума использовался алгоритм, описанный в статье [13]. Исходя из упомянутого алгоритма, оценка мощности шума $|\hat{N}(k, m)|$ производится по формуле:

$$|\hat{N}(k, m)| = \tilde{\alpha}_d(k, m-1) |\hat{N}(k, m-1)| + (1 - \tilde{\alpha}_d(k, m-1)) |\bar{\bar{X}}(k, m)|, \quad (4.8)$$

где $\bar{\bar{X}}(k, m)$ – зашумленный сигнал; $\tilde{\alpha}_d(k, m-1) = \alpha_d + (1 - \alpha_d) \hat{p}(k, m)$ – изменяющийся параметр сглаживания, который зависит от условной вероятности присутствия речи $\hat{p}(k, m)$; α_d – параметр сглаживания, который определяет время усреднения при отсутствии речи ($0 < \alpha_d < 1$). Вероятность

присутствия речи оценивается путем рекурсивного усреднения бинарного признака $I(k, m)$:

$$\hat{p}(k, m) = \alpha_p \hat{p}(k, m-1) + (1 - \alpha_p) I(k, m), \quad I(k, m) = \begin{cases} 1, & S_r(k, m) > \delta \\ 0, & S_r(k, m) \leq \delta \end{cases}, \quad (4.9)$$

где $S_r(k, m) = |\bar{X}(k, m)| / |\bar{X}_{min}(k, m)|$ – соотношение между текущей амплитудой зашумленного сигнала и минимальной амплитудой.

Для определения значения минимальной амплитуды $|\bar{X}_{min}(k, m)|$ используется две переменных со следующими начальными значениями: $|\bar{X}_{min}(k, m)| = |\bar{X}(k, 0)|$, $|\bar{X}_{tmp}(k, m)| = |\bar{X}(k, 0)|$. При каждой новой оценке шума текущее значение переменных получается путем сравнения текущей амплитуды и существующего минимума:

$$\begin{aligned} |\bar{X}_{min}(k, m)| &= \min (|\bar{X}_{min}(k-1, m)|, |\bar{X}(k, m)|) \\ |\bar{X}_{tmp}(k, m)| &= \min (|\bar{X}_{tmp}(k-1, m)|, |\bar{X}(k, m)|). \end{aligned} \quad (4.10)$$

Каждые $L = 0,5$ с переменные минимума обновляются следующим образом:

$$\begin{aligned} |\bar{X}_{min}(k, m)| &= \min (|\bar{X}_{tmp}(k-1, m)|, |\bar{X}(k, m)|) \\ |\bar{X}_{tmp}(k, m)| &= |\bar{X}(k, m)|. \end{aligned} \quad (4.11)$$

4.5 Синтез сигнала

Синтез выходного сигнала из буфера фреймов осуществляется методом перекрытия с суммированием. Перекрытие соседних фреймов составляет 50 %. Перед суммированием перекрывающихся участков каждый фрейм умножается на окно Хэмминга. Полученный выходной сигнал умножается на нормирующий коэффициент.

5 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

Программная реализация системы шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания осуществлялась с помощью пакета MATLAB.

5.1 Описание пакета MATLAB

MATLAB – это платформа, предназначенная для программирования и числовых вычислений, которая используется для разработки алгоритмов, создания моделей, а также для анализа данных. 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 редактор для создания сценариев, которые объединяют код, выходные данные и форматированный текст в исполняемом блокноте.

MATLAB предоставляет профессионально разработанные наборы инструментов, которые могут быть использованы для следующих целей [14]:

- проектирование, тестирование и внедрение систем управления;
- подготовка данных, проектирование, моделирование и развертывание глубоких нейронных сетей;
- получение, обработка и анализ изображений и видео для разработки алгоритмов и проектирования систем;
- обучение моделей, настройка параметров и развертывайте их в рабочей среде или на периферии для машинного обучения;
- разработка и внедрение программного обеспечения для мониторинга состояния и профилактического обслуживания;
- преобразование идей и концепций робототехники в автономные системы;
- анализ сигналов и данных временных рядов;
- моделирование и проектирование систем обработки сигналов;
- сбор и анализ данных;
- проектирование и тестирование систем беспроводной связи.

Язык MATLAB является высокоуровневым интерпретируемым языком программирования, который включает в себя структуры данных, основанные на матрицах, широкий спектр функций, а также объектно-ориентированные возможности и интерфейсы к программам, написанным с помощью других

языков программирования. В то время как другие языки программирования обычно работают с числами по одному, MATLAB работает с целыми матрицами и массивами.

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, в то время как скрипты используют общее рабочее пространство. Оба типа программ сохраняются в виде текстовых файлов и компилируются в машинный код динамически. Существует также возможность сохранять программы, обработанные в вид, удобный для машинного исполнения. В общем случае такие программы выполняются быстрее обычных, особенно если функция содержит команды построения графиков [15].

5.2 Программная реализация системы шумоподавления для речевых сигналов

На основе алгоритма, приведенного в разделе 4, был сформирован алгоритм программной реализации системы, содержащий следующие этапы:

- 1 Инициализация констант.
- 2 Считывание входного сигнала функцией *audioread*.
- 3 Разбиение сигнала на фреймы с перекрытием равным 50% функцией *signal_to_frames*.
- 4 Умножение каждого фрейма на окно Хэмминга с использованием функции *hamming*.
- 5 Оценка начальной мощности шума функций *start_noise*.
- 6 Определение фреймов, содержащих речь, с помощью функции *voice_activity_detection*.
- 7 Обработка сигнала фрейм за фреймом:
 - 7.1 Выполнение БПФ с помощью функции *fft*;
 - 7.2 Вычисление отношения сигнал/шум во фрейме функцией *snr*;
 - 7.3 Выполнение субполосного спектрального вычитания:
 - 7.3.1 Разбиение фрейма в частотной области на непересекающиеся полосы функцией *bands_borders*;
 - 7.3.2 Вычисление отношения сигнал/шум в каждой полосе;
 - 7.3.3 Определение коэффициента α для каждой полосы;
 - 7.3.4 Спектральное вычитание в каждой полосе;
 - 7.3.5 Ограничение отрицательных компонентов.

7.4 Дополнение симметричной части БПФ с помощью функций *flipud* и *conj*;

7.5 Выполнение ОБПФ с помощью функции *ifft*;

7.6 Выполнение новой оценки шума, в случае если во фрейме отсутствовала речь.

8 Синтез выходного сигнала.

9 Запись выходного сигнала в файл с помощью функции *audiowrite*.

Листинг программной реализации алгоритма работы системы шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания, приведенного в разделе 4, приведен в приложении Б.

5.3 Описание вспомогательных функций пакета MATLAB

5.3.1 Функция *audioread*

Функция *audioread* считывает данные из звукового файла с именем *filename* и возвращает матрицу данных *y* и частоту дискретизации *Fs*.

При указании параметра *samples* функция возвращает не весь файл целиком, а только выбранный диапазон. Параметр должен иметь форму [начало конец]. Необязательный параметр *dataType* указывает в каком формате будут считаны данные из файла.

Описание параметров данной функции представлено в таблице 5.1.

Таблица 5.1 – Описание параметров функции *audioread*

Параметр	Возможные значения	Описание
<i>filename</i>	Строка	Название считываемого файла
<i>samples</i> (необязательный)	числовой вектор [начало конец]	Номера начального и конечного отсчетов считываемого диапазона
<i>dataType</i> (необязательный)	'double'	Считываемые данные будут нормированы в диапазоне [-1; 1]
	'native'	Данные будут считаны в формате, в котором представлены в файле

Описание возвращаемых значений функцией *audioread* представлено в таблице 5.2.

Таблица 5.2 – Описание возвращаемых значений функцией *audioread*

Параметр	Возможные значения	Описание
F_s	Число	Частота дискретизации
y	Матрица, число строк которой соответствует числу отсчетов считанного сигнала, а число столбцов – числу звуковых каналов в файле	Считанный сигнал

5.3.2 Функция *hamming*

Функция *hamming* возвращает окно Хэмминга необходимой длины. Параметр L задает желаемый размер окна.

5.3.3 Функция *fft*

Функция *fft* вычисляет дискретное преобразование Фурье для параметра x , используя алгоритм быстрого преобразования Фурье.

В случае указания параметра n длина ДПФ будет равна этому параметру. Если длина параметра x меньше n , то вектор дополняется нулями, если больше – усекается до длины n . В случае, когда параметр n не указан, длина ДПФ равна длине вектора x .

Необязательный параметр *dim* определяет направление, в котором будет выполнено ДПФ, в случае, когда x – многомерный массив. Если этот параметр не указан, то преобразование выполняется по первому измерению массива, не равному единице.

Описание параметров функции *fft* представлено в таблице 5.3.

