

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»**
(**Н И У « Б е л Г У »**)

ИНСТИТУТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ И ЕСТЕСТВЕННЫХ НАУК

КАФЕДРА ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ
СИСТЕМ И ТЕХНОЛОГИЙ

ВЛИЯНИЕ ЭМОЦИЙ НА ХАРАКТЕРИСТИКИ РЕЧЕВОГО СИГНАЛА

Выпускная квалификационная работа
обучающегося по направлению подготовки
11.04.02 Инфокоммуникационные технологии и системы связи,
магистерская программа «Системы и устройства радиотехники и связи»
очной формы обучения, группы 07001636
Дорецкого Игоря Александровича

Научный руководитель
кандидат технических наук, доцент,
доцент кафедры ИТСиТ Прохоренко Е.И.

Рецензент
Начальник отдела серийного
сопровождения АО НПП «Спец-Радио»,
кандидат физико-математических наук
Туяков С. В.

БЕЛГОРОД 2018

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	3
ГЛАВА 1 ОПИСАНИЕ ХАРАКТЕРИСТИК И ПРИЗНАКОВ РЕЧЕВОГО СИГНАЛА.....	5
1.1 Характеристики речевого сигнала	5
1.2 Признаки ЭОР	10
1.2.1 Спектрально-временные признаки	10
1.2.2 Кепстральные коэффициенты.....	12
1.2.3 Амплитудно-частотные признаки.....	15
1.2.4 Признаки нелинейной динамики.....	17
ГЛАВА 2 ЛИНЕЙНЫЕ ХАРАКТЕРИСТИКИ РЕЧИ ХАРАКТЕРИЗУЮЩИЕ ЭМОЦИОНАЛЬНОЕ СОСТОЯНИЕ.....	20
2.1 Линейные характеристики речи	20
2.1.1 Частоты основного тона.....	21
2.1.2 Громкость речи.....	24
2.1.3 Паузы.....	24
ГЛАВА 3 ИССЛЕДОВАНИЕ ЛИНЕЙНЫХ ХАРАКТЕРИСТИК ОЦЕНИВАЮЩИХ ЭМОЦИОНАЛЬНУЮ ОКРАСКУ ЗВУКОВ	26
3.1 Подготовка речевых данных.....	26
3.2 Исследование параметра частоты основного тона.....	32
3.3 Исследование количества пауз	39
3.4 Исследование громкости речевого сигнала	41
ЗАКЛЮЧЕНИЕ	44
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	46

ВВЕДЕНИЕ

Существуют системы, для которых имеет значение эмоциональный окрас звукового сигнала.

Эмоционально окрашенная речь (ЭОР) применяется в многочисленных сферах человеческой деятельности и играет роль в функциях систем автоматизированного управления федеральной службы безопасности (ФСБ), реабилитации пациента и срочного оповещения для предотвращения неприятных ситуаций, а возможно даже катастрофических. Уже давно есть интерес к анализу речевого сигнала (РС) как действительного показателя эмоционального состояния (ЭС) человека, выполняющего ответственную роль космонавта, летчика, оператора АЭС, диспетчера аэропорта, психолога, сотрудника спецслужб и т.д. Проблема автоматического распознавания эмоционального состояния звуков на данный момент не решена полностью, в связи с отсутствием модели описания речевых образцов в условиях проявления различных видов эмоций.[3] ЭОР должна отражать взаимосвязь вида эмоций и объективных характеристик РС. На настоящий момент времени определение такой взаимосвязи вызывает затруднение у большинства исследователей в этой области. Существует несколько определений понятия эмоция прямо или косвенно относящихся к определению эмоционального окраса звуков речи:

Эмоции – одна из форм отражения. Эмоции отражают отношения к потребностям, целям и мотивам деятельности человека, переживающего эти эмоции.

Процесс, который оценивает внешнюю и внутреннюю ситуацию жизнедеятельности в форме переживаний называется эмоцией.

Процесс отражающий оценочное отношение к существующей ситуации, который напрямую связан с понятием речевой сигнал называется эмоцией (с точки зрения речевого сигнала).

Речевой сигнал содержит признаки, характеризующие эмоциональные и физиологические аспекты функциональных состояний.

Классифицируют эмоции на две категории:

Категория положительных эмоций (веселье, радость, смех и т. д.)

Категория отрицательных эмоций (грусть, печаль, угнетение и т. д.)

Соотнесение распознаваемого речевого сигнала с базой звуков, которые необходимо распознать, проходит в три этапа:

- 1) выделение признака объекта;
- 2) объединение признаков в группы;
- 3) выбор предполагаемого значения из ряда вариантов.

Актуальностью данной работы является использование характеристик для разработки алгоритма определения эмоционального состояния человека. Это является важной частью для многих сфер деятельности человека, и представляет собой востребованную функцию в современных системах имеющих автоматизированное управление, реабилитацию, срочное оповещение.

Целью работы является исследование характеристик, зависящих от эмоционального состояния звуков.

Задачи исследования:

Проанализировать характеристики и признаки речевого сигнала;

Изучить линейные характеристики речи характеризующие эмоциональное состояние;

Исследовать линейные характеристики оценивающие эмоциональную окраску звуков.

ГЛАВА 1 ОПИСАНИЕ ХАРАКТЕРИСТИК И ПРАЗНАКОВ РЕЧЕВОГО СИГНАЛА

1.1 Характеристики речевого сигнала

Существует множество характеристик ЭОР. Для анализа влияния эмоций на характеристики речевого сигнала следует знать определения:

1. Звуковое давление
2. Частота основного тона
3. Спектральная плотность
4. Белый шум
5. Розовый шум
6. Высота звука
7. Динамический диапазон
8. Пик-фактор
9. Форманты
10. Слитность звучания
11. Прямое и диффузное звучание
12. Индекс направленности слуха

Ниже приведены определения и описания характеристик речевого сигнала касемо или прямо относящихся к ЭОР.[12]

Сила, с которой звуковая волна, оказывает давление на одну единицу площади поверхности, расположенной перпендикулярно к губам воспроизводящего на расстоянии один метр от него-называется звуковым давлением речи. Уровни звукового давления речи находятся в интервале 0-65 дБ для негромкой речи и 0-80 дБ для речи с высокой громкостью. С увеличением дистанции от источника звука значение уровня звукового давления речи становится меньше. К примеру, удаление от источника звука на расстояние в два раза большее, позволяет уменьшить уровень на 6 дБ, в 4 раза – на 12 дБ, в 8 раз – на 18 дБ и т.д.

Частота основного тона. Основная часть звуков речи формируется с участием голоса. С работой голосовых связок связано голосообразование, которое вызывает периодические изменения площади голосовой щели. Так как голосовые связки обладают определенными физиологическими особенностями, то для их размыкания и смыкания требуется определенное время. Периодом колебаний называется участок времени, определяющий частоту основного тона звуков речи. Частота основного тона в свою очередь определяет высоту голоса. Эта частота для всех звуков определяется в диапазоне частот 70-450 Гц. Небольшие отличия характерны для мужского и женских голосов. Частота женского голоса несколько выше мужского, это связано с физиологическими особенностями организма.

При воспроизведении слов частота основного тона изменяется, поскольку человек произносит слова с ударением, подчеркивает звуки интонацией, а также проявляет эмоции. Таким образом, само изменение частоты основного тона называют интонацией. Каждый человек обладает индивидуальным диапазоном изменения частоты основного тона, а, следовательно, и интонация. Все эти параметры (частота основного тона, интонация, окраска голоса) могут служить для идентификации личности человека в различных автоматических системах. Частота основного тона оказывает непосредственное влияние так же на спектральный состав голоса конкретного диктора.

Спектральная плотность. Речевой сигнал имеет шумоподобную природу. Спектральной плотностью интенсивности речевого сигнала называют отношение средней интенсивности сигнала в заданной полосе частот ΔI_{cp} к ширине этой полосы Δf : [9]

$$W = \Delta I_{cp} / \Delta f \quad (1.1)$$

Спектральная плотность может быть измерена в Вт/(м*Гц) (ватт на метр квадратный на герц) и равна интенсивности шума в диапазоне частот полосой 1 Гц.[25]

«Белый шум» - это случайный процесс, спектр которого равномерно распределен в диапазоне от нуля до бесконечности, таким образом, его спектральная плотность которого не зависит от частоты. При решении практических задач необходимо, чтобы это требование выдерживалось в полосе частот слуховой системы человека.

Кроме «белого шума» существует еще ряд «цветных» шумов.

Высота голоса – это восприятие звука, оцениваемая субъективно и зависит от частоты. За объективную единицу измерения этого параметра, принята октава. Октавы характеризуются двукратным отношением частот – 1, 2, 4, 8, 16, 32 и т.д. При практическом использовании октава может делиться на полуоктавы и третьоктавы.[12]

Динамический диапазон. В время воспроизведения речевого звукового сообщения, акустический уровень звукового сигнала изменяется. Уровнеграммой называется зависимость уровня сигнала от времени. Динамический диапазон-это разность максимального и минимального уровня сигнала:

$$D = L_{\max} - L_{\min} \quad (1.2)$$

Динамический диапазон звуков речи 25-35 дБ, а телефонных разговоров от 35 до 45 дБ.[

Пик-фактор представляет собой разность максимального и среднего уровня звукового сигнала:

$$\Pi = L_{\max} - L_{\text{ср}} \quad (1.3)$$

Форманты – это участки скопления энергии в звуковом диапазоне частот. Форманты полностью заполняют весь диапазон частот речи 125-8000 Гц. В некоторых диапазонах частот меняется частота появления формант, зависящая от количества повторения звуковых сигналов. Форманты имеют свою часть информации о звуковом сигнале, и эти части не имеют зависимости друг от друга. Исходя из этого, можно суммировать вероятности появления формант. Спектр гласных можно определить 2-3 формантами и имеет диапазоны: 1)300-1000 Гц, 2)90-2300 Гц, 4)2200-2500 Гц. Спектр согласных имеет одну форманты и определяется минимумом.[15]

Слитность звучания. Громкость звука пропадает не сразу, а плавно снижаясь до нулевого значения. Длительность значения снижения громкости звука в среднем составляет 150-200 мс. Поэтому наблюдается слитность восприятия звуков, запаздывающих друг относительно друга. Для того что бы наблюдалась слитность восприятия двух звуков нужно, чтобы следующий звук запаздывал относительно предыдущего на отрезок времени не более 50 мс. Так же при запаздывании с большим значением слитность звучания может не нарушаться, если следующий звук имеет уровень ниже первого. Звуки, запаздывающие на 60 мс и менее, полностью суммируются с основным звуком, а запаздывающие более чем на 60 мс, являются помехой. Оба звука воспринимаются отдельно если интервал запаздывания больше, так как интервал первого звука становится малым и не покрывает второй.

Прямое и диффузное звучание. Звук распространяется прямолинейно, до того момента пока не отразится от поверхности, которой он достигнет. Данный звук называют прямым звуком, а звуковым полем, в котором он распространяется называют свободным звуковым полем. Диффузное звучание-звук имеющий одинаковую звуковую энергию в каждой точке и многократно отражающий от поверхности.

Индекс направленности слуха -это единица измерения, которая измеряет повышение звукового давления и измеряется в децибелах (дБ). При

падении звуковой волны под 90 градусов от уха имеет место быть отражению ее волны и ее дифракция. Отношение между отраженной волны и ее дифракцией, зависит от отношения длины звуковых волн и плоскостью с которой происходит взаимодействие. Так как волна отражается, то звуковое давление у уха повышается. Повышение составляет 1-6 дБ. В случае падения звуковой волны спереди-явление отражения почти не сказывается. Когда звуковая волна падает под различными углами, то на низких частотах давление возле плоскости, на которую оно падает, равно звуковому давлению диффузного поля. При высоких частотах это давление удваивается.

