Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования   
«Санкт-Петербургский политехнический университет Петра Великого»

Институт компьютерных наук и технологий

**Кафедра компьютерных систем и программных технологий**

**ВЫПУСКНАЯ РАБОТА БАКАЛАВРА**

**Исследование методов обработки речи для передачи по каналу связи**

направление: 09.03.01 - Информатика и вычислительная техника

|  |  |
| --- | --- |
|  | Выполнил:  Алексеев Даниил Михайлович Подпись\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
|  | Руководитель: доцент, к.т.н.,  Богач Наталья Владимировна Подпись\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |

Санкт-Петербург  
2016

# Реферат

РЕЧЕВОЙ КОДЕР, ПОМЕХОЗАЩИЩЁННОЕ КОДИРОВАНИЕ, КАНАЛЬНЫЕ КОДЕКИ, ПОДКЛЮЧАЕМАЯ БИБЛИОТЕКА, МЕТОДОЛОГИЯ ВЫБОРА СПОСОБОВ ОБРАБОТКИ РЕЧИ.

В выпускной работе рассматривается разработка библиотек для канального и речевого кодирования и методы обработки речевого сигнала.

Содержание

[Реферат 3](#_Toc451277121)

[Введение 7](#_Toc451277122)

[Постановка задачи 8](#_Toc451277123)

[1. Исследование методов обработки речевого сигнала 10](#_Toc451277124)

[1.1. Канальный уровень: блочные кодеки 10](#_Toc451277125)

[1.1.1. Проверка на чётность 10](#_Toc451277126)

[1.1.2. Код Хэмминга 12](#_Toc451277127)

[1.1.3. Циклические коды 13](#_Toc451277128)

[1.2. Канальный уровень: свёрточные кодеки 13](#_Toc451277129)

[1.2.1. Ввод контрольных бит 13](#_Toc451277130)

[1.2.2. Кодирование полиномом с задержкой 14](#_Toc451277131)

[1.3. Канальный уровень: перемежение (interleaving) 15](#_Toc451277132)

[1.4. Речевые кодеки 15](#_Toc451277133)

[1.4.1. Вокодер 16](#_Toc451277134)

[1.4.2. Липридер 16](#_Toc451277135)

[1.4.3. Гибрид из вокодера и липридера 16](#_Toc451277136)

[2. Проектирование библиотек 17](#_Toc451277137)

[2.1. Модули для обработки речевого сигнала 18](#_Toc451277138)

[2.1.1. Модуль для кодирования речевого сигнала 18](#_Toc451277139)

[2.1.2. Модуль для декодирования речевого сигнала 19](#_Toc451277140)

[2.2. Модули для работы на канальном уровне 19](#_Toc451277141)

[2.2.1. Модуль для канальных кодеков 19](#_Toc451277142)

[2.2.2. Модуль интерливинга и деинтерливинга 20](#_Toc451277143)

[3. Разработка библиотек 22](#_Toc451277144)

[3.1. Структура библиотек и описание 22](#_Toc451277145)

[3.2. Разработка канальной библиотеки 23](#_Toc451277146)

[3.2.1. Описание функций канальной библиотеки 23](#_Toc451277147)

[3.3. Разработка библиотеки для речевого сигнала 26](#_Toc451277148)

[3.3.1. Описание функций библиотеки для речевой обработки сигнала 26](#_Toc451277149)

[Заключение 34](#_Toc451277150)

[Список использованных источников 35](#_Toc451277151)

[Приложение 1 37](#_Toc451277152)

[Приложение 2 39](#_Toc451277153)

[Приложение 3 40](#_Toc451277154)

[Приложение 4 42](#_Toc451277155)

**Список рисунков**

[Рис. 1.1. Последовательность кодирования речи 7](#_Toc451205564)

[Рис. 1.2. Алгоритм дополнения до чётности, двумерный вариант 11](#_Toc451205565)

[Рис. 1.3. Представление закодированных бит в трёхмерном варианте контроля на чётность 12](#_Toc451205566)

[Рис. 4. Рисунок на который ссылаемся из текста 42](#_Toc451205567)

# Введение

Не существует настолько надёжных каналов связи, которые могут обеспечить полное отсутствие помех, воздействующих на передаваемую речь. Поэтому были изобретены различные методы защиты от помех. Из них можно выделить три наиболее популярных: аппаратное улучшение аппаратуры и канала связи, повышение отношения сигнал-шум, обнаружение и исправление ошибок в принятых сообщениях. Первый способ защиты позволяет улучшить качество передаваемой речи только до определённых пределов, второй подразумевает работу на физическом уровне (что также имеет свои пределы). Поэтому мы рассмотрим методы обнаружения и исправления ошибок в речевом сигнале.