Таблица 5.3 – Описание параметров функции *fft*

Параметр	Возможные значения	Описание
1	2	3
x	Вектор-строка, вектор-столбец, многомерный массив	Входной параметр
n	Число	Длина ДПФ

Продолжение таблицы 5.3

1	2	3
<i>dim</i>	Число, не превышающее количество измерений массива	Определение направления вычисления ДПФ

Возвращаемое значение Y функции *fft* представляет собой результаты вычисления ДПФ. Если входной параметр x представляет собой вектор, то функция возвращает результат ДПФ вектора. Если x – матрица, то функция обрабатывает столбцы матрицы как векторы и возвращает ДПФ для каждого из них. В случае, когда x – многомерный массив, функция обрабатывает значения вдоль первого измерения массива, размер которого не равняется единице, как векторы и возвращает ДПФ каждого вектора. В случае, когда указан параметр *dim*, функция обрабатывает значения *dim* первого измерения массива.

5.3.4 Функция *flipud*

Функция *flipud* возвращает переданный параметр вектор-столбец A , записанный в обратном порядке. В случае, если передается матрица A , строки записываются в обратном порядке. Если A – вектор-строка, то возвращается неизмененный параметр A .

5.3.5 Функция *conj*

Параметр Z может представлять собой число, вектор или многомерную матрицу. Функция *conj* работает поэлементно и возвращаемое значение Z_c содержит комплексно-сопряженное число для каждого отдельного элемента передаваемого параметра Z .

5.3.6 Функция *ifft*

Функция *ifft* вычисляет обратное дискретное преобразование Фурье для параметра Y , используя алгоритм быстрого преобразования Фурье. Фурье-образ Y дополняется до длины n нулями, чтобы получить ОДПФ

необходимой длины. Необязательный параметр *dim* определяет направление, в котором будет выполнено ОДПФ, в случае, когда *x* – многомерный массив.

Описание параметров функции *ifft* представлено в таблице 5.4.

Таблица 5.4 – Описание параметров функции *ifft*

Параметр	Возможные значения	Описание
<i>Y</i>	Вектор-строка, вектор-столбец, многомерный массив	Входной параметр
<i>n</i>	Число	Длина ОДПФ
<i>dim</i>	Число, не превышающее количество измерений массива	Определение направления вычисления ОДПФ

Возвращаемое значение *x* функции *ifft* представляет собой результаты вычисления ОДПФ. Если входной параметр *Y* представляет собой вектор, то функция возвращает результат ОДПФ вектора. Если *Y* – матрица, то функция обрабатывает столбцы матрицы как векторы и возвращает ОДПФ для каждого из них. В случае, когда *Y* – многомерный массив, функция обрабатывает значения вдоль *dim* измерения массива, как векторы и возвращает ДПФ каждого вектора. В случае, когда *dim* не указан, происходит обработка вдоль первого измерения массива, размер которого превышает единицу.

5.3.7 Функция *audiowrite*

Функция *audiowrite* записывает матрицу аудиоданных *y* с частотой дискретизации *Fs* в файл с именем *filename*. В параметре *filename* также задается формат записываемого файла.

Описание параметров функции *ifft* представлено в таблице 5.5.

Таблица 5.5 – Описание параметров функции *audiowrite*

Параметр	Возможные значения	Описание
1	2	3
<i>y</i>	Матрица <i>mхn</i>	Аудиоданные длиной <i>n</i> с количеством каналов <i>m</i>

Продолжение таблицы 5.5

1	2	3
F_s	Натуральное число	Частота дискретизации
<i>filename</i>	Строка	Имя файла, в который будет производиться запись

У функции также необязательные парные параметры *Name*, *Value*, которые приведены в таблице 5.6.

Таблица 5.6 – Описание параметров *Name*, *Value* функции *audiowrite*

Параметр <i>Name</i>	Возможные значения параметра <i>Value</i>	Описание
<i>BitsPerSample</i>	16 (по умолчанию), 8, 24, 32, 64	Количество бит на отсчет. Только для форматов WAVE и FLAC
<i>BitRate</i>	128 (по умолчанию), 64, 96, 160, 192, 256, 320	Килобиты в секунду. Только для формата MPEG-4
<i>Quality</i>	75 (по умолчанию), число в интервале [0; 100]	Чем больше число, тем выше качество и ниже степень сжатия. Только для формата OGG
<i>Title</i>	Строка	Название
<i>Artist</i>		Исполнитель
<i>Comment</i>		Комментарий

Функция *audiowrite* не имеет возвращаемых значений.

5.4 Описание пользовательских функций

5.4.1 Функция *signal_to_frames*

Функция *signal_to_frames* разбивает входящий параметр *input_signal* на фреймы заданного размера с перекрытием 50 % и возвращает буфер фреймов *frame_array*. В случае, когда длина последнего фрейма, не совпадает с заданной длиной, этот фрейм дополняется нулями до необходимой длины, а параметр *remainder* показывает какой длины был последний фрейм.

Алгоритм работы функции представлен в подпункте 4.2.

5.4.2 Функция *start_noise*

Функция *start_noise* производит начальную оценку мощности шума по первым k фреймам из буфера *frame_array*. Возвращаемым значением функции является начальная оценка шума *noise*.

Алгоритм работы функции представлен в подпункте 4.3.

5.4.3 Функция *processing_frames*

Функция *processing_frames* выполняет обработку буфера *frame_array* фрейм за фреймом.

Описание параметров и возвращаемых значений функции *processing_frames* приведены в таблицах 5.7 и 5.8 соответственно.

Таблица 5.7 – Описание параметров функции *processing_frames*

Параметр	Описание
<i>frame_array</i>	Буфер фреймов
<i>noise_abs</i>	Начальная оценка мощности шума
F_s	Частота дискретизации

Таблица 5.8 – Описание возвращаемых значений функцией *processing_frames*

Параметр	Описание
<i>frame_array</i>	Буфер фреймов, прошедший обработку фрейм за фреймом
<i>segSNR_array</i>	Буфер, содержащий значение ОСШ для каждого фрейма из буфера <i>frame_array</i>
<i>speech</i>	Буфер, показывающий присутствует ли речь во фрейме из буфера <i>frame_array</i> .

5.4.4 Функция *voice_activity_detection*

В функцию входным параметром передается буфер фреймов *frame_array*. Возвращаемым значением функции является буфер *speech*, значения которого соответствуют фреймам из буфера *frame_array* и содержат информацию присутствует или отсутствует речь в конкретном фрейме.

Алгоритм работы функции описан в подпункте 4.4.1.

5.4.5 Функция *processing_frame*

Функция *processing_frame* осуществляет обработку фрейма *frame*, в том числе определяет ОСШ *segSNR* для фрейма, выполняет спектральное вычитание и осуществляет новую оценку мощности шума, в случае если во фрейме отсутствует речь. Алгоритм работы функции описан в подпункте 4.4.

Описание параметров и возвращаемых значений функции *processing_frame* приведены в таблицах 5.9 и 5.10 соответственно.

Таблица 5.9 – Описание параметров функции *processing_frame*

Параметр	Описание
<i>frame</i>	Текущий фрейм
<i>noise_abs</i>	Оценка мощности шума
<i>Fs</i>	Частота дискретизации
<i>p</i>	Вероятность присутствия речи в предыдущем фрейме
<i>Xmin</i>	Переменные минимума, используемые при новой оценке шума
<i>Xtmp</i>	
<i>time_counter</i>	Счетчик времени, используемый управления обновлением переменных <i>Xmin</i> и <i>Xtmp</i>
<i>i</i>	Номер текущего фрейма
<i>speech</i>	Параметр, указывающий наличие или отсутствие речи во фрейме

Таблица 5.10 – Описание возвращаемых значений функцией *processing_frame*

Параметр	Описание
<i>frame</i>	Фрейм, полученный после выполнения субполосного спектрального вычитания
<i>new_noise_abs</i>	Новая оценка мощности шума
<i>segSNR</i>	Значение ОСШ для фрейма
<i>p</i>	Обновление входных параметров, в случае если производилась новая оценка шума
<i>Xmin</i>	
<i>Xtmp</i>	
<i>time_counter</i>	Обновленный параметр <i>time_counter</i>

5.4.6 Функция *snr*

Функция *snr* принимает векторы с мощностью звукового сигнала *input_signal_abs2* и с мощностью шума *noise_abs2* и возвращает их ОСШ.

5.4.7 Функция *spectral_subtraction*

Функция *spectral_subtraction* принимает вектор мощности звукового сигнала *input_signal_abs2*, оценку мощности шума *noise_abs2* и частоту дискретизации *Fs*.

Функция разбивает сигналы *input_signal_abs2* и *noise_abs2* на заданное количество полос, после чего осуществляет спектральное вычитание в каждой из полос, ограничивает отрицательные компоненты и возвращает вектор мощности очищенного сигнала *speech_abs2*.

5.4.8 Функция *band_borders*

Функция *band_borders* принимает входной параметр *Fs*, соответствующий частоте дискретизации звукового сигнала, и определяет коэффициенты для неперекрывающихся полос спектральной мощности сигнала. Возвращаемый параметр *band* представляет собой тип данных *cell*, каждая ячейка которого содержит диапазон коэффициентов для отдельной полосы.