1.2 Признаки ЭОР

Основной задачей получения признаков эмоционально окрашенной речи является, преобразование звуковой волны в такое признаковое пространство, в котором большинство объектов одного класса будет в одной группе, а объекты похожих классов отделено. Существует 5 групп признаков, которые определяют речевые сигналы:

1. спектральные
2. временные
3. кепстральные
4. АЧХ (амплитудно-частотные)
5. признаки не линейной динамики

1.2.1 Спектрально-временные признаки

Спектрально-временные признаки характеризуют звуки речи с помощью различных характеристик и участков речевого сигнала имеющих:

1. основной тон;
2. шум;
3. паузы;

Спектральные признаки речевого сигнала:

1. Относительное время нахождения сигнала в полосах спектра;
2. Нормализованное время нахождения сигнала в полосах спектра;
3. Медианное значение спектра речи в полосах;
4. Относительная мощность спектра речи в полосах;
5. Вариация огибающих спектра речи;
6. Нормализованные величины вариации огибающих спектра речи;
7. Коэффициенты кросскорреляции спектральных огибающих между полосами спектра;[1]

Временные признаки включают в себя:

1. Длительность сегмента, фонемы;
2. Высота сегмента;
3. Коэффициент формы сегмента.[27]

Спектрально-временные признаки отражают различные формы временного ряда и спектры голосовых импульсов у разных звуков и особенности фильтрации речевых трактов. Особенность речевого потока, связанная с перестройкой органов речи, отражает синхронность движения этих органов, и является интегральной характеристикой.

Речевой сигнал рассматривается как процесс неизменный процесс с очень коротким промежутком времени в группе спектрально-временных признаков. Были выделены параметры, относящиеся к повышению уровня сигнала. Они описывают статистические характеристики речевого сигнала, основного тона и особенность спектра. РС имеет последовательность значений энергетических спектров имеющих короткий промежуток времени. Измерения были проведены в моменты времени $j = 0, 1, \dots, J$, каждые пмс. Речевой сигнал с помощью быстрого преобразования Фурье (БПФ)

подвергается спектральному анализу. Спектры вычисляются с помощью набора 24 фильтров, соответствующим 24 критическим полосам. Таким образом, РС может быть представлен в виде

$$\{x(0, j), \dots, x(i, j), \dots, x(23, j)\}, j = 0, 1, \dots, J \quad (1.4)$$

где $x(i, j)$ – значение сигнала на выходе i -го полосового фильтра в j -м кратковременном энергетическом спектре; J – общее количество спектральных срезов на анализируемом отрезке.

В результате проведенного исследования была экспериментально доказана взаимосвязь временных параметров модели, а именно длительности сегментов l , их отношений dl и средних значений длительности сегментов l со степенью эмоционального окраса речевого сигнала по группе спектрально-временных признаков. Это дает возможность использовать временные характеристики РС в качестве признаков, которые позволяют распознавать ЭОР. Доказано наличие взаимосвязи между объективными количественными параметрами РС и субъективной (воспринимаемой слушателями) качественной оценкой образцов речи. Применять для распознавания и идентификации эмоций только спектральные характеристики.

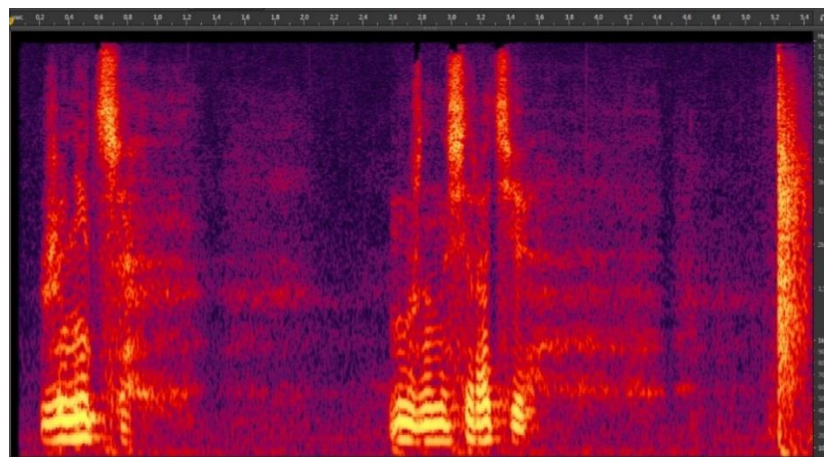


Рисунок 1.1- Спектр звукового сигнала

1.2.2 Кепстральные коэффициенты

Включают в себя:

Мел-частотные кепстральные коэффициенты;

Коэффициенты линейного предсказания с коррекцией на неравномерность чувствительности человеческого уха;

Коэффициенты мощности частоты регистрации;

Коэффициенты спектра линейного предсказания;

Коэффициенты кепстра линейного предсказания.

В настоящее время современные автоматические системы синтеза и распознавания речи усложняются на извлечении такой характеристики как частотная. Это можно объяснить таким образом, что коэффициенты первой модели наделяются лучшей делимостью звуков. Для того чтобы отделить сигналы возбуждения и речевого тракта, прибегают к кепстральному анализу. Схематически этот метод представлен на рисунке 1.2:



Рисунок 1.2 - Схема кепстрального анализа

Общая схема кепстрального анализа сигнала:

FFT–блок быстрого преобразования Фурье сигнала(БПФ); LOG-блок логарифмирования спектра;

IFFT–блок обратного быстрого преобразования Фурье (ОБПФ)

Одним из наиболее эффективных методов анализа речевого сигнала является линейное предсказание. Этот метод является одним из главных при оценке основных параметров РС: период основного тона, форманты, спектр, функция площади речевого тракта. Сокращенном представлении речи с

целью ее передачи и хранения, так же играет роль в оценке этих параметров. Этот метод позволяет получить довольно точную оценку, при его простоте вычисления. Основным принцип линейного предсказания состоит в том что бы была возможность аппроксимировать линейной комбинацией текущие отсчеты речевого сигнала. Коэффициенты вычисляются минимизацией среднего квадрата разности между отсчетами РС и их предсказанными значениями.

Мел-частотные кепстральные коэффициенты (MFCC) широко используются в качестве набора признаков РС. Данный метод подразумевает собой максимальное приближение информации поступающей в головной мозг слушателя. Признаки, построенные на основе MFCC, учитывают психологические и акустические принципы восприятия речи, поскольку используют мел-шкалу, связанную с полосами слуха (критические полосы). Данные признаки вычисляются следующим образом:

$$MFCC = \sum_{l=1}^L E_1 \cos(k(l - 0,5) \frac{\pi}{l}), 1 \leq k \leq L, \quad (1,5)$$

Где, $S(k)$ -спектр Фурье; k_1, k_2 -границы частотных диапазонов l -й мелко частотной полосы; $\Delta K = K_2 - K_1$ – четное число; $\omega(x)$ – оконная функция; L – количество мел-полос.

E_1 Вычисляется по формуле:

$$E_1 = \lg \left(\sum (S(k))^2 \omega \left(k - \left(kl_l + \frac{\Delta K_l}{2} \right) \right) \right) \quad (1.6)$$

Все перечисленные кепстральные коэффициенты позволяют уменьшить размер признакового пространства. Это оказывает влияние на быстроту вычисления параметров ЭОР.

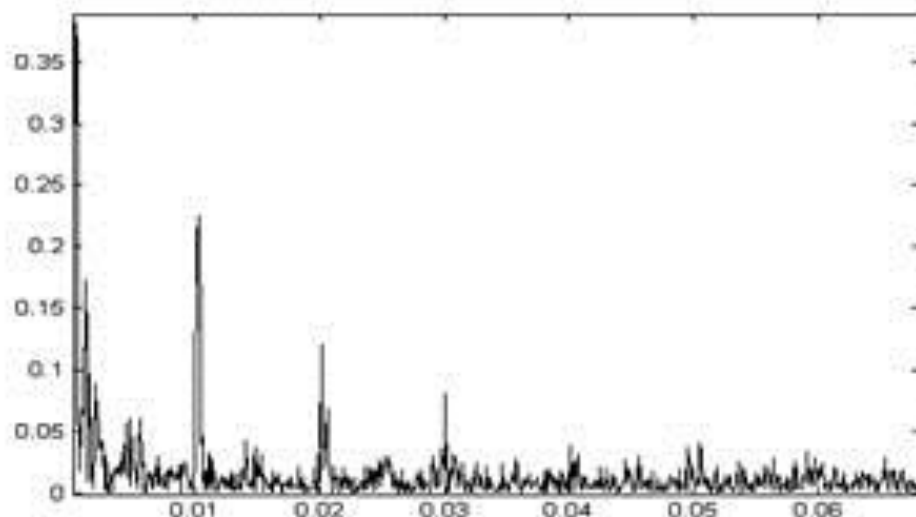


Рисунок 1.3 - Кепстр вокализованного сигнала

1.2.3 Амплитудно-частотные признаки

Амплитудно-частотные признаки оценивают значения, которые меняются в зависимости от параметров ДПФ (вида и ширины окна), а так же при незначительных сдвигах окна по выборке, и включают в себя такие понятия как:

1. Интенсивность, амплитуда
2. Энергия
3. Частота основного тона (ЧОТ)
4. Формантные частоты
5. Частотная модуляция основного тона (шумовой параметр);
6. Амплитудная модуляция на основном тоне (шумовой параметр);
7. Радиальная базисная ядерная функция

8. Нелинейный оператор

РС представляет собой звуковые колебания, которые характеризуют частоту (число колебаний в секунду), интенсивность (амплитуда колебаний) и длительности (участок анализируемого звукового сигнала). Данные характеристики фиксируются и измеряются с помощью специальных электронно-акустических приборов, таких как осциллограф и спектрограф. С помощью амплитудно-частотных признаков возможно определить достаточное количество информации для человека используя минимальное количество времени для восприятия.