5.4.9 Функция *new_noise*

Параметр *input_signal_abs* функции *new_noise* является вектором мощности звукового сигнала. Функция производит новую оценку мощности шума для фрейма.

Параметры *noise_abs*, *Fs*, *p*, *Xmin*, *Xtmp*, *time_counter*, *i* и выходные значения *new_noise_abs*, *time_counter*, *p*, *Xmin*, *Xtmp* функции *new_noise* соответствуют параметрам и выходным значениям с таким же именами функции *processing_frame* и описаны в таблицах 5.9, 5.10.

Описание алгоритма работы функции приведено в подпункте 4.4.4.

5.4.10 Функция *synthesis_signal*

Функция *synthesis_signal* принимает буфер фреймов *frame_array* и параметр, позволяющий восстановить изначальную длину сигнала, *remainder*.

Функция формирует выходной сигнал *signal* из буфера фреймов методом суммирования с перекрытием, восстанавливает изначальную длину сигнала и умножает получившийся сигнал на нормирующий коэффициент.

Алгоритм работы функции описан в подпункте 4.5.

6 ТЕХНИКО-ЭКОНОМИЧЕСКОЕ ОБОСНОВАНИЕ РАЗРАБОТКИ СИСТЕМЫ ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

Проводимая научно-исследовательская работа заключается в проектировании системы шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания.

В большинстве случаев при записи речь подвержена воздействию шумов различного рода, что вызывает ухудшение разборчивости речи человеком при воспроизведении речевых сообщений. Исходя из этого, создание систем шумоподавления является актуальной задачей.

Разрабатываемая система предназначена для использования в программах, осуществляющих запись речевых сообщений и предоставляющих пользователю конечный аудиофайл. В результате пользователь будет получать сигнал, очищенный от посторонних шумов.

Преимущество данной системы над рядом существующих аналогов заключается в том, что она позволяет очищать речевой сигнал от различных типов шумов, а также использует алгоритм с низкой ресурсоемкостью, путем чего достигается высокая скорость обработки сигнала от шума.

6.1 Расчет затрат на проведение исследования системы шумоподавления для речевых сигналов

6.1.1 Разработка плана проведения научно-исследовательской работы

Данная работа проводится в рамках внутреннего исследования университета с дальнейшим развитием в научных статьях и журналах. План проведения научно-исследовательской работы представлен в таблице 6.1. В данном плане отражены этапы проектирования системы шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания, а также количество и состав исполнителей: научный руководитель и ответственный исполнитель.

Основными методами расчета трудоемкости являются система аналогов, метод прямого счета и метод экспертных оценок. В данной работе при определении трудоемкости использовался метод прямого счета, при котором трудоемкость обосновывается руководителем совместно с исследователем [16].

Таблица 6.1 – План проведения научно-исследовательской работы

Наименование этапа и вида работы	Исполнитель (должность, квалификация)	Численность исполнителей, чел.	Продолжительность выполнения работы, д.
1	2	3	4
1 Сопоставление и утверждение ТЗ на НИР	Научный руководитель, ответственный исполнитель	2	0,5
2 Сбор научно-технической литературы и других материалов, относящихся к теме исследования	Ответственный исполнитель	1	2
3 Изучение собранных материалов и научно-технической литературы, относящихся к теме исследования	Ответственный исполнитель	1	5
4 Проведение патентных исследований	Ответственный исполнитель	1	2
5 Составление аналитического обзора состояния вопросов по теме	Ответственный исполнитель	1	2
6 Формулирование возможных направлений решения задач, поставленных в ТЗ на НИР	Ответственный исполнитель	1	2
7 Сравнительная оценка возможных направлений решения задач, поставленных в ТЗ на НИР	Ответственный исполнитель	1	2
8 Выбор и обоснование принятого направления проведения исследований способов решения поставленных задач	Научный руководитель, ответственный исполнитель	2	2
9 Разработка общей методики проведения исследований	Ответственный исполнитель	1	2,5

Продолжение таблицы 6.1

1	2	3	4
10 Проектирование структуры системы шумоподавления для речевых сигналов	Ответственный исполнитель	1	4
11 Проектирование алгоритма работы системы шумоподавления для речевых сигналов	Ответственный исполнитель	1	5
12 Программная реализация структуры системы шумоподавления для речевых сигналов	Ответственный исполнитель	1	3
13 Программная реализация алгоритма работы системы шумоподавления для речевых сигналов	Ответственный исполнитель	1	5
14 Составление промежуточного отчета и его рассмотрение	Научный руководитель, ответственный исполнитель	2	2
15 Выявление необходимости проведения экспериментов для подтверждения отдельных положений теоретических исследований	Ответственный исполнитель	1	1
16 Разработка частных методик проведения экспериментальных исследований.	Ответственный исполнитель	1	4
17 Подготовка моделей, а также испытательного оборудования, необходимых для проведения экспериментальных исследований	Научный руководитель, ответственный исполнитель	2	2
18 Проведение экспериментальных исследований	Ответственный исполнитель	1	2
19 Сопоставление результатов экспериментов с теоретическими исследованиями	Научный руководитель, ответственный исполнитель	2	1
20 Корректировка теоретических моделей исследований	Научный руководитель, ответственный исполнитель	2	1

Продолжение таблицы 6.1

1	2	3	4
21 Проведение дополнительных экспериментов	Ответственный исполнитель	1	2
22 Обобщение результатов предыдущих этапов работы	Научный руководитель, ответственный исполнитель	2	1
23 Оценка полноты решения поставленных задач	Ответственный исполнитель	1	4
24 Составление и оформление отчета	Ответственный исполнитель	1	11
25 Рассмотрение результатов проведенной НИР	Научный руководитель	1	1
Итого			81,5

Результат выполнения дипломного проектирования научно-исследовательского характера относится к научно-технической продукции, которая в условиях рыночных отношений является товаром. Поэтому важным вопросом технико-экономического обоснования выступает определение цены основного результата дипломного проекта [16].

6.1.2 Расчет затрат на топливно-энергетические ресурсы

В данный подпункт включаются затраты на топливно-энергетические ресурсы, потребляемые оборудованием, используемым для научно-экспериментальных целей.

Таблица 6.2 – Расчет затрат на топливно-энергетические ресурсы

Наименование оборудования, используемого для научно-экспериментальных и технологических целей	Установочная мощность, кВт	Время использования, ч	Тариф за 1 кВт·ч	Сумма, р.
Ноутбук ASUS TUF Gaming FX505DD-BQ067	0,12	652	0,2321	18,16
Итого				18,16

6.1.3 Расчет затрат на спецоборудование, приобретаемое для научно-исследовательской работы

Для проведения исследовательских работ необходимо приобретение устройств и программного обеспечения, необходимого для осуществления научно-исследовательской работы.

Информация о ценах на программное обеспечение, используемое во время выполнения научно-исследовательской работы, получена на официальном сайте поставщика [17].

Таблица 6.3 – Расчет затрат на спецоборудование

Наименование специальных инструментов, приспособлений, приборов, стендов, устройств и другого специального оборудования	Количество, шт.	Цена, р.	Сумма, р.
1 Ноутбук ASUS TUF Gaming FX505DD-BQ067	1	2 195,95	2 195,95
2 Программное обеспечение Matlab (Student license) для исполнителя	1	82,67	82,67
3 Программное обеспечение Matlab (Annual license) для научного руководителя	1	783,92	783,92
Итого			3 062,54

6.1.4 Расчет основной заработной платы научно-технического персонала

Согласно плану проведения научно-исследовательской работы в исследовании задействовано два специалиста: ответственный исполнитель и научный руководитель.

Величина затрат определяется исходя из численности различных категорий исполнителей и трудоемкости выполнения отдельных видов работ (см. табл. 6.1), дневных тарифных ставок, должностных окладов, используемых премиальных систем оплаты труда исполнителей.

Тарифные ставки работников были взяты по данным организации, в которой проводились исследования.

Таблица 6.4 – Расчет основной заработной платы исследователей

Категории исполнителей	Численность исполнителей, чел.	Дневной оклад (тарифная ставка), р.	Трудоемкость, д.	Сумма, р.
Научный руководитель	1	19,37	10,5	203,39
Ответственный исполнитель	1	11,40	71	809,4
Итого				1 012,79
Премия и иные стимулирующие выплаты (25 %)				253,20
Всего				1 265,99

6.1.5 Расчет дополнительной заработной платы научно-технического персонала

Дополнительная заработная плата Z_d научно-технического персонала рассчитывается по формуле 6.1:

$$Z_d = \frac{Z_o \cdot H_d}{100} = \frac{1\,265,99 \cdot 15}{100} = 189,90 \text{ р.}, \quad (6.1)$$

где Z_o – основная заработная плата, рассчитанная в таблице 6.4; H_d – норматив дополнительной заработной платы ($H_d = 10 - 20\%$).

6.1.6 Расчет отчислений на социальные нужды

Отчисления в фонд социальной защиты населения и по обязательному страхованию от несчастных случаев на производстве вычисляется по формуле 6.2:

$$P_{\text{соц}} = \frac{(Z_o + Z_d) \cdot H_{\text{соц}}}{100} = \frac{(1\,265,99 + 189,90) \cdot 34,6}{100} = 503,74 \text{ р.}, \quad (6.2)$$

где $H_{\text{соц}}$ – норматив отчислений в ФСЗН и Белгосстрах ($H_{\text{соц}} = 34,6\%$).

6.1.7 Расчет накладных затрат

Расчет накладных затрат выполняется по формуле 6.3:

$$P_{\text{накл}} = \frac{Z_o \cdot H_{\text{накл}}}{100} = \frac{1\,265,99 \cdot 65}{100} = 822,89 \text{ р.}, \quad (6.3)$$

где $H_{\text{накл}}$ – норматив накладных расходов ($H_{\text{накл}} = 50 - 80\%$).

6.1.8 Расчет полной себестоимости

Полная себестоимость научно-исследовательской работы C_{Π} рассчитывается по следующей формуле:

$$C_{\Pi} = Z_{\text{тэр}} + Z_{\text{сп}} + Z_o + Z_d + P_{\text{соц}} + P_{\text{накл}} = 18,16 + 3\,062,54 + 1\,265,99 + 189,90 + 503,74 + 822,89 = 5\,863,22 \text{ р.}, \quad (6.4)$$

где $Z_{\text{тэр}}$ – затраты на топливно-энергетические ресурсы, рассчитанные в таблице 6.1; $Z_{\text{сп}}$ – затраты на спецоборудование, рассчитанные в таблице 6.2.

6.1.9 Расчет плановой прибыли

Плановая прибыль Π научно-исследовательской работы определяется по формуле 6.5:

$$\Pi = \frac{C_{\Pi} \cdot P_{\text{н.т.п}}}{100} = \frac{5\,863,22 \cdot 30}{100} = 1\,758,97 \text{ р.}, \quad (6.5)$$

где $P_{\text{н.т.п.}}$ – рентабельность научно-технической продукции ($P_{\text{н.т.п.}} = 30 - 50\%$).

6.1.10 Расчет сметной стоимости научно-технической продукции

Сметная стоимость научно-технической продукции $\Pi_{\text{н.т.п}}$ вычисляется с помощью формулы 6.6:

$$\Pi_{\text{н.т.п.}} = C_{\Pi} + \Pi = 5\,863,22 + 1\,758,97 = 7\,622,19 \text{ р.} \quad (6.6)$$

Данная научно-исследовательская работа проводилась исследовательской командой в составе двух специалистов (научный руководитель и ответственный исполнитель) и в результате сметная стоимость научно-технической продукции составит 7 622,19 р.

6.2 Оценка уровня качества разработанной системы шумоподавления для речевых сигналов

Для расчета уровня качества разработанной системы будут использованы следующие критерии: новизна, значимость для науки и практики, объективность. Каждому критерию дается характеристика и уровень в баллах, чем выше балл, тем выше результат исследований по критерию. Осуществляется оценка значимости каждого из критериев. Полученные оценки значимости нормируются, чтобы сумма всех нормированных коэффициентов значимости $K_{н.зи}$ i -го критерия была равна единице. Описание критериев и оценок результатов приведено в таблице 6.5.

Таблица 6.5 – Оценка критериев исследования и достигнутых результатов

Критерий	Оценка критерия	Уровень	Оценка	$K_{н.зи}$
1	2	3	4	5
Новизна	Сделан элементарный анализ связей и взаимозависимостей между фактами. Дана классификация фактов.	2	0,2	0,07
Значимость для науки и практики	Результат окажет положительное влияние на развитие отдельного научного направления	3	0,6	0,2
Объективность на основе учета квалификации и компетенции разработчиков и экспертов	Результат получен без участия научных работников высшей квалификации. Рассмотрен и оценен секцией Ученого совета.	3	0,8	0,27
Объективность на основе форм результатов	Результат одобрен Ученым советом.	2	0,4	0,13

Продолжение таблицы 6.5

1	2	3	4	5
Доказательность	Результат получен на основе экспериментальных и теоретических исследований, лабораторных испытаний с использованием математических методов обработки данных	3	1	0,33

Количественная оценка уровня научно-технического результата, полученного в дипломном проекте, осуществляется на основе комплексного показателя достигнутого уровня качества K_k , рассчитываемого по формуле

$$K_k = \sum_{i=1}^n K_{нzi} \cdot B_{ди} , \quad (6.7)$$

где n – количество критериев, $B_{ди}$ – достигнутый уровень по i -му критерию.

На основе показателей, приведенных в таблице 6.5, рассчитывается комплексный показатель достигнутого уровня качества дипломного проекта:

$$K_k = 0,07 \cdot 2 + 0,2 \cdot 3 + 0,27 \cdot 3 + 0,13 \cdot 2 + 0,33 \cdot 3 = 2,8. \quad (6.8)$$

Максимальное возможное значение комплексного показателя качества равно 5. Чем ближе значение K_k к значению 5, тем больше полученные результаты соответствуют современным требованиям. Значения K_k не менее 3 свидетельствуют о том, что результаты соответствуют современным требованиям, иначе они не удовлетворяют по уровню качества современным требованиям.

На основании полученного значения K_k , равного 2,8, можно сделать вывод, что результаты дипломного проекта на данный момент не удовлетворяет современным требованиям. Такие результаты были получены, из-за отсутствия публикаций результатов научно-исследовательской работы и отсутствия масштабного научного эксперимента, а также из-за невысокого показателя новизны. Дальнейшее проведение более масштабных исследований, а также внедрение алгоритма в устройства позволят увеличить комплексный показатель качества до более высоких значений.

7 АНАЛИЗ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ СИСТЕМЫ ШУМОПОДАВЛЕНИЯ ДЛЯ РЕЧЕВЫХ СИГНАЛОВ

7.1 Тестирование разработанной системы шумоподавления

Тестирование разработанной системы шумоподавления для речевых сигналов произведено путём MATLAB моделирования. На чистый речевой сигнал накладывался белый шум (с различными значениями ОСШ в промежутке от 0 до 25 дБ), после чего полученный зашумленный сигнал подавался на вход разработанной системы шумоподавления.

После тестирования системы было установлено, что при очистке речевого сигнала с ОСШ не менее 20 дБ формируется выходной сигнал высокого качества, где музыкальные шумы минимальны. При очистке сигнала с ОСШ = 15 дБ качество очищенной речи остается на достаточно высоком уровне, однако музыкальные шумы усиливаются. В случаях, когда ОСШ не более 10 дБ, музыкальные шумы в очищенном сигнале начинают влиять на качество речи, однако задача по очистке сигнала от шума выполняется.

На рисунке 7.1 показаны результаты тестирования системы на речевом сигнале. На рисунке отображены спектрограммы зашумленного (ОСШ = 15 дБ) и очищенного (ОСШ = 33 дБ) сигналов, а также показаны значения ОСШ для каждого фрейма.

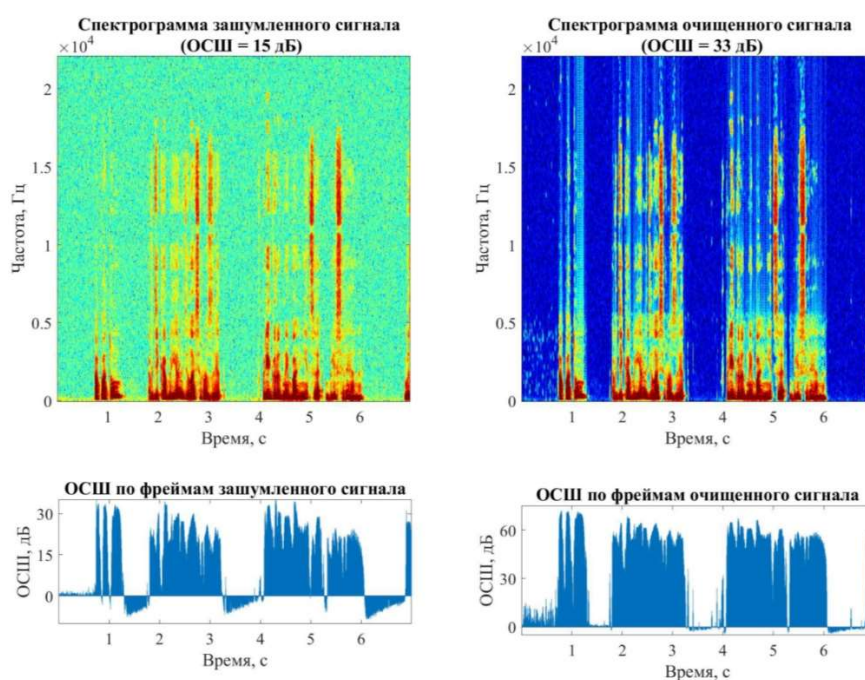


Рисунок 7.1 – Тестирование системы шумоподавления на речевом сигнале

Спектрограммы исходного, зашумленного и очищенного сигналов, а также их временные представления и значения ОСШ для каждого фрейма приведены на плакате ГУИР.460818.002 ПЛ. Листинг программы для тестирования представлен в приложении Б.