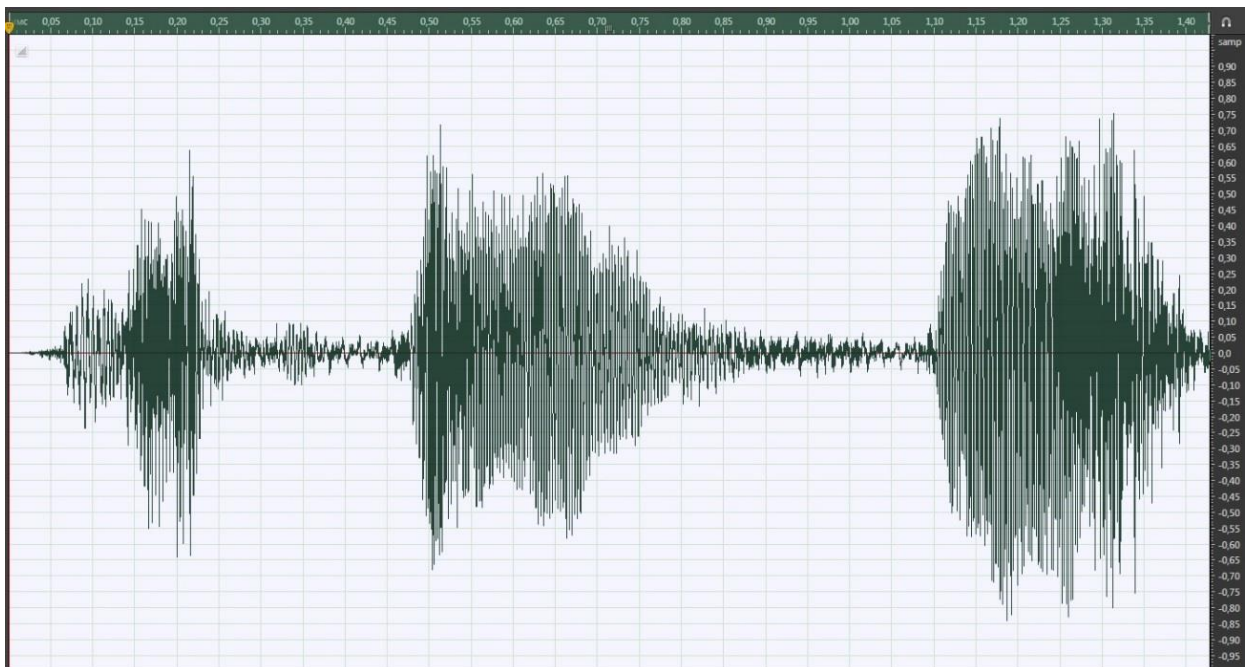


Рисунок 1.4 - Осциллограмма звукового сигнала

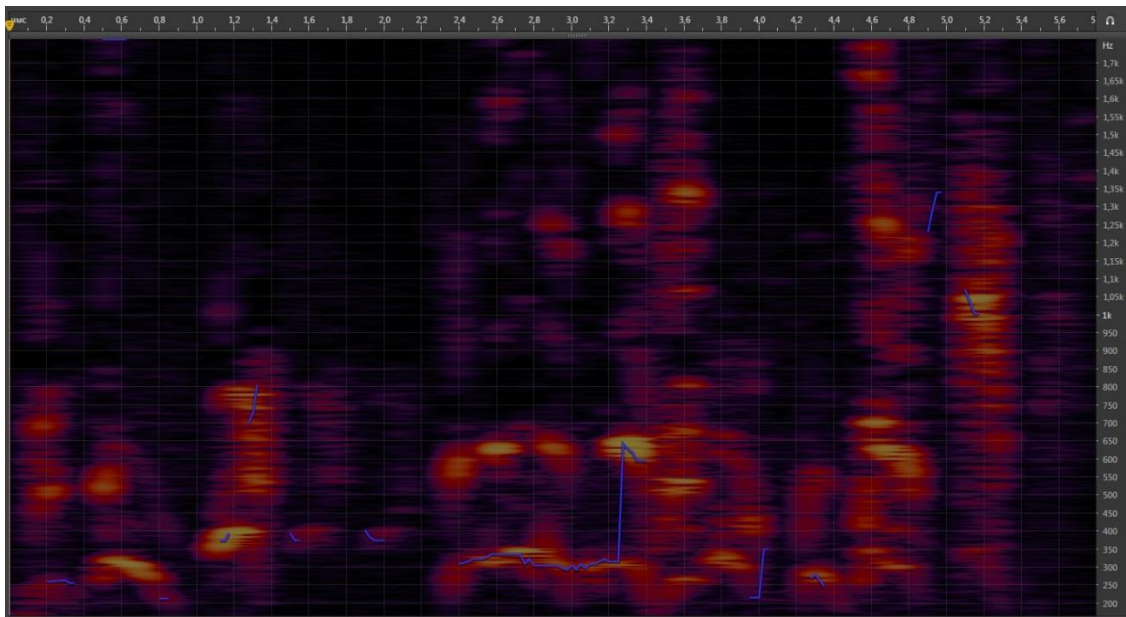


Рисунок 1.5 - Участки основного тона звукового сигнала

1.2.4 Признаки нелинейной динамики

В образовании речи важную роль играют процессы, которые называются нелинейными. Образование речи считается нелинейным, если его анализировать методами нелинейной динамики.

Исследование базовых математических моделей и реальных систем относится к задачам нелинейной динамики. Данные задачи составляют системы отдельных элементов и законы взаимодействия между ними.

В настоящее время математическая теория основывается на методах нелинейной динамики, которая представляет собой теорему Такенса, подводящая математическую модель под идеи нелинейной авторегрессии и предоставляет возможность восстановления фазового портрета аттрактора по одной координате или временному ряду. Аттрактор-большое количество точек или подпространство в фазовом пространстве, к которому приближается фазовая траектория после затухания переходных процессов.

При рассмотрении оценки характеристик сигнала используются модели временного ряда.

Это позволяет распознать и правильно произвести идентификацию различных эмоций в ЭОР. Выявленные отличия используются в форме аттракторов.

Точность нелинейных систем зависит от величины внешних воздействий, что нельзя сказать про линейные системы. Еще существует зависимость по отношению к устойчивости и качества переходных процессов в нелинейных системах, имеющая такую же зависимость как в линейных и нелинейных системах. Точность, устойчивость и качество переходных процессов не зависят от величины внешних воздействий у линейных систем. Благодаря применимости суперпозиции к линейным системам, пропорциональное абсолютное отклонение выходной величины, изменяет величину внешнего воздействия, не изменяя его качество и длительность.

Устойчивость линейных вне зависимости от внешних воздействий, остается неизменной. Она характеризуется только собственными параметрами. Линейная система является устойчивой, когда она устойчива в любом установившемся режиме и основа любого процесса всегда будет сходиться.

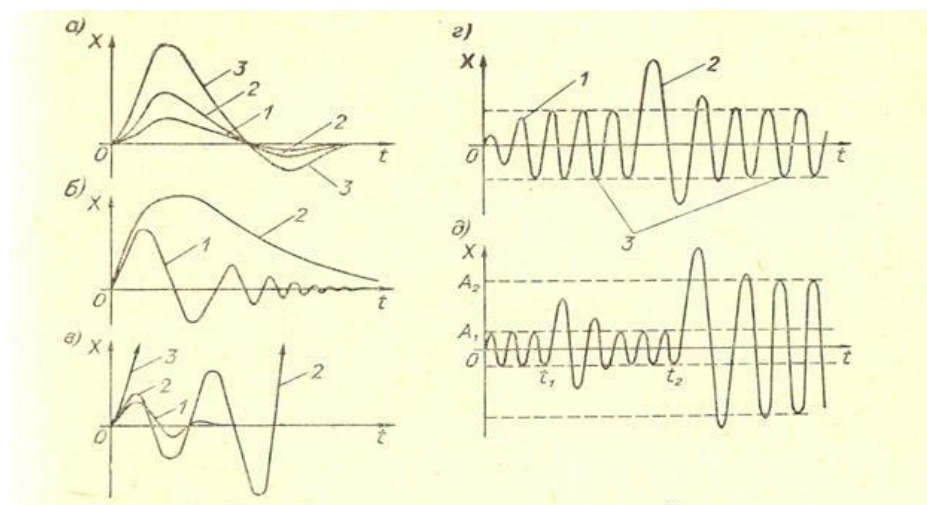


Рисунок 1.6 - Динамика нелинейных систем

В нелинейных системах дело обстоит иначе. При изменении величины переходных воздействий качество переходных процессов изменяется. В нелинейных системах период колебаний процессов не постоянный. По мере изменения отклонения этот период уменьшается с уменьшением отклонения. Кроме того, нелинейная система, устойчива или неустойчива от значений внешнего воздействия. Она может оказаться неустойчивой при возникновении расходящегося переходного процесса при других значениях этого воздействия. Можно сделать вывод, что в установившемся состоянии в системе будут существовать незатухающие колебания, у которых амплитуда ограничена зоной нечувствительности, несколько превышая ее.

Автоколебания-это собственные колебания, имеющие постоянную амплитуду, нелинейностями системы. Автоколебания имеют новый вид, и представляют собой установившийся режим, с отсутствием внешних воздействий, наряду со статическим установившимся режимом и характерен только для нелинейных систем. Собственные незатухающие колебания находятся на границе устойчивости в линейных системах. Амплитуда этих колебаний зависит от внешних воздействий. При таких условиях данная линейная система будет являться неработоспособной.

ГЛАВА 2 ЛИНЕЙНЫЕ ХАРАКТЕРИСТИКИ РЕЧИ

ХАРАКТЕРИЗУЮЩИЕ ЭМОЦИОНАЛЬНОЕ СОСТОЯНИЕ

Для описания эмоций, существует такой термин, который имеет название- базовая эмоция. Существуют различные определения базовых эмоций, предложенных множеством известных ученых, и их объединяет выделение ограниченного количества эмоций. Математической составляющей является определение базиса векторного пространства.

Количество базовых эмоций в ранних исследованиях составляет меньше 10, в современных это число насчитывает 0-20. Как пример можно привести 10 базовых эмоций: радость, интерес, удивление, печаль, гнев, отвращение, презрение, страх, стыд, вина. В качестве основных критериев базовых эмоций приводятся следующие:

Базовая эмоция определяется мимикой лица;

Базовая эмоция осознается человеком;

Базовые эмоции появились в результате эволюции и течения биологических процессов;

Базовая эмоция служит адаптации человека;

Исследования в области эмоциональной речи строятся таким образом, что основные эмоции являются частью базовых, и это все строится различным образом.

2.1 Линейные характеристики речи

Комплексной задачей, характеризующих эмоциональное состояние, является идентификация. Речь является составляющей этой системы. Большинство акустических характеристик содержат в себе большое количество информации, относящейся к эмоциональному состоянию.

Известно, что некоторые параметры, имеющие огромное значение для определения эмоционального окраса, до сих пор не изучены. Это связано, с технической частью, а не теоретической.

2.1.1 Частота основного тона

Гласные и согласные звуков речи в голосе человека определяется частотой колебания, задающей основной тон (ОТ) звуков речи. Для каждого человека распределение основного тона по частотному диапазону индивидуально. Характеристикой основного тона является усредненный показатель интервала значения частоты основного тона.

Измерять частоту основного тона можно с помощью осциллограмм, используя высококачественные и довольно точные приборы для измерения. С вероятностью 0,95 основной тон мужских голосов имеет частоту от 97 до 195 Гц. Основной тон женских голосов имеет интервал 195-320 Гц.

Были проведены исследования в получении распределения длительности мужских и женских звуков, вокализованных и невокализованных сигналов. По результатам исследования было выявлено, что средняя длительность невокализованных участков шумового сигнала составляет 65 мс при максимальной длительности 160 мс. Средняя длительность гласных составляет 180 мс, согласных-95 мс, слога дикторской речи-260 мс.

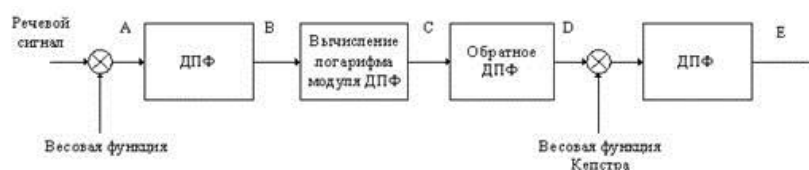


Рисунок 2.1 - Вычисление частоты основного тона

Периодические колебания связок называют периодом основного тона. Данная величина имеет индивидуальные параметры голоса воспроизводящего. Меняется она в зависимости от эмоциональной окраски речи. Примерное значение частоты основного тона человека находится в диапазоне 80 – 400 Гц. Существуют различные методы оценки основного тона, которые можно разделить на 3 группы:

- амплитудный;
- корреляционный;
- частотный;

Можно описать более подробно каждую группу алгоритмов.

Амплитудная селекция

На выбранном диапазоне речевого звука при довольно малом уровне шумов форма речевого колебания, как правило точно повторяется на каждом очередном периоде основного тона. Расстояние между максимумами звуков речи можно посчитать равным периоду основного тона. Трудностью этого алгоритма составляет поиск локальных ложных максимумов. Этого можно добиться за счет увеличения поиска максимума в схеме поиска, повысив его порог срабатывания. Повысить надежность определения периода основного тона в данной схеме можно, если добавить второй канал амплитудной селекции, определяющий и выделяющий положение минимумов речевого сигнала. Главным достоинством устройств временной селекции является чрезвычайная простота реализации. Основным недостатком является низкая точность и устойчивость определения основного тона звуков речи.

Корреляционные методы определения периода основного тона

Представим что речевой сигнал имеет вид последовательности отсчетов S_i , $i = 0, 1, 2, \dots$. Предположим, что энергия речевого сигнала не меняется на участке квазистационарности. Тогда оценка периода основного тона должна максимизировать корреляционную функцию. Это дает возможность повысить оценку определения основного тона по отношению к

методу временной селекции. При этом следует отметить значительную вычислительную сложность данного алгоритма. Существуют его модифицированные аналоги, основополагающей которой является взаимная корреляционная функция. Если функция разделяет речевой сигнал на три уровня: $\{-1, 0, 1\}$, можно построить вычислитель взаимной корреляционной функции без умножителя. Корреляционные методы оценивания периода основного тона имеют общий недостаток: При модулировании сигнала по амплитуде наблюдается неустойчивая работа. Энергия эмоционально окрашенной речи меняется и на квазистационарных участках, соответствующих одной фонеме речевого сигнала. Параметр имеет смысл коэффициента усиления, когда модифицируем целевую функцию. Этот метод позволяет получить довольно точную оценку основного тона, которая изменяется во времени при изменении голоса человека.