Также разработанная система шумоподавления была протестирована на других звуковых сигналах:

- звуковой сигнал пения птиц;
- звуковой сигнал игры на фортепиано.

Результаты моделирования показывают, что разработанная система шумоподавления позволяет очистить звуковой сигнал с пением птиц с ОСШ не менее 15 дБ и сигналы с игрой на фортепиано с ОСШ не менее 20 дБ с высоким уровнем качества (итоговые значения ОСШ равны 34 и 54 дБ соответственно). При очистке сигналов с указанными значениями ОСШ музыкальные шумы в очищенных сигналах были минимальны.

Результаты тестирования разработанной системы на звуковых сигналах с пением птиц и игрой на фортепиано представлены на рисунках 7.2 и 7.3 соответственно.

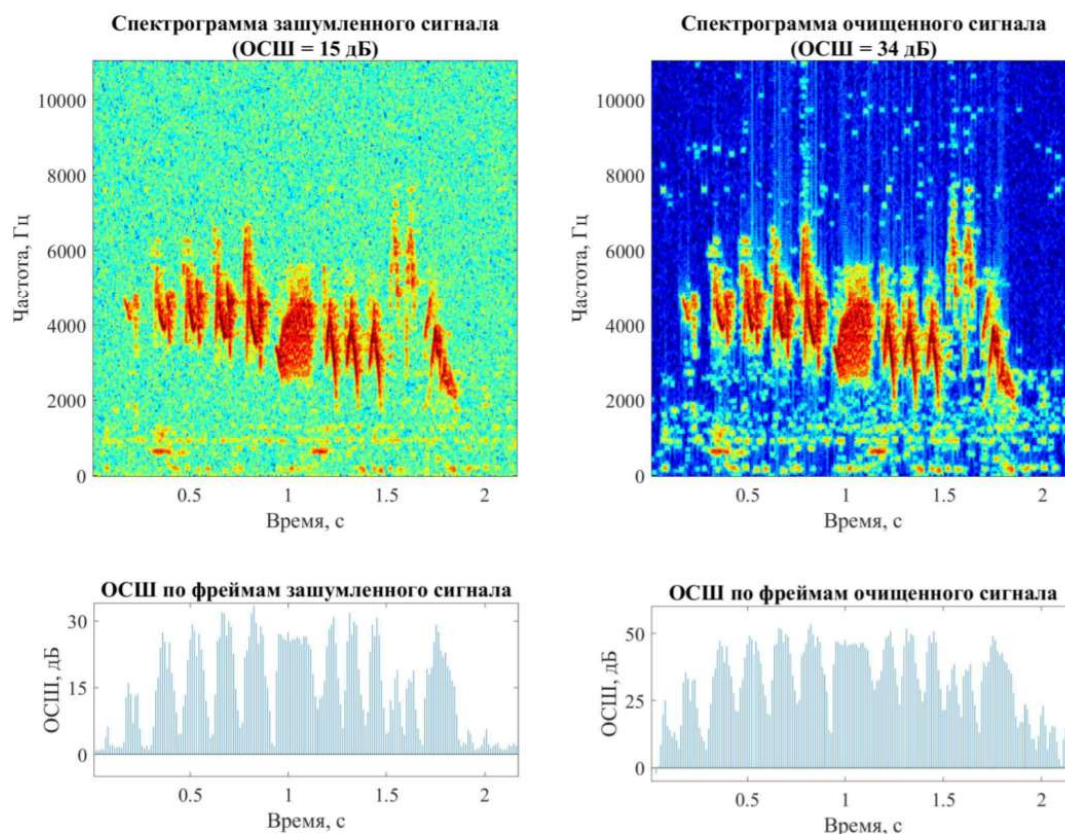


Рисунок 7.2 – Тестирование системы шумоподавления на сигнале с пением ПТИЦ

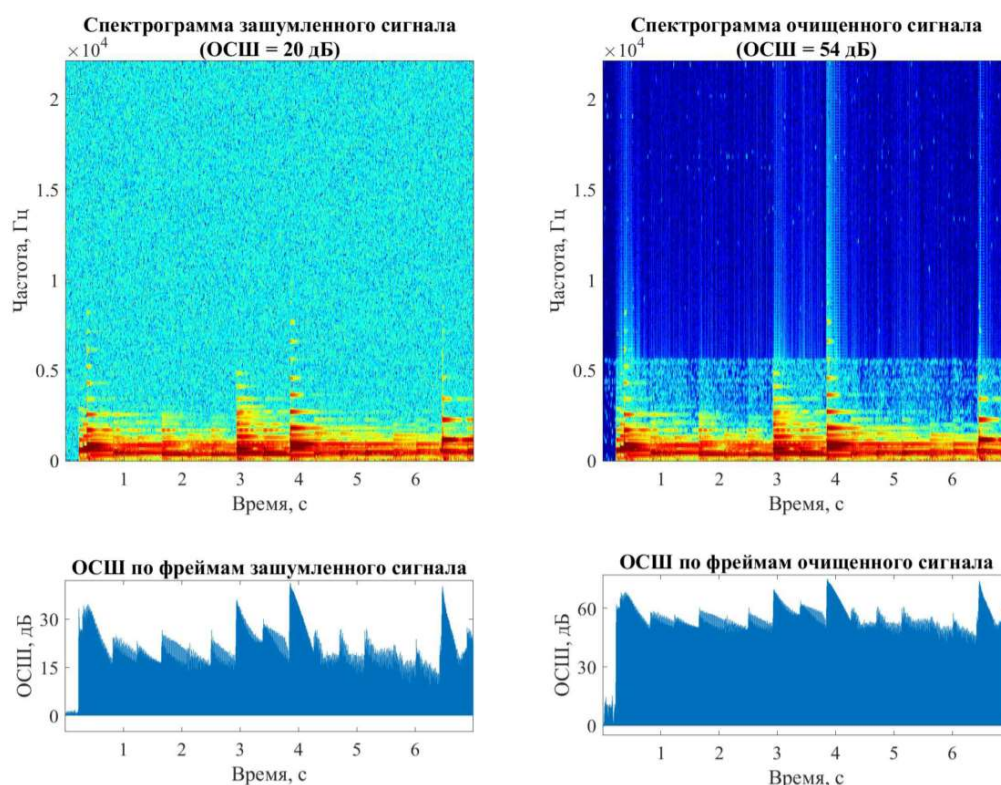


Рисунок 7.3 – Тестирование системы шумоподавления на сигнале с игрой на фортепиано

Таким образом, разработанная система шумоподавления может быть использована не только для шумоподавления в речевых сигналах, но и в других звуковых сигналах.

7.2 Экспериментальное исследование системы шумоподавления для речевых сигналов

Также для тестирования разработанной системы шумоподавления для речевых сигналов был проведен эксперимент. Целью эксперимента было определить влияние работы системы шумоподавления на изменение разборчивости речи в зашумлённых и очищенных речевых сигналах.

7.2.1 Описание проводимого эксперимента

Каждому участнику предлагается прослушать шесть заранее подготовленных записей. Одна запись содержит артикуляционную таблицу, состоящую из 50 слов. После каждого произнесенного слова из таблицы в записи делается пауза, во время которой участнику необходимо произнести услышанное им слово, правильность которого фиксируется. Далее

подсчитывается количество правильно повторенных слов и определяется процент разборчивости речи.

7.2.2 Подготовка данных для проведения эксперимента

Для проведения эксперимента было необходимо сформировать для каждого участника набор из шести обработанных аудиосигналов:

- 1 Сигнал, зашумленный белым шумом с ОСШ 5 дБ (далее white).
- 2 Сигнал, зашумленный розовым шумом с ОСШ 5 дБ (далее pink).
- 3 Сигнал, зашумленный шумом метрополитена с ОСШ 5 дБ (далее metro).

4 Сигнал, зашумленный белым шумом с ОСШ 5 дБ и затем очищенный разработанной системой шумоподавления (далее white+NR).

5 Сигнал, зашумленный розовым шумом с ОСШ 5 дБ и затем очищенный разработанной системой шумоподавления (далее pink+NR).

6 Сигнал, зашумленный шумом метрополитена с ОСШ 5 дБ и затем очищенный разработанной системой шумоподавления (далее metro+NR).

Все шесть записей для одного участника содержат разные наборы слов, чтобы исключить фактор запоминания, который может повлиять на конечный результат эксперимента. Каждая из используемых таблиц была обработана шестью способами, перечисленными выше, таким образом было сформировано 36 тестовых записей (см. рисунок 7.4).

Таблица для каждого типа обработки определялась случайным образом для каждого участника. Информация о соответствии типа обработки номеру таблицы для каждого участника приведена в таблице 7.1.

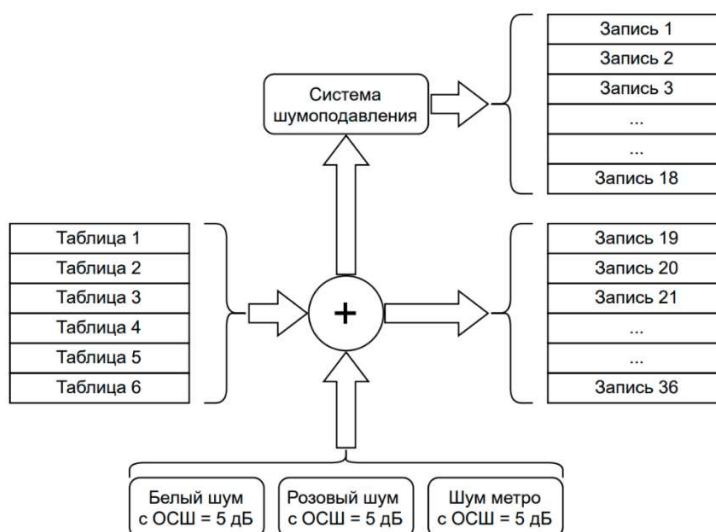


Рисунок 7.4 – Формирование тестовых записей

Схема данных для экспериментального исследования эффективности работы системы повышения шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания представлена в графическом материале ГУИР.460818.003 ПД.

Таблица 7.1 – Соответствие номера таблицы типу обработки для каждого участника

Участник	White	Pink	Metro	White+NR	Pink+NR	Metro+NR
1	2	3	4	5	6	7
1	3	4	1	2	6	5
2	6	3	2	1	4	5
1	2	3	4	5	6	7
3	2	4	5	3	6	1
4	1	2	6	3	5	4
5	3	4	6	1	5	2
6	5	4	2	3	1	6
7	4	2	3	1	6	5
8	5	6	1	4	2	3
9	6	4	3	5	2	1

Для проведения эксперимента было использовано шесть артикуляционных таблиц из ГОСТ 16600-72 [16], представленных в таблицах 7.2 – 7.7.

Таблица 7.2 – Первая артикуляционная таблица

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	месть	срам	брак	флюс	мал	хан
бук	вид	путь	пуст	туп	бант	мост	цен	сруб	млад

Таблица 7.3 – Вторая артикуляционная таблица

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 7.4 – Третья артикуляционная таблица

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
борт	сerp	наш	лаз	рать	кол	нам	май	печь	зам

Таблица 7.5 – Четвертая артикуляционная таблица

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
мрём	льём	пьёшь	таз	пёк	месть	рис	врёт	вьют	полк

Таблица 7.6 – Пятая артикуляционная таблица

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
тёк	час	пьёшь	пьём	трём	долг	пьют	трёт	лис	месть

Таблица 7.7 – Шестая артикуляционная таблица

вставь	сноб	лень	груз	там	пол	знал	стар	весь	нам
трав	штос	рак	вас	быт	сов	чур	гроб	час	вон
толь	краг	гад	как	дыр	кум	гак	горд	пал	лет
мех	тень	бот	парт	хрен	бас	щель	сечь	быть	мот
рать	баз	баш	герб	порт	гам	печь	лай	вам	зол

После осуществления подготовки данных был проведен эксперимент, результаты которого приведены в пункте 7.2.3.

7.2.3 Результаты проведенного эксперимента

В проведенном эксперименте приняло участие девять человек в возрасте от 17 до 19 лет. После проведения эксперимента были получены данные об изменении разборчивости речи. Далее были проведены расчеты для определения влияния присутствия шума на разборчивость речи,

результаты которых представлены в таблице 7.8. Разборчивость речи определяется соотношением правильно услышанных слов ко всем словам, содержащимся в тестовой записи.

Таблица 7.8 – Изменение разборчивости речи для каждого участника

Участник	Разборчивость речи, %					
	White	Pink	Metro	White+NR	Pink+NR	Metro+NR
1	82	76	84	88	78	92
2	56	88	72	78	82	76
3	86	84	94	90	94	98
4	84	84	88	92	88	96
5	88	88	94	88	92	96
6	84	86	88	90	90	92
7	88	80	90	92	82	90
8	90	84	92	98	96	96
9	82	88	84	80	92	92

В таблице 7.9 приведены данные о процентах разборчивости речи для различных типов обработки исходных сигналов, а в таблице 7.10 – данные о приросте разборчивости.

Таблица 7.9 – Изменение разборчивости речи при различных типах обработки исходных сигналов

Тип обработки сигнала	Разборчивость речи, %		
	Минимальная	Максимальная	Средняя
White	56	90	82
Pink	76	88	84
Metro	72	94	87
White+NR	78	98	88
Pink+NR	78	96	88
Metro+NR	76	98	92

Таблица 7.10 – Изменение прироста разборчивости речи для различных шумов

Тип шума	Прирост разборчивости речи, %		
	Минимальная	Максимальная	Средняя
1	2	3	4
Белый шум	-2	22	3,7

Продолжение таблицы 7.10

1	2	3	4
Розовый шум	-6	12	2,4
Шум метрополитена	0	8	2,8

Исходя из данных, полученных при проведении эксперимента, следует, что разработанная система снижает уровень шума в речевых сигналах и тем самым увеличивает разборчивость речи.

После проведения экспериментального исследования была проведена проверка значимости полученных результатов для исключения возможности получения случайных результатов из-за небольшой выборки. В данной работе проверка выполнена при помощи функции *ttest* пакета MATLAB, использующей парный *t*-критерий. Алгоритм такой проверки заключается в том, что для каждого участника эксперимента вычисляется разница между разборчивостью речи для зашумленного и очищенного сигналов. При этом нулевая гипотеза подразумевает, что истинная разность равна нулю. В результате проверки значимости результатов эксперимента нулевая гипотеза была отвергнута. Итоговый уровень значимости составил $p = 0.045$, что говорит о том, что вероятность того, что результаты эксперимента не являются значимыми очень мала (4,5%).

Результаты проведенного эксперимента, а также спектрограммы используемых для эксперимента сигналов, представлены в графическом документе ГУИР.460818.003 ПЛ к данному проекту.

7.3 Анализ результатов тестирования

Исходя из данных, полученных при тестировании разработанной системы шумоподавления для речевых сигналов в среде MATLAB и при проведении эксперимента, следует, что разработанная система снижает уровень шума в речевых сигналах, тем самым уменьшает негативное воздействие шума на слуховую систему человека. Также использование разработанной системы помогает увеличить разборчивость речи в сигналах с низким ОСШ.

Главным недостатком разработанной системы является появление музыкальных шумов, которые искажают обработанный сигнал. Во время тестирования системы было обнаружено, что большие музыкальные шумы возникают при очистке сигналов с низким ОСШ, в то время как при очистке сигнала с относительно высоким ОСШ музыкальные шумы не существенны.

ЗАКЛЮЧЕНИЕ

В дипломном проекте была разработана система шумоподавления для речевых сигналов на основе метода субполосного спектрального вычитания.

Реализация системы соответствует требованиям технического задания. В процессе разработки были учтены следующие требования:

- независимость от типа аддитивного шума;
- высокая скорость обработки сигнала от шума.

Перечисленные требования были удовлетворены путем выбора алгоритма субполосного спектрального вычитания для речевых сигналов для осуществления шумоподавления в речевых сигналах.

Выполненная программная реализация проекта системы удовлетворяет следующим требованиям технического задания:

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

Для проверки работоспособности системы было осуществлено ее моделирование в среде MATLAB. С помощью смоделированной системы были подготовлены данные для эксперимента, который показал, что разработанная система выполняет очистку сигнала от шума и тем самым повышает разборчивость речи. Недостатком разработанной системы является появление музыкальных шумов, влияющих на качество очищенного сигнала, при очистке сигналов с низким ОСШ.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

- [1] Исследование методов очистки русской речи от шумов / Е. Г. Жилияков [и др.] // Информационные системы и технологии. – 2010. – №5. – С. 17-25.
- [2] Шумоподавление на основе перцептуальных алгоритмов спектрального вычитания и обработки сигналов в подпространствах / А. А. Петровский [и др.] // Речевые технологии. – 2012. – №4. – С. 4-15.
- [3] Айфичер, Э. У. Цифровая обработка сигналов: практический подход, 2-е издание. / Э. С. Айфичер, Б. У. Джерви. – М. : Издательский дом «Вильямс», 2004. – 992 с.
- [4] Upadhyay, N. Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study / N. Upadhyay, A. Karmakar // Eleventh International Multi-Conference on Information Processing, 2015 – P. 574–584.
- [5] Kamath, S. A Multi-Band Spectral Subtraction Method for Enhancing Speech Corrupted by Colored Noise / S. Kamath, P. Loizou // IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002.
- [6] Сергиенко, А. Б. Цифровая обработка сигналов: учеб. Пособие. – 3-е изд. – СПб. : БХВ-Петербург, 2011. – 768 с.
- [7] Королева, И. В. Введение в аудиологию и слухопротезирование / И. В. Королева. – СПб. : КАРО, 2012. – 400 с.
- [8] Wikipedia [электронный ресурс]. – Режим доступа : <https://ru.wikipedia.org/wiki/Шум>
- [9] Wikipedia [электронный ресурс]. – Режим доступа : https://ru.wikipedia.org/wiki/Стационарный_шум
- [10] Wikipedia [электронный ресурс]. – Режим доступа : https://ru.wikipedia.org/wiki/Белый_шум
- [11] Moattar, M. H. A simple but efficient real-time voice activity detection algorithm / M. H. Moattar, M. M. Homayounpour // 17th European Signal Processing Conference, 2009. – P. 2549-2553.
- [12] Петровский, А. А. Цифровая обработка аудио- и видеоданных: пособие / А. А. Петровский, М. И. Вашкевич, И. С. Азаров. – Минск : БГУИР, 2015. – 64 с.
- [13] Вашкевич, М. И. Алгоритм подавления шума и акустической обратной связи на основе спектрального вычитания в слуховом протезе на базе смартфона / М. И. Вашкевич, И. С. Азаров, А. А. Петровский // Доклады БГУИР. – 2016. – №7. – С. 313-317.