Частотная селекция

При воспроизведении звукового сигнала, в его спектре присутствуют пики на частотах, кратных частоте основного тона. Если построить дискретное преобразование Фурье с достаточно малым шагом дискретизации по частоте, то можно рассмотреть это в качестве оценки частоты основного тона. Эта частота, имеет максимум энергии спектра. Поиск максимума производится в интервале 80 – 400 Гц. Довольно часто появляется такая ситуация, когда в заданном диапазоне появляется вторая гармоника, возможно даже с большей энергией. И тогда она принимается как основной тон ошибочно. Представляется это как функция сжатия по частотной характеристике логарифмов спектра мощности. Суть идеи состоит в том, что для истинной частоты основного тона вторая гармоника второго слагаемого сложится с первой гармоникой первого слагаемого и усилит ее. Аналогией будет для третьего слагаемого и т. д. В результате для звукового сигнала появится ярко выраженный пик функции на частоте основного тона. Для

невокализованного звука суммирование будет иметь не упорядоченный характер.

2.1.2 Громкость речи

Громкость – единица измерения, оценивающая амплитуду сигнала, его интенсивность, по слуховому восприятию. Измеряется в децибелах (дБ). Обычным уровнем громкости (нормальный) является величина равная 50-80 дБ (при постоянном фоновом шуме до 10 дБ). Различают уровень громкости на такие значения как: умеренное повышение (80-90 дБ), значительное повышение (90-110 дБ), крик (выше 110 дБ), умеренное понижение (40-50 дБ), значительное снижение (20-40 дБ) и шёпот (менее 20 дБ).

Громкость оказывает влияние на восприятие звуков и их различия. Эта способность и является главным объектом исследований. Большое внимание уделено для исследования интенсивности сигнала, с помощью которой измеряются различные характеристики речевого сигнала, такие как амплитуда, среднеквадратичное отклонение и т.д. Наибольшее распространение получил метод оценки «абсолютной громкости».

2.1.3 Паузы

Паузой – является перерыв в звуковых сообщениях. Разделить паузы можно на 3 группы: 1)короткие - до 3 с; 2) средние - 3-7 с; 3) длинные - более 7 с. Важно учитывать участки высказывания, в котором выявлена пауза, так как она может определить реальное эмоциональное состояние. Количество

пауз и их длительность служит косвенным указателем на депрессию, диссимуляцию переживаний и т.д. Но иногда она может передавать противоположные оценки окраски речи, поэтому нужно ориентироваться еще и на контекст.

Проводились исследования, в которых использовались паузы для определения трех эмоциональных окрасов звуков речи: нормального, депрессивного и маниакального. Использовались 10 переменных, которые измеряются в единицах времени (мс): общее время разговора, общее время пауз, общее время фрагмента речи (мс), скорость речи (фонем/с), скорость артикуляции (фонем/с), задержка сигнала (мс), средняя длина паузы (мс), длина сегмента, длина паузы и максимальный уровень речи. По результатам измерения этих параметров, можно дать довольно точную оценку окраски речевого сигнала по трём состояниям (нормальное, депрессивное, маниакальное).

ГЛАВА 3 ИССЛЕДОВАНИЕ ЛИНЕЙНЫХ ХАРАКТЕРИСТИК ОЦЕНИВАЮЩИХ ЭМОЦИОНАЛЬНУЮ ОКРАСКУ ЗВУКОВ

В данной работе проводились исследования зависимости линейных характеристик эмоциональной окраски голоса. Будет использоваться речевая база, состоящая из различных звуковых файлов с положительной и отрицательной эмоциональной окраской, а так же нейтрального состояния звуков речи, для определения нормальной. Исследования проводились в несколько этапов. Начальным этапом является подготовка речевых данных с помощью программного пакета adobe audition. Последующими этапами являются анализ характеристик, а так же их исследования.

3.1 Подготовка речевых данных

Подготовка речевых данных заключается в записи голосов с разными эмоциональными окрасками голоса диктора. Для этого используем программный пакет adobe audition.

Для анализа использовались 8 звуковых файлов-4 с положительной эмоциональной окраской и 4 с отрицательной (рис. 3.1-3.8)

Данные для сигналов с положительной эмоциональной окраской представлены в таблице 3.1:

Таблица 3.1 - Звуковые файлы и их параметры положительной окраски звукового файла

Звуковой файл	Длительность, с	Разрядность, бит	Частота дискретизации, Гц
1	7,674	32	22050
2	5,332	32	44100
3	7,972	32	44100
4	3,870	32	44100

Данные для сигнала с отрицательной эмоциональной окраской представлены в таблице 3.2:

Таблица 3.2 - звуковые файлы и их параметры отрицательной окраски звукового файла

Звуковой файл	Длительность, с	Разрядность, бит	Частота дискретизации, Гц
1	7,540	32	22050
2	6,496	32	16000
3	4,759	32	44100
4	7,050	32	24000



Рисунок 3.1 - Звуковой сигнал звукового файла 1 с положительной эмоциональной окраской

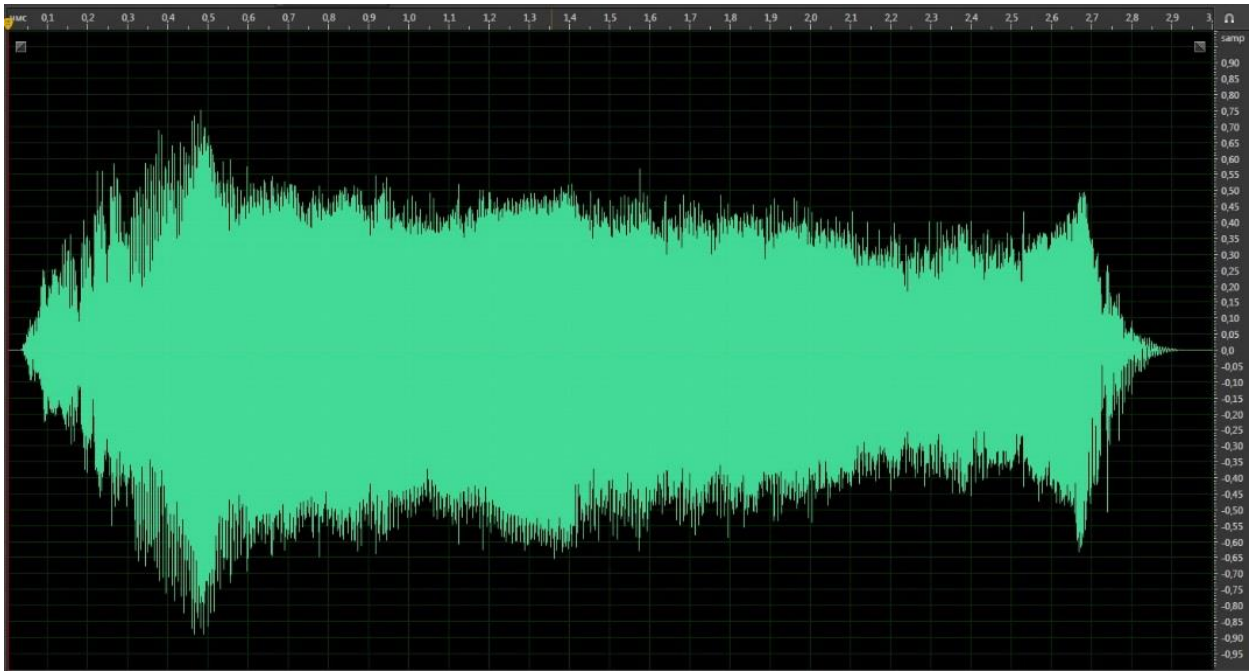


Рисунок 3.2 - Звуковой сигнал звукового файла 2 с положительной эмоциональной окраской



Рисунок 3.3 - Звуковой сигнал звукового файла 3 с положительной эмоциональной окраской



Рисунок 3.4 - Звуковой сигнал звукового файла 4 с положительной эмоциональной окраской



Рисунок 3.5 - Звуковой сигнал звукового файла 1 с отрицательной эмоциональной окраской

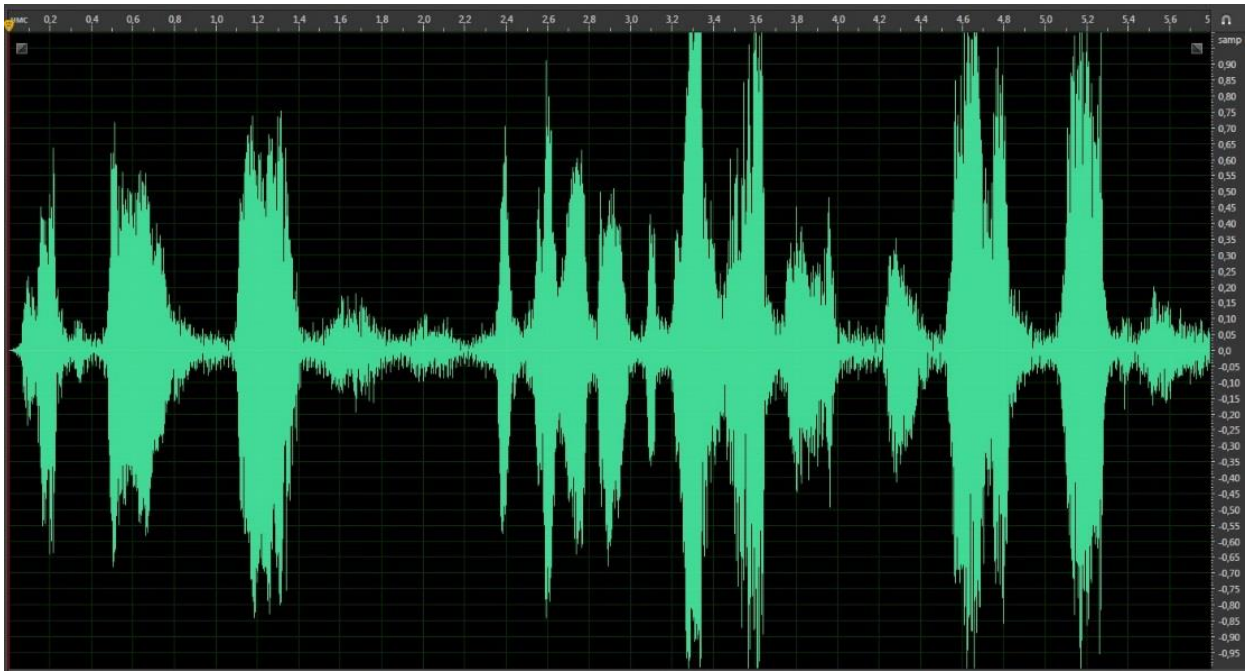


Рисунок 3.6 - Звуковой сигнал звукового файла 2 с отрицательной эмоциональной окраской

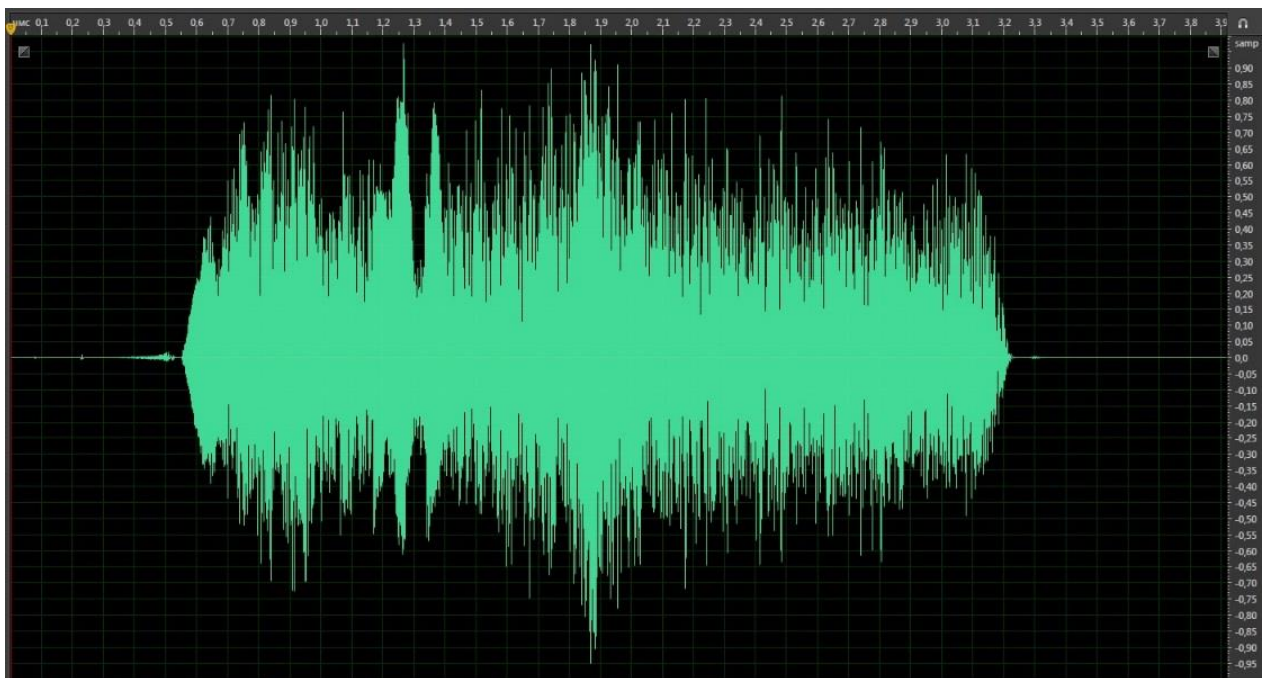


Рисунок 3.7 - Звуковой сигнал звукового файла 3 с отрицательной эмоциональной окраской

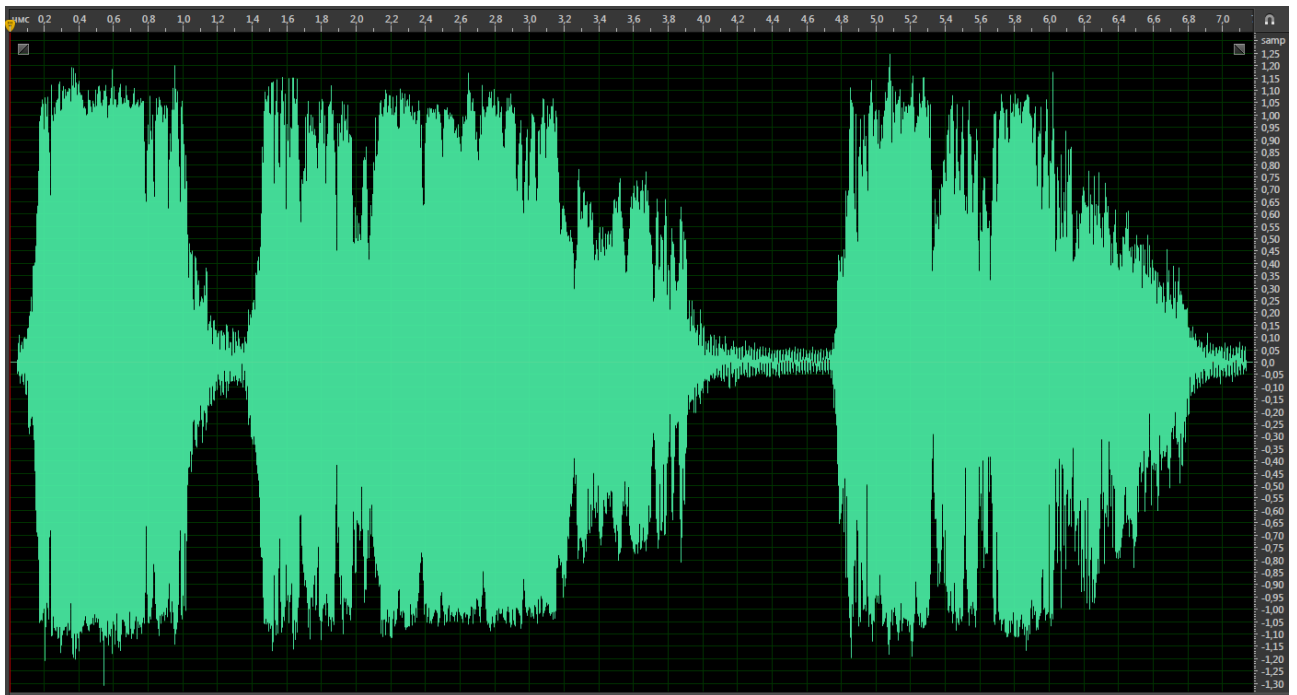


Рисунок 3.8 - Звуковой сигнал звукового файла 4 с отрицательной эмоциональной окраской

На основе подготовленных речевых данных можно произвести исследование линейных характеристик: 1) Частота основного тона, 2) Длительность и количество пауз, 3) Громкость звука.

На основе подготовленных речевых данных возможно исследование линейных характеристик таких как:

-Исследование Частоты основного тона

Для каждого файла нужно анализировать спектр и диапазоны изменения участков основного тона для положительного и отрицательного эмоционального состояния. На основе данного анализа построить таблицы и рассчитать коэффициент отношения участков основного тона для положительного и отрицательного эмоционального состояния.

-Исследование пауз

Исследование количества пауз представляет собой измерение количества и длительности пауз для оценки эмоционального окраса звукового сигнала, а так же расчета их средних значений.

-Исследование громкости звука

Громкость звука так же играет немаловажную роль в определении эмоционального состояния. Поэтому берется звуковой файл со средним значением, представляющим собой нейтральный окрас звукового сигнала, и от него отталкиваемся, для определения эмоционального окраса остальных звуковых файлов.

3.2 Исследование параметра частоты основного тона

Для определения динамического изменения частоты основного тона нужно знать диапазон изменения частоты основного тона для каждого файла (представлены на рисунках 3.9-3.16), а так же количество участков с их \max и \min .

На рисунках 3.9-3.12 представлены спектры диапазона изменения частоты основного тона для звуковых файлов с положительным эмоциональным окрасом речевого сигнала.

На рисунках 3.13-3.16 представлены спектры диапазонов изменения участков основного тона для звуковых файлов с отрицательным эмоциональным окрасом речевого сигнала.

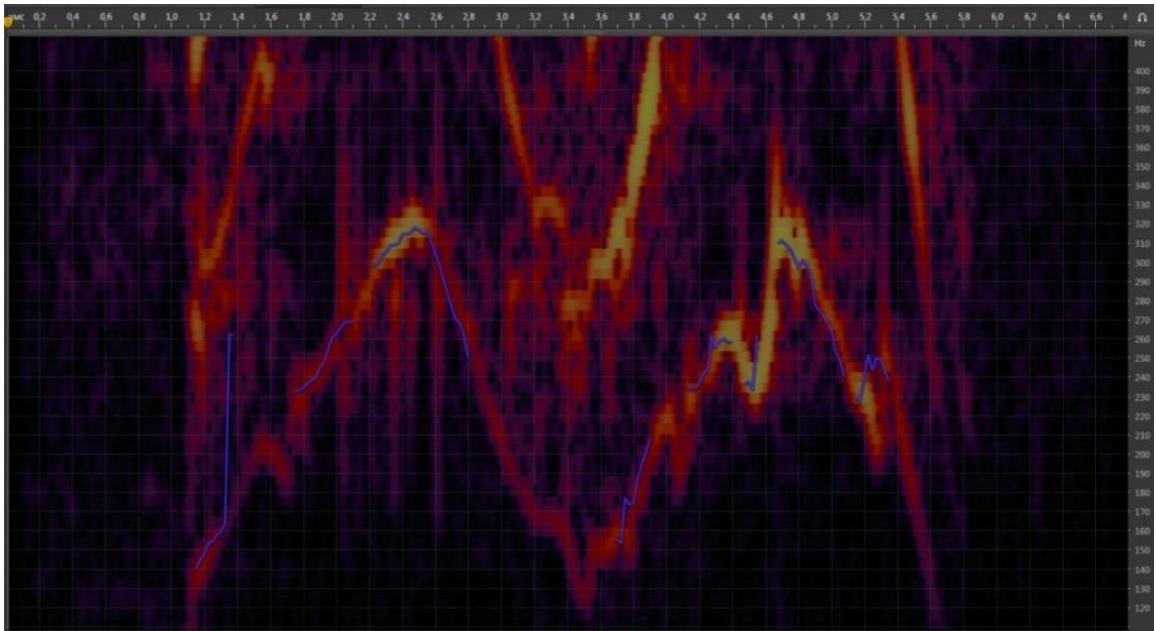


Рисунок 3.9 - Спектр и диапазоны изменения участков основного тона файла 1 при положительном эмоциональном состоянии

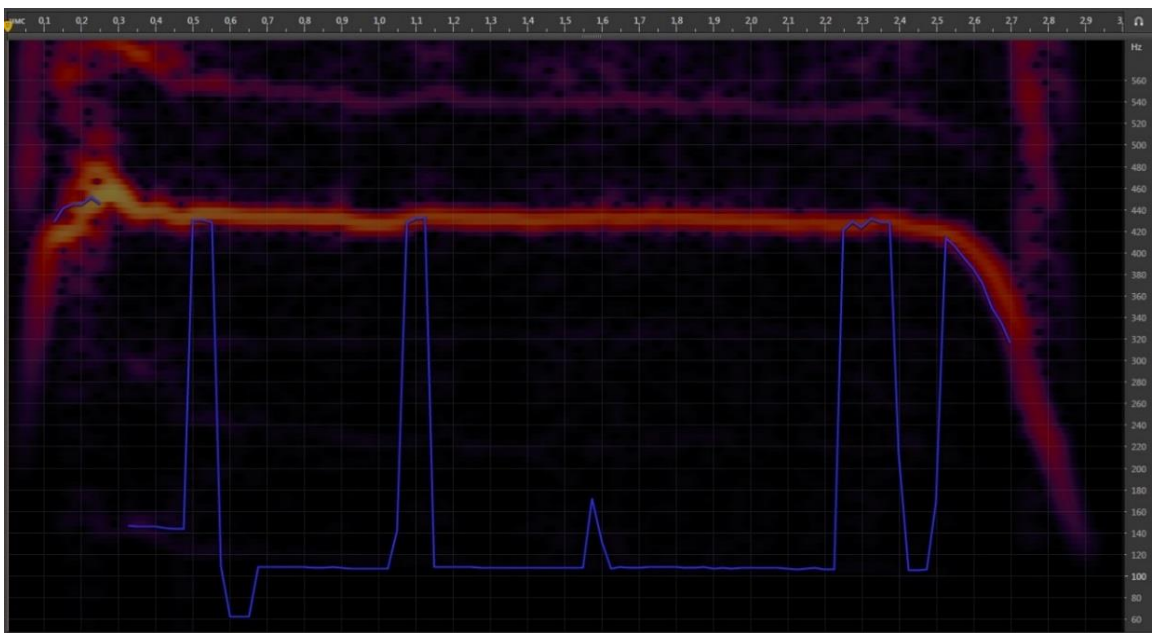


Рисунок 3.11 - Спектр и диапазоны изменения участков основного тона файла 2 при положительном эмоциональном состоянии