[14] MATLAB [электронный ресурс]. – Режим доступа : <https://www.mathworks.com/products/matlab.html> .

[15] Wikipedia [электронный ресурс]. – Режим доступа : <https://ru.wikipedia.org/wiki/MATLAB> .

[16] Экономика проектных решений: методические указания по экономическому обоснованию дипломных проектов : учеб.-метод. пособие / В. Г. Горовой [и др.] – Минск : БГУИР, 2021. – 107 с.

[17] MATLAB [электронный ресурс]. – Режим доступа : <https://www.mathworks.com/pricing-licensing.html> .

[18] ГОСТ 16600-72, Требования к разборчивости речи и методы артикуляционных измерений, Москва. – 1972. – 92 с.

ПРИЛОЖЕНИЕ А

(обязательное)

Отчет о проверке на заимствования



Отчет о проверке на заимствования №1



Автор: Рогова Яна Александровна
Проверяющий:

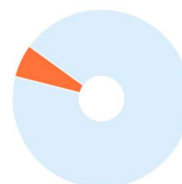
Отчет предоставлен сервисом «Антиплагиат» - <http://users.antiplagiat.ru>

ИНФОРМАЦИЯ О ДОКУМЕНТЕ

№ документа: 8
Начало загрузки: 21.05.2022 19:25:10
Длительность загрузки: 00:00:01
Имя исходного файла: Пояснительная записка.pdf
Название документа: Пояснительная записка
Размер текста: 90 кБ
Тип документа: Дипломный проект
Символов в тексте: 92607
Слов в тексте: 10849
Число предложений: 492

ИНФОРМАЦИЯ ОБ ОТЧЕТЕ

Начало проверки: 21.05.2022 19:25:11
Длительность проверки: 00:00:03
Корректировка от 21.05.2022 19:37:33
Комментарии: не указано
Модули поиска: Интернет Free



ЗАИМСТВОВАНИЯ
6,34%

САМОЦИТИРОВАНИЯ
0%

ЦИТИРОВАНИЯ
0%

ОРИГИНАЛЬНОСТЬ
93,66%

Рисунок А.1 – Отчет о проверке на заимствования

____ Я. А. Рогова

ПРИЛОЖЕНИЕ Б

(обязательное)

Листинг

Б.1 Листинг программы шумоподавления для речевых сигналов

```
% инициализация констант
classdef NamedConst
    properties (Constant)
        N_thr = 5           % Пороговое значение присутствия речи
        Frame_size = 512    % Размер фрейма
        Number_frames = 5   % Фреймы для определения начального шума

% Для спектрального вычитания
        N_bands = 3
        Subtraction_delta = [1 3 2.5 5.5 7.8];
        Subtraction_betta = 0.001;

% Для новой оценки шума
        noise_alpha_p = 0.1;
        noise_alpha_d = 0.99;
        noise_delta = 1;
    end
end

% Выполнение программы
%close all; clear all;

str = 'audiobook';
[input_signal, Fs] = audioread([str '.wav']);
[frame_array, remainder] = signal_to_frames(input_signal);
[noise] = start_noise(frame_array, NamedConst.Number_frames);
[frame_array, segSNR_array, speech] = ...
    processing_frames(frame_array, noise, Fs);
output_signal = synthesis_signal(frame_array, remainder);
audiowrite([str '_out.wav'], output_signal, Fs);
[~, N] = size(frame_array);
in_out_spectrogram(input_signal, output_signal, segSNR_array, N, ...
    Fs, speech);

% разбиение сигнала на фреймы
function [frame_array, remainder] = signal_to_frames(signal)
    frame_size = NamedConst.Frame_size;
    half_frame = frame_size / 2;
    N = fix(length(signal) / half_frame); % деление без остатка
    frame_array = zeros(frame_size, N);
```



```

window = hamming(frame_size);

remainder = mod(length(signal), half_frame);
if remainder == 0
    N = N - 1;
else
    %дополнение сигнала нулями до размера, кратного половине фрейма
    signal = vertcat(signal, zeros(half_frame - remainder, 1));
end

for i = 0:N - 1
    current_index = i * half_frame + 1;
    frame_array(:, i + 1) = signal(current_index:current_index + ...
        frame_size - 1, 1);
end
frame_array = window .* frame_array;
end

% оценка начальной мощности шума
function [noise_abs] = start_noise(frame_array, k)
    [frame_size, ~] = size(frame_array);
    N = frame_size/2 + 1;
    noise_abs = zeros(N, 1);

    for i = 1:k
        current_frame = abs(fft(frame_array(:, i), frame_size));
        noise_abs = noise_abs + current_frame(1:N);
    end

    noise_abs = (noise_abs / k);
end

% обработка сигнала фрейм за фреймом
function [frame_array, segSNR_array, speech] = ...
    processing_frames(frame_array, noise_abs, Fs)
    [frame_size, N] = size(frame_array);
    segSNR_array = zeros(N, 1);
    speech = zeros(N, 1);
    Xmin = abs(fft(frame_array(:, 1)));
    Xmin = Xmin(1:frame_size/2+1);
    Xtmp = Xmin;
    p = zeros(1, frame_size/2+1);
    time_counter = 0;

    %определение фреймов, содержащих речь
    noise_abs2 = noise_abs.^2;
    buf_speech = 0;

```

```

buf_noise = 0;

%обработка фрейм за фреймом
for i = 1:N
    [speech, buf_speech, buf_noise] = voice_activity_detection
        (frame_array, buf_speech,buf_noise, speech, i);

    [frame_array(:, i), new_noise_abs, segSNR_array(i), speech(i), p, ...
        Xmin, Xtmp, time_counter] = processing_frame ...
        (frame_array(:, i), noise_abs, Fs, p, Xmin, Xtmp, ...
        time_counter, i, speech(i));

    noise_abs = new_noise_abs;
    noise_abs2 = noise_abs.^2;
end
end

% детектор речи
function [speech] = voice_activity_detection (frame_array, ...
    buf_speech,buf_noise, speech, i)
    [~, N] = size(frame_array);
    speech = zeros(1, N);
    buf_speech = 0;
    buf_noise = 0;
    k = NamedConst.Number_frames;

    %нормированная краткосрочная энергия
    E_thr = sum(frame_array(:, 1:k), 'all') / N/k;
    %доминантная частота
    F_thr = 0;
    %тональный коэффициент
    SFM_thr = 0;
    for i = 1:k
        s = abs(fft(frame_array(:, i)));
        F_buf_thr = max(s);
        F_thr = max(F_buf_thr, F_thr);
        SFM_thr = SFM_thr + abs(10 * log10((sum(log(s(:))) / N) / ...
            (sum(s(:)) / N)));
    end
    SFM_thr = SFM_thr/k;

    for i = 1:N
        %нормированная краткосрочная энергия
        E_frame = sum(frame_array(:, i)) / N;
        %нормализованная автокорреляционная функция

        %доминантная частота
        S_frame = abs(fft(frame_array(:, i)));
    end

```

```

F_frame = max(S_frame);
%тональный коэффициент
SFM_frame = abs(10 * log10((sum(log(S_frame(:))) / N) / ...
    (sum(S_frame(:)) / N)));
if i == 1
    E_min = E_frame;
    F_min = F_frame;
    SFM_min = SFM_frame;
elseif i <= 30
    E_min = min(E_min, E_frame);
    F_min = min(F_min, F_frame);
    SFM_min = min(SFM_min, SFM_frame);
end
counter = 0;
if E_frame > E_thr*5
    counter = counter + 1;
end
if F_frame > F_thr*1.7
    counter = counter + 1;
end
if SFM_frame > SFM_thr*1.1
    counter = counter + 1;
end
if counter > 0
    buf_speech = buf_speech + 1;
    if buf_speech == 5
        buf_noise = 0;
        speech(i - 4:i) = 1;
    elseif buf_speech > 5
        speech(i) = 1;
    else
        speech(i) = 0;
    end
else
    buf_noise = buf_noise + 1;

    if buf_noise == 10
        buf_speech = 0;
        speech(i - 9:i) = 0;
    elseif buf_noise > 10
        speech(i) = 0;
    else
        if i > k
            speech(i) = 1;
        end
    end
end
end
end
end