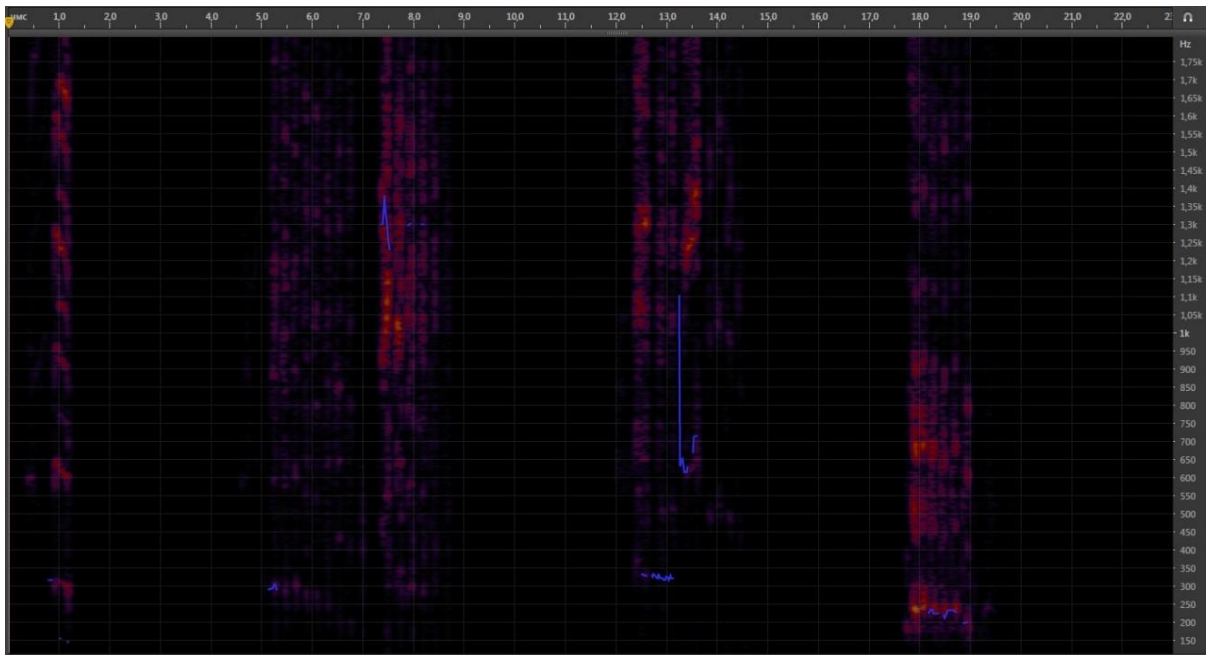


Рисунок 3.12 - Спектр и диапазоны изменения участков основного тона файла 3 при положительном эмоциональном состоянии

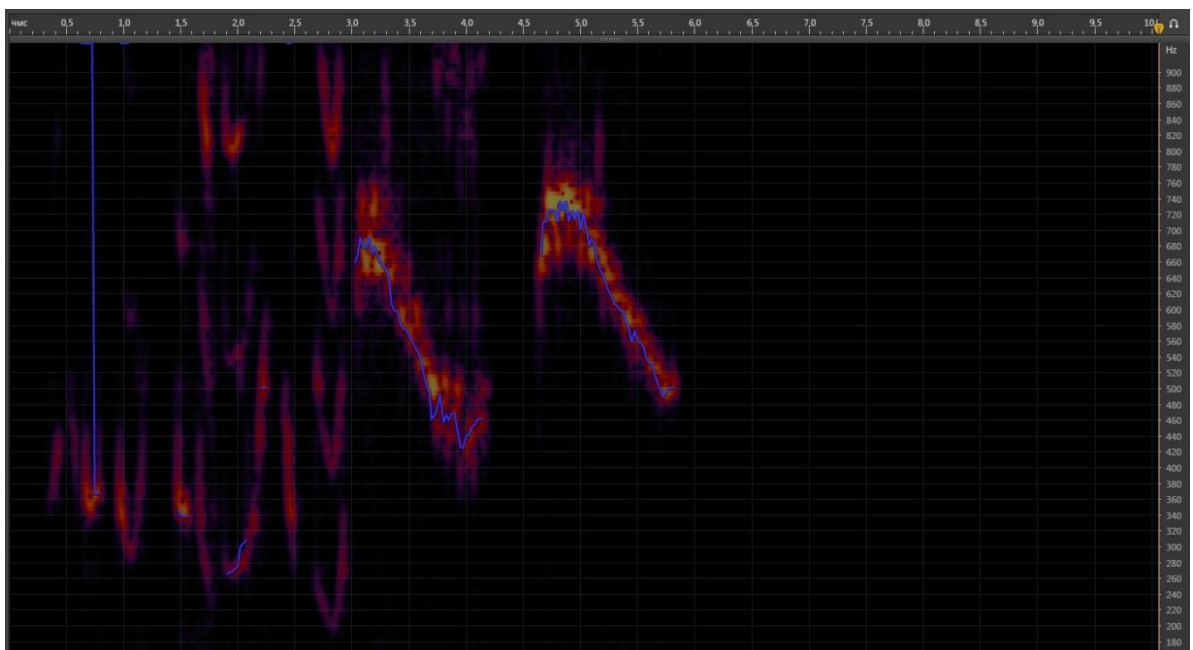


Рисунок 3.13 - Спектр и диапазоны изменения участков основного тона файла 4 при положительном эмоциональном состоянии

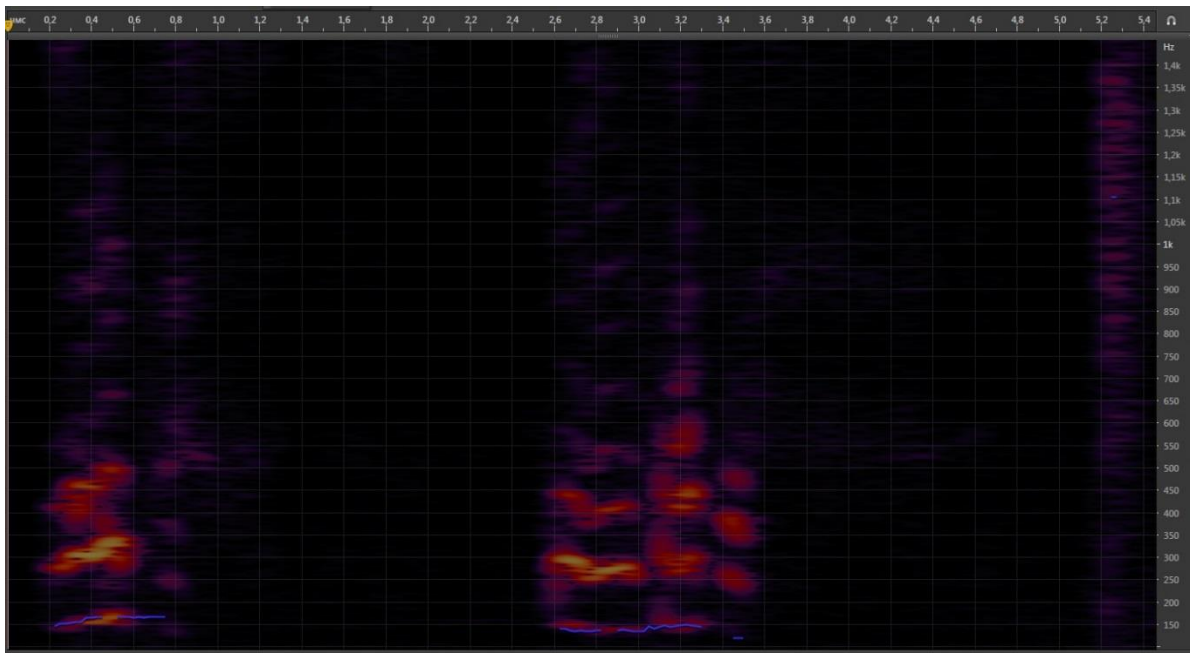


Рисунок 3.14 - Спектр и диапазоны изменения участков основного тона файла 1 при отрицательном эмоциональном состоянии

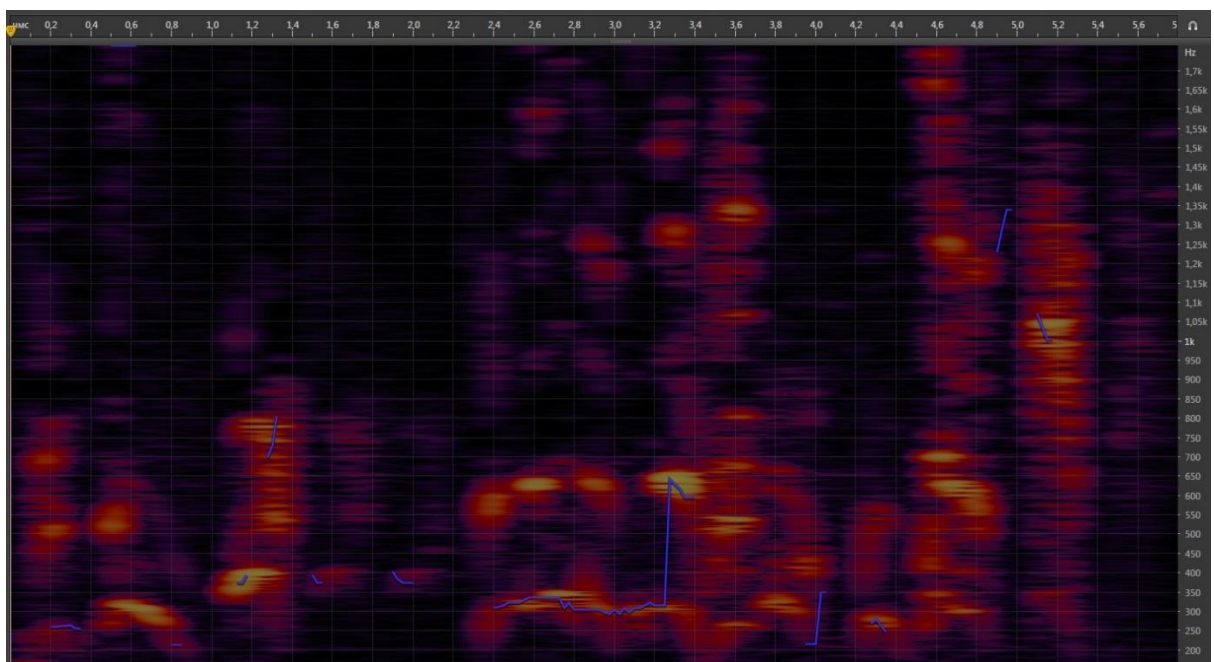


Рисунок 3.15 - Спектр и диапазоны изменения участков основного тона файла 2 при отрицательном эмоциональном состоянии

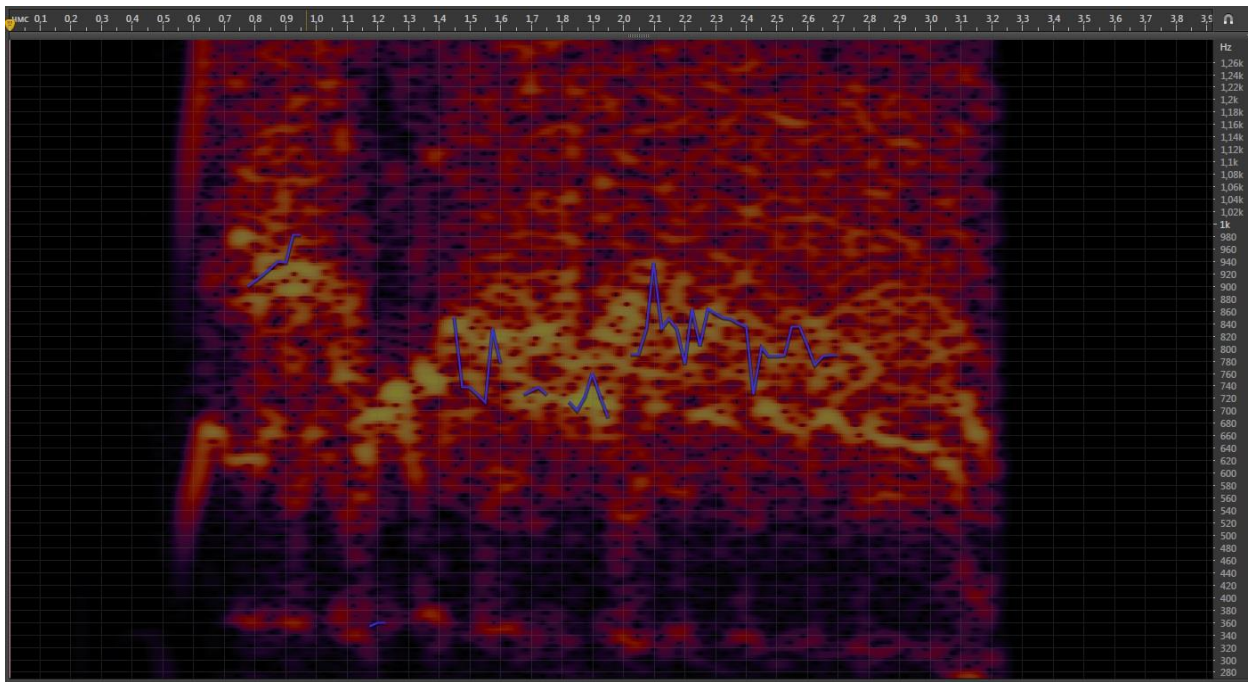


Рисунок 3.16 - Спектр и диапазоны изменения участков основного тона файла 3 при отрицательном эмоциональном состоянии