```

```

% обработка одного фрейма
function [frame, new_noise_abs, segSNR, speech, p, Xmin, Xtmp, ...
        time_counter] = processing_frame(frame, noise_abs, Fs, p, ...
        Xmin, Xtmp, time_counter, i, speech)
frame_size = NamedConst.Frame_size;
N = frame_size/2 + 1;
fft_frame = fft(frame, frame_size);
theta = angle(fft_frame(1:N));
input_signal_abs2 = abs(fft_frame(1:N)) .^ 2;
input_signal_abs = abs(fft_frame(1:N));
noise_abs2 = noise_abs.^2;

segSNR = snr(input_signal_abs2, noise_abs2);
[speech_abs2] = spectral_subtraction(input_signal_abs2, noise_abs, Fs);
%дополнение симметричной части ДПФ
signal_to_ifft = sqrt(speech_abs2).*(cos(theta)+1j*(sin(theta)));
signal_to_ifft(frame_size/2+2:frame_size) = ...
        flipud(conj(signal_to_ifft(2:frame_size/2)));

frame = real(ifft(signal_to_ifft));
time_counter = time_counter + frame_size / Fs;
% новая оценка мощности шума, если во фрейме отсутствует речь
if (speech == 1)
    new_noise_abs = noise_abs;
else
    [new_noise_abs, time_counter, p, Xmin, Xtmp] =
new_noise(input_signal_abs, noise_abs, p, Xmin, Xtmp, time_counter,
i);
end
end

% вычисление отношения сигнал/шум
function [SNR] = snr(input_signal_abs2, noise_abs2)
    SNR = 10 * log10 (sum(input_signal_abs2) / sum(noise_abs2));
end

% субполосное спектральное вычитание
function [speech_abs2] = spectral_subtraction(input_signal_abs2,
noise_abs, Fs)
    betta = NamedConst.Substraction_betta;
    d = NamedConst.Substraction_delta;

    speech_abs2 = zeros(length(input_signal_abs2), 1);
    bands = bands_borders(Fs);
    a = zeros(NamedConst.N_bands, 1);

```

```

subSNR = zeros(NamedConst.N_bands, 1);

for i = 1:NamedConst.N_bands
    sub_noise_abs2 = noise_abs(bands{i, 1}) .^2;
    subSNR(i) = snr(input_signal_abs2(bands{i, 1}), sub_noise_abs2);

    if (subSNR(i) < -5)
        a(i) = 8;
    elseif (subSNR(i) <= 20)
        a(i) = 5 - 3 / 20 * subSNR(i);
    else
        a(i) = 1;
    end

    speech_abs2(bands{i, 1}) = input_signal_abs2(bands{i, 1}) - ...
        a(i) * d(i) * sub_noise_abs2;
end

z = find(speech_abs2 < 0);
if ~isempty(z)
    speech_abs2(z) = betta*input_signal_abs2(z);
end
end

% разбиение ДПФ на полосы по частоте
function [bands] = bands_borders(Fs)
    fs_fft = (0:Fs/(NamedConst.Frame_size+1):Fs/2)';
    freq = Fs / (2^NamedConst.N_bands);
    bands{1, 1} = 1:length(fs_fft);
    for i = 2:NamedConst.N_bands
        num = find(fs_fft > freq);
        if ~isempty(num)
            bands{i, 1} = num;
            bands{i - 1, 1} = bands{i - 1, 1}(1):bands{i, 1}(1)-1;
        else
            break;
        end
        freq = freq * 2;
    end
end

% новая оценка мощности шума
function [new_noise_abs, time_counter, p, Xmin, Xtmp] = ...
    new_noise(input_signal_abs, prev_noise, prev_p, Xmin, Xtmp, ...
        time_counter, frame_number)

alpha_d = NamedConst.noise_alpha_d;

```

```

delta = NamedConst.noise_delta;
alpha_p = NamedConst.noise_alpha_p;
if time_counter >= 0.5
    time_counter = 0;
    Xmin = min(Xtmp, input_signal_abs);
    Xtmp = input_signal_abs;
else
    Xmin = min(Xmin, input_signal_abs);
    Xtmp = min(Xtmp, input_signal_abs);
end

Sr = input_signal_abs(:) / Xmin(:);
Sr = Sr(:, 1);

if frame_number > 2
    p = alpha_p * prev_p;
    index = find(Sr > delta);
    if ~isempty(index)
        p(index) = p(index) + (1 - alpha_p);
    end
else
    p = prev_p;
end

alpha = alpha_d + (1 - alpha_d) * p;
new_noise_abs = alpha .* prev_noise + (1 - alpha) .* ...
    input_signal_abs;

new_noise_abs = new_noise_abs(:, 1);
end

% синтез сигнала из фреймов
function [signal] = synthesis_signal(frame_array, remainder)
    [frame_size, N] = size(frame_array);
    half_frame = frame_size / 2;
% нормирующий коэффициент
    norm_win = half_frame / sum(hamming(frame_size));
    signal = zeros(half_frame * N, 1);

    signal(1:half_frame) = frame_array(1:half_frame, 1);
    for i = 2:N
        current_index = (i - 1) * half_frame + 1;
        signal(current_index:current_index + half_frame - 1, 1) = ...
            (frame_array(half_frame + 1:frame_size, i - 1) + ...
            frame_array(1:half_frame, i));
    end

% восстановление изначальной длины сигнала

```

```

if remainder == 0
signal = vertcat(signal, frame_array(half_frame + 1:frame_size, N));
else
signal = vertcat(signal, frame_array(half_frame + 1:half_frame + ...
remainder, N));
end

signal = signal * norm_win;
end

```

Б.2 Листинг программы для тестирования системы шумоподавления

```

function [] = in_out_spectrogram(input_signal, output_signal, ...
segSNR_array, N, Fs, speech)
% построение гистограмм

frame_size = NamedConst.Frame_size;
% исходный сигнал
figure;
spectrogram(input_signal, 512, Fs , kaiser(512,7), 475);
set(gca,'Clim', [-65 15]);
graph_labels([" " "Время, с" "Частота, Гц"]);

figure; subplot(211);
stem((1:N)*frame_size/2/Fs, segSNR_array, 'Marker', 'none');
graph_labels([" " "Время, с" "ОСШ, дБ"]);
axis ([frame_size/2/Fs, N*frame_size/2/Fs, -25, 60]);

subplot(212);
stem((1:length(input_signal))/Fs, input_signal, 'Marker', 'none');
graph_labels([" " "Время, с" "Амплитуда, дБ"]);
axis ([frame_size/2/Fs, N*frame_size/2/Fs, -1, 1]);
speech_detector(0.5, N, speech, Fs);

% очищенный сигнал
% получение значений ОСШ для очищенного сигнала
buf = signal_to_frames(output_signal);
[noise] = start_noise(buf, 5);
[~, segSNR_array_out] = processing_frames(buf, noise, Fs);

figure;
spectrogram(output_signal, 512, Fs , kaiser(512,7), 475);
set(gca,'Clim', [-65 15]);
graph_labels([" " "Время, с" "Частота, Гц"]);

figure; subplot(211);
stem((1:N)*frame_size/2/Fs, segSNR_array_out(1:N), 'Marker', ... 'none');
graph_labels([" " "Время, с" "ОСШ, дБ"]);
axis ([frame_size/2/Fs, N*frame_size/2/Fs, -25, 60]);

```

```

subplot(212);
stem((1:length(output_signal))/Fs, output_signal, 'Marker', 'none');
graph_labels([" " "Время, с" "Амплитуда, дБ"]);
axis ([frame_size/2/Fs, N*frame_size/2/Fs, -1, 1]);
speech_detector(0.5, N, speech, Fs);
end

```

```

function [] = graph_labels(str_array)
% функция, настраивающая шрифт и подписи графика
title(str_array(1));
xlabel(str_array(2));
ylabel(str_array(3));
set(gca, 'FontName', 'Times New Roman', 'FontSize', 18);
end

```

```

function [] = speech_detector(height, N, speech, Fs)
% функция, показывающая на графике присутствие речи во фрейме
detector_array = zeros(N, 1);
speech_c = find(speech > 0);
if ~isempty(speech)
    detector_array(speech_c) = height;
end

dbl_detector_array = zeros(2*N-1, 1);
dbl_detector_array(1:2:2*N-1) = detector_array(1:N);
dbl_detector_array(2:2:2*N-1) = detector_array(2:N);

hold on;
stairs((1:0.5:N)*NamedConst.Frame_size/2/Fs, dbl_detector_array, ...
'LineWidth', 1);
end

```