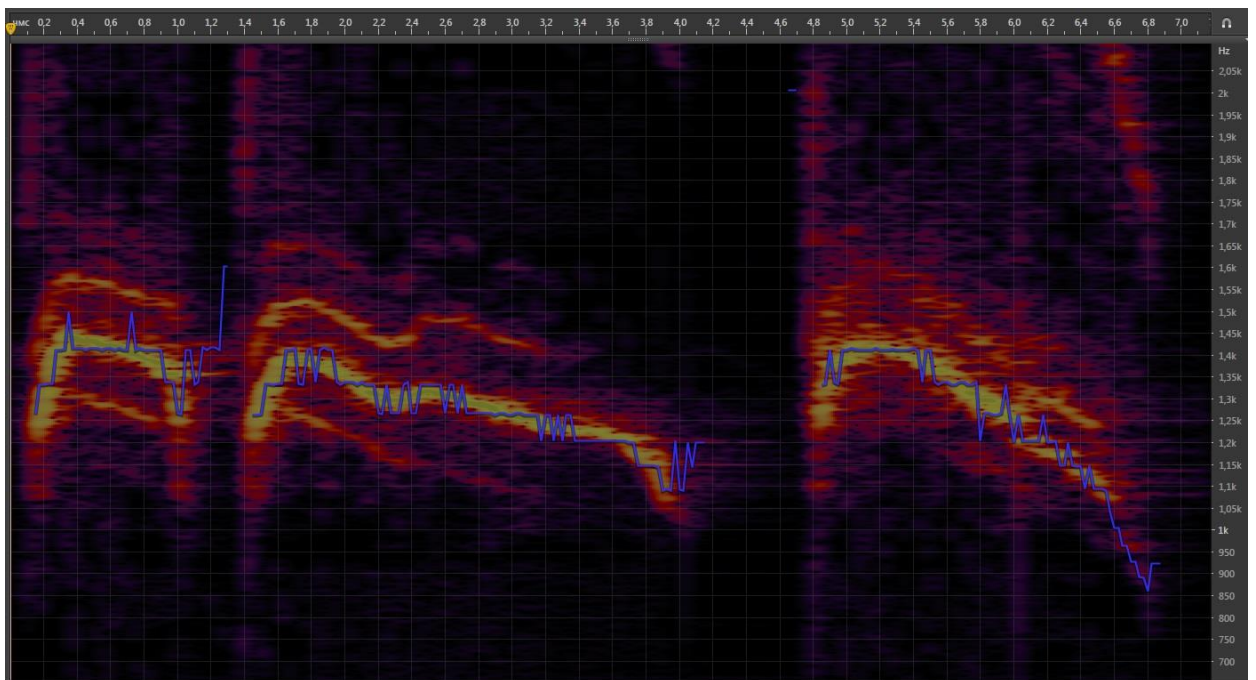


Рисунок 3.17 - Спектр и диапазоны изменения участков основного тона файла 4 при отрицательном эмоциональном состоянии

На основе анализа рисунков 3.9-3.16 можно определить участки основного тона и их количество (таблица 3.3-3.4), а так же зафиксировать чередование min и max значений частоты основного тона, определить количество min и max представленных в таблице 3.5.

Таблица 3.3 - Участки основного тона Звуковых файлов при веселом эмоциональном состоянии

Время С	0,11-0,22	0,38-0,45	1,6-1,75	1,85-1,95	2,25-2,33	2,47-2,53	2,87-2,95
Частота Гц	115	113-110	150-135	122-120	131	113	125-122
Время С	3,1-3,5	4,05-4,18	4,3-4,42	4,5-4,6	5,55-5,85	6,15-6,25	7,05-7,13
Частота Гц	130	131-128	119-120	130-127	151-127	130	210-198
Время С	8,35-8,47	9,15-9,2					
Частота Гц	112-108	150					

Таблица 3.4 - Участки основного тона Звуковых файлов при отрицательном эмоциональном состоянии

Время С	0,4-0,48	0,65-0,95	1,88-1,95	2,1-2,22	2,4-2,47	2,55-2,61	2,78-2,87
Частота Гц	210-205	118-113	132-130	119-117	120-122	120	115-112
Время С	3,25-3,3	3,42-3,47	4,4-4,45	6,41-6,67	7,87-,1	8,67-8,81	8,86-8,9
Частота Гц	112	112	112	121-117	119-120	137-129	118

Таблица 3.5 - чередование min и max значений частоты основного тона

Весело		Грустно	
Min Гц	Max Гц	Min Гц	Max Гц
110	150	113	132
120	131	117	122
113	125	112	121
122	131	117	137
119	130	-	-
127	151	-	-
127	210	-	-
108	-	-	-
118,2 ср	146,8 ср	114,7 ср	128

На основе представленных данных используя таблицы можно определить коэффициент изменения K_v и K_g .

K_v -параметр отношения изменения количества участков записанного звука с веселым эмоциональным состоянием.

N_{iv} -количество изменений min и max записанного звука при веселом эмоциональном состоянии.

N_{ov} -общее количество участков веселого эмоционального состояния.

$$K_v = \frac{N_{iv}}{N_{ov}} = \frac{15}{16} = 0,93 \quad (3.1)$$

K_g -параметр отношения изменения количества участков записанного звука с грустным эмоциональным состоянием.

N_{ig} —количество изменений min и max записанного звука при веселом эмоциональном состоянии

N_{og} -общее количество участков веселого эмоционального состояния

$$Kg = \frac{N_{ig}}{N_{og}} = \frac{8}{14} = 0,57 \quad (3.2)$$

Можно сделать вывод, что от эмоционального состояния голоса диктора зависят параметры записанного звукового файла. При положительном эмоциональном состоянии средние значения частот основного тона $\min=118,2$ Гц и $\max=146,8$ Гц отличаются от средних значений частот основного тона грустного эмоционального состояния и соответственно равны $\min=114,7$ Гц и $\max=128$ Гц на $3,5$ Гц и $18,8$ Гц для \min и \max соответственно. Так же значительно отличается параметр отношения изменения количества участков записанного звука. Для положительного эмоционального состояния он равен $0,93$, а для отрицательного $0,57$. Это связано с тем, что при отрицательном эмоциональном состоянии не такое частое повышение частот основного тона, как при положительном. Даже если посмотреть на рисунки диапазона изменения спектра и основного тона, то можно увидеть, что значения частот основного тона при отрицательном эмоциональном состоянии несколько ниже, чем при положительном, а так же есть отличия в количестве диапазонов изменения участков частот основного тона. При отрицательном эмоциональном состоянии их чуть-чуть меньше, но это не главное, главное то что значений \min и \max в два раза меньше по отношению к общему количеству участков, где показывается в параметрах отношения изменения количества участков записанного звука при отрицательном эмоциональном состоянии.

3.3 Исследование количества пауз

Исследование количества пауз представляет собой измерение количества и длительности пауз для оценки эмоционального окраса звукового сигнала. Для анализа паузы менее 0,1 с не учитывались, а входили в непрерывные звуко сочетания, которые могли представлять собой, как сочетание слов, так и единичные звуки. Данные представлены в таблицах 3.6-3.11.

Таблица 3.6 - Количество пауз и непрерывных звуко сочетаний для звуковых файлов с положительной эмоциональной окраской

звук	Общая длительность, с	Частота, Гц	Количество пауз	Количество непрерывных звуко сочетаний
1	7,674	22050	3	5
2	5,332	44100	3	4
3	7,972	44100	3	2
4	3,870	44100	2	1

Таблица 3.7 - Длительность пауз для звуковых файлов с положительной эмоциональной окраской

звук	Длительность пауз, с			
	1	2	3	Общ.
1	0,340	1,211	0,408	1,959
2	0,320	0,201	0,134	0,655
3	1,121	0,266	0,995	2,382
4	0,376	0,402	-	0,778

Таблица 3.8 - Длительность непрерывных звуко сочетаний для звуковых файлов с положительной эмоциональной окраской

звук	Длительность непрерывных звуко сочетаний, с					
	1	2	3	4	5	Общ.
1	0,833	1,234	0,896	1,220	1,532	5,715
2	0,856	1,071	1,526	1,224	-	4,677
3	2,111	2,247	1,232	-	-	5,590
4	3,092	-	-	-	-	3,092

Таблица 3.9 - Количество пауз и непрерывных звуко сочетаний для звуковых файлов с отрицательной эмоциональной окраской

звук	Общая длительность, с	Частота, Гц	Количество пауз	Количество непрерывных звуко сочетаний
1	7,540	22050	4	3
2	6,496	16000	3	3
3	4,759	44100	3	1
4	7,050	24000	4	3

Таблица 3.10 - Длительность пауз для звуковых файлов с отрицательной эмоциональной окраской

звук	Длительность пауз, с				
	1	2	3	4	Общ.
1	0,206	1,737	1,620	2,111	5,674
2	0,800	0,201	0,546	-	1,547
3	0,551	0,690	0,856	-	2,097
4	0,047	0,233	0,754	0,324	1,358

Таблица 3.11 - Длительность непрерывных звукоочетаний для звуковых файлов с отрицательной эмоциональной окраской

звук	Длительность непрерывных звукоочетаний, с			
	1	2	3	Общ.
1	0,634	0,990	0,242	1,866
2	1,428	1,767	1,754	4,949
3	2,662	-	-	2,662
4	1,051	2,581	2,060	5,692

На основе данных приведенных в таблицах 3.6-3.11 можно высчитать средние значения представленные в таблице 3.12 для анализа зависимости характеристик и расчета их средних значений пауз и непрерывных звукоочетаний звуковых файлов с положительной и отрицательной эмоциональной окраской.

Таблица 3.12 - Средние значения пауз и непрерывных звукоочетаний для звуковых файлов с положительной и отрицательной окраской звуковых файлов

	Положительные эмоции		Отрицательные эмоции	
	Ср. Длительность звука, с	6,212	24,848	6,46125
Ср. количество пауз	2,75	11	3,5	14
Ср количество непрерывных звукоочетаний	3	12	2,25	10
Ср. длит пауз, с	1,4435	5,774	2,669	10,676
Ср. длит. Непрерывных звукоочетаний, с	4,7685	19,074	3,792	15,169

Можно сделать вывод что при примерно одинаковой длительности звука среднее количество и длительность пауз положительных эмоций

меньше, а длительность и количество непрерывных звуко сочетаний наоборот больше.

3.4 Исследование громкости речевого сигнала

За основу берется таблица восприятия разности в физической силе произносимости звуков (таблица 3.13) и осциллограмма звукового файла с нормальным значением (рисунок 3.18). За нормальное значение берем звуковой файл с уровнем 0,4 samp-единица измерения по вертикальной оси графика осциллограммы звукового сигнала с нейтральным эмоциональным окрасом звука.

Таблица 3.13 - Восприятие разности в физической силе произносимости звуков

категория	Громкость, дб	% от нормального
Норм.	50-80	100
Умеренное повышение	80-90	160
Значительное повышение	90-110	180
Крик	<110	<220
Умеренное понижение	40-50	80
Значительное понижение	20-40	40
Шепот	>20	>40

Для определения категории звуковых сигналов возьмем за нормальное значение звуковой файл с нейтральным эмоциональным окрасом (рис. 3.18).

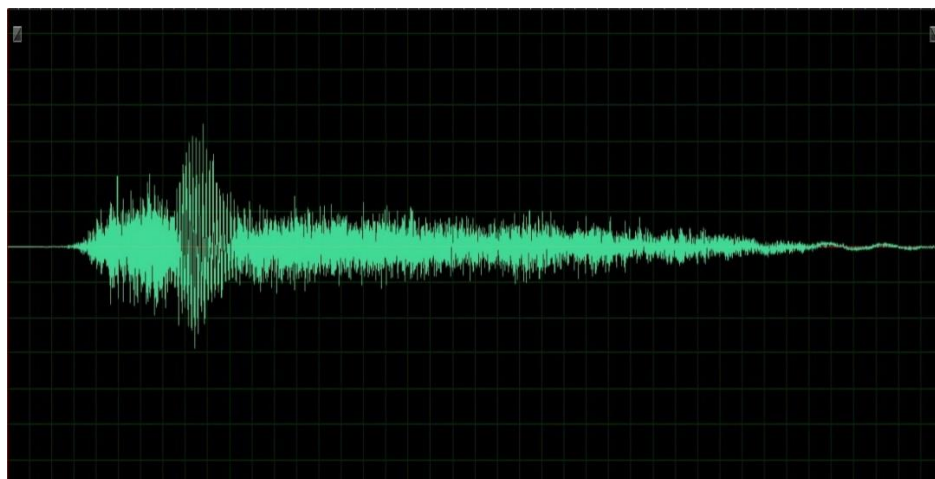


Рисунок 3.18 - Звуковой файл с нейтральным эмоциональным окрасом

Звуковой файл с нейтральной эмоциональной окраской имеет 40% уровня громкости (0,4 samp)

На основе этих данных и данных с осциллограммы звуковых файлов построим таблицу 3.14 определения категории звукового файла.

Таблица 3.14 - Категория определения звука относительно нормальной

	Значение уровня громкости, Samp	% от нормальной	категория
Звуковой файл положительного эмоционального окраса			
1	0,7	175	Умеренное повышение
2	0,6	150	Умеренное повышение
3	0,95	230	крик
4	0,75	187	Значительное повышение
Звуковой файл отрицательного эмоционального окраса			
1	0,4	100	Норм.
2	1	250	крик
3	0,75	185	Значительное повышение
4	1,1	275	крик

По уровню громкости нельзя точно определить какой именно эмоциональный окрас имеет звуковой файл, так как при радостном всплеске эмоций по таблице есть определение по категории эмоций-крик, что не является действительностью, а при угнетенном состоянии показывает нормальный, то есть нейтральное состояние. Но по большому счету при радостном всплеске положительного эмоционального окраса и разрывающимся крике отрицательного окраса, возможно совпадение уровня громкости, такое же совпадение возможно при угнетенном состоянии.

ЗАКЛЮЧЕНИЕ

В ходе выполнения выпускной квалификационной работы было выявлено, что для определения эмоционального окраса звуков речи важно применение сразу несколько комбинаций характеристик, чем их больше, тем больше вероятность точного определения эмоционального окраса. Данными характеристиками являются: основной тон, количество пауз, громкость речи.

Проведено исследование влияние эмоций на характеристики речевого сигнала. Изучены признаки и характеристики звуковых сигналов

Проведен сравнительный анализ характеристик речевого сигнала. Самой эффективной характеристикой оказалась характеристика оценивающая основной тон. Самой неэффективной громкость звука. Громкость звука может меняться в зависимости от эмоционального состояния будто это крик положительный или отрицательный.

Проведено исследование влияние эмоций на характеристики речевого сигнала. Изучены признаки и характеристики звуковых сигналов

Можно сделать вывод, что от эмоционального состояния голоса диктора зависят параметры записанного звукового файла. При положительном эмоциональном состоянии средние значения частот основного тона $\min=118,2$ Гц и $\max=146,8$ Гц отличаются от средних значений частот основного тона грустного эмоционального состояния и соответственно равны $\min=114,7$ Гц и $\max=128$ Гц на $3,5$ Гц и $18,8$ Гц для \min и \max соответственно. Так же значительно отличается параметр отношения изменения количества участков записанного звука. Для положительного эмоционального состояния он равен $0,93$, а для отрицательного $0,57$. Это связано с тем, что при отрицательном эмоциональном состоянии не такое частое повышение частот основного тона, как при положительном. Даже если посмотреть на рисунки диапазона изменения спектра и основного тона, то можно увидеть, что значения частот основного тона при отрицательном

эмоциональном состоянии несколько ниже, чем при положительном, а так же есть отличия в количестве диапазонов изменения участков частот основного тона. При отрицательном эмоциональном состоянии их чуть-чуть меньше, но это не главное, главное то что значений \min и \max в два раза меньше по отношению к общему количеству участков, где показывается в параметрах отношения изменения количества участков записанного звука при отрицательном эмоциональном состоянии.

При примерно одинаковой длительности звука среднее количество и длительность пауз положительных эмоций меньше, а длительность и количество непрерывных звукосочетаний наоборот больше.

По уровню громкости нельзя точно определить какой именно эмоциональный окрас имеет звуковой файл, так как при радостном всплеске эмоций по таблице есть определение по категории эмоций-крик, что не является действительностью, а при угнетенном состоянии показывает нормальный, то есть нейтральное состояние. Но по большому счету при радостном всплеске положительного эмоционального окраса и разрывающимся крике отрицательного окраса, возможно совпадение уровня громкости, такое же совпадение возможно при угнетенном состоянии.

Проведен сравнительный анализ характеристик речевого сигнала. Самой эффективной характеристикой оказалась характеристика оценивающая основной тон. Самой неэффективной-громкость звука. Громкость звука может меняться в зависимости от эмоционального состояния будто это крик положительный или отрицательный.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Леонтьев В.О. Классификация эмоций [Текст] // Одесса, 2001;
2. Изард К. Психология эмоций [Текст] // Спб, 1999;
3. Андреева И. Эмоциональный интеллект: исследование феномена [Текст] // Вопросы психологии, №3, 2006.
4. Засыпкин, А.В. О дикторнезависимой системе голосового телефонного номеронабирателя [Текст] // А.В. Засыпкин, А.Т. Мицевич, М.В. Овецкий, В.Ю. Шелепов// Труды международной конференции “Знание-Диалог-Решение”. – Ялта. – 1995. – С.427-430.
5. Афифи А., Эйзен С. Статистический анализ. Подход с использованием ЭВМ [Текст] // М: Мир, 1982, с. 234-244;
6. Гусев А.Н. Дисперсионный анализ в экспериментальной психологии [Текст] // Методическое пособие для студентов факультетов психологии вузов, М: Психология, 2000;
7. Лабоцкий В.В. Анализ экономических данных [Текст] // использование Statistica методическое пособие, 2006 ;
8. Минковский Г. Пространство и время [Текст]// Успехи Физических Наук (УФН), т.69,в.2, 1959;
9. Статистика и обработка данных в психологии [Электронный ресурс] // Данные в психологических аспектах URL:<http://psyfactor.org/lib/stat3.htm>; (дата обращения 18.05.2018)
10. Станиславский К. С. Работа актёра над собой [Текст] //М: Художественная литература, 1938;
11. Витт Н.В. Личностно-ситуационная опосредованность выражения [Электронный ресурс] // /Распознавание эмоции в речи - URL:<http://www.voppsy.ru/issues/1991/911/911095.htm>; (дата обращения 07.05.2018)

12. Пашина А.Х. К проблеме распознавания эмоционального контекста звуковой речи // Эмоциональный окрас звуков- URL:<http://www.voppsy.ru/issues/1991/911/911088.htm>; (Дата обращения 29.05.2018)
13. Маслобоев Ю.П., Рычагов М.Н. Нейронные сети [Текст] // Методическое пособие, М: МИЭТ, 2006;
14. Уоссермен Ф. Нейрокомпьютерная техника [Электронный ресурс] // neurnews.iu4.bmstu.ru/ Нейрокомпьютеры - архитектура и реализация – URL:<http://neurnews.iu4.bmstu.ru/book/nkt/> (дата обращения 04.05.2018)
15. Яхъяева Р.Э. Основы нейронных сетей [Электронный ресурс]// Пространства нейронных сетей URL:www.intuit.ru/department/ds/neuronnets/;
16. Видяпин В.И., Борисов А.С., Данько Т.П. Бакалавр экономики т.2 [Текст]// Триада-Х, 1999;
17. Моисеева Н.К., Костина Г.Д. Маркетинговые исследования при создании и использовании программных продуктов [Текст] //Методические указания для выполнения курсовых и дипломных работ по специальности «Менеджмент» М.: МГИЭТ (ТУ), 1996;
18. Короткова Т.Л., Лукичева Л.И. Методические указания по выполнению курсовых работ и организационно-экономической части дипломных проектов по тематике курса «Основы маркетинга» [Электронный ресурс] // М.: МГИЭТ (ТУ), 1994. Багиев Г.Л., Богданова Е.Л. Маркетинг-статистика // Электронный учебник, URL: <http://www.marketing.spb.ru/read/m9/index.htm> (дата обращения 04.04.2018)
19. Affective Computing: техника не разделяет наши чувства [Текст]// Интернет-журнал Мембрана, 2003;
20. под ред. Белова С.В. охрана окружающей среды [Текст] // М., Высшая школа, 1983;
21. Засыпкин, А.В. О дикторонезависимой системе голосового телефонного номеронабирателя [Текст] // А.В. Засыпкин, А.Т. Мицевич, М.В.

Овецкий, В.Ю. Шелепов [Текст] // Труды международной конференции "Знание-Диалог-Решение". – Ялта. – 1995. – С.427-430.

22. Кавальчук, А.Н. (2011), "Формула для перехода из области частот к шкале барков и обратно," А.Н. Кавальчук, Ал.А. Петровский [Текст] // Информатика, 2011, 4(32), стр. 71-81

23. Каганов, А.Ш. Криминалистическая экспертиза [Текст] // Экспертиза звукозаписей. – М.: "Юрлитинформ", 2005. - 272с.

24. Кипяткова И.С. Автоматическая обработка разговорной русской речи: монография [Текст] // И.С. Кипяткова , А.Л. Ронжин, А.А. Карпов. СПИИРАН – СПб.: ГУАП, 2013. – 314 с.

25. Колерс, П.А. Распознавание образов. Исследование живых и автоматических распознающих систем [Текст] // П.А. Колерс, Е.Д. Мюрей, пер. Л.И. Титомира – М.: «Мир», 1970. – 288 с.

26. Ле, Н.В. Распознавание речи на основе искусственных нейронных сетей [Текст] // Н.В. Ле, Д.П. Панченко Технические науки в России и за рубежом: материалы междунар. заоч. науч. конф.– Москва. – 2011. – С.8-11.

27. Леонович, А.А. Современные технологии распознавания речи [Текст] //А.А. Леонович // Материалы конференции «Диалог: Компьютерная лингвистика и интеллектуальные технологии». – Звенигород. – 2005.

28. Мазуренко, И.Л. Компьютерные системы распознавания речи [Текст] // И.Л. Мазуренко // Интеллектуальные системы. – Москва. – 1998. – т.3. вып. 1-2. – С.117-134.

29. Мазуренко, И.Л. Одна модель распознавания речи [Текст] // И.Л. Мазуренко // Компьютерные аспекты в научных исследованиях и учебном процессе. – Москва – 1996 – С.107-112.

30. Малла, С. Вэйвлеты в обработке сигналов [Текст] // М.: Мир, 2005. — 672 с.

Выпускная квалификационная работа выполнена мной совершенно самостоятельно. Все использованные в работе материалы и концепции из опубликованной научной литературы и других источников имеют ссылки на них.

«__» _____ г.

(подпись)

(Ф.И.О.)