Кодирование речи можно разбить на два основных этапа: непосредственно речевое кодирование и канальное кодирование (которое, как правило, делится на блочное и свёрточное кодирования, а также деинтерливинг (перемежение)). На Рис. 1 приведена поясняющая схема последовательности кодирования.

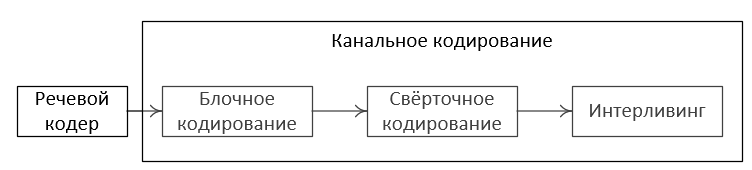


Рис. 1.1. Последовательность кодирования речи

Речевое кодирование не является помехоустойчивым, поэтому данные с речевого кодера поступают на канальный. Его задача – сделать передаваемую информацию помехоустойчивой, т.е. дать возможность приёмнику обнаружить (и, в некоторых случаях, исправить) ошибки, которые возникают при передаче информации. Помимо этого, канальное кодирование может выполнять такие функции, как добавление управляющей информации и шифрование.

При блочном кодировании информация делится на блоки определенной длины, и каждый блок кодируется отдельно. Простейшим примером блочного кодирования является дополнение до четности. В этом случае каждый блок делится на группы, которые дополняются одним битом со значением единица или ноль, в зависимости от четного или нечетного количество единиц в исходной группе.

При свёрточном кодировании работа ведётся сразу со всей поступившей информацией (без разделения её на части). Самый простой способ свёрточного кодирования – это добавление между информационных бит проверочных, которые зависят от рядом стоящих информационных бит.

Оба описанных выше метода кодирования позволяют не только определить наличие ошибки в передаваемом коде, но и устранить её. Последнее не всегда является обязательным требованием: бывают ситуации, когда намного проще повторно отправить данные, или же оставить всё, как есть (в таком случае мы получим небольшие помехи во время разговора). К минусам данных методов можно отнести то, что они не могут восстановить правильную последовательность, если ошибкам подверглось слишком много подряд идущих бит. Решением данной проблемы является интерливинг.

Интерливинг нужен для того, чтобы переставить местами закодированную последовательность: если при передаче информация подверглась пачечным ошибкам, то на приёмной стороне после сборки последовательности для декодирования повреждённые биты окажутся на значительном расстоянии друг от друга. Для надёжности можно выполнить процедуру интерливинга несколько раз подряд.

## Постановка задачи

Исходя из вышесказанного, мы можем сформулировать задачи, которые требуется решить при выполнении данной работы.

Необходимо исследовать существующие методы обработки речи и предложить свою реализацию библиотек, предназначенных для обработки речи, а также методологию выбора более подходящего варианта обработки с учётом имеющегося канала связи.

К канальным кодекам выдвигаются следующие требования:

* простота реализации;
* допустимая избыточность;
* возможность коррекции ошибок.

К речевым кодекам выдвигаются следующие требования:

* минимальное снижение качества сигнала;
* небольшие задержки.

Стоит отметить, что не всегда удаётся соответствовать всем требованиям. Чаще всего они даже противоречат друг другу.

# 1. Исследование методов обработки речевого сигнала

## 1.1. Канальный уровень: блочные кодеки

Блочному кодированию подвергается не весь речевой сигнал, а только его часть. Это связано с тем, что блочными кодами, как правило, не могут кодировать длинные последовательности бит, или при повышении числа информационных бит избыточность становится неприемлемой.

По ходу выполнения работы было выделено три наиболее популярных вида блочных кодов, а также предложены конкретные варианты их применения (по числу информационных и контрольных бит).

### 1.1.1. Проверка на чётность

Данный вид кодирования выделяется следующими свойствами:

* простота реализации;
* возможность находить и восстанавливать ошибки;
* не защищён от ошибок в контрольных битах.

Нами предлагается проверка на чётность через матрицу; при этом можно выделить два основных вида матриц: двумерные и трёхмерные.

В ходе исследований было установлено, что для первого вида наиболее оптимальным является размер матрицы 2х2 информационных бит: в таком случае мы передаём 4 информационных и 4 контрольных бита, и можем исправить 1 ошибку. На Рис. 2 изложена основная идея алгоритма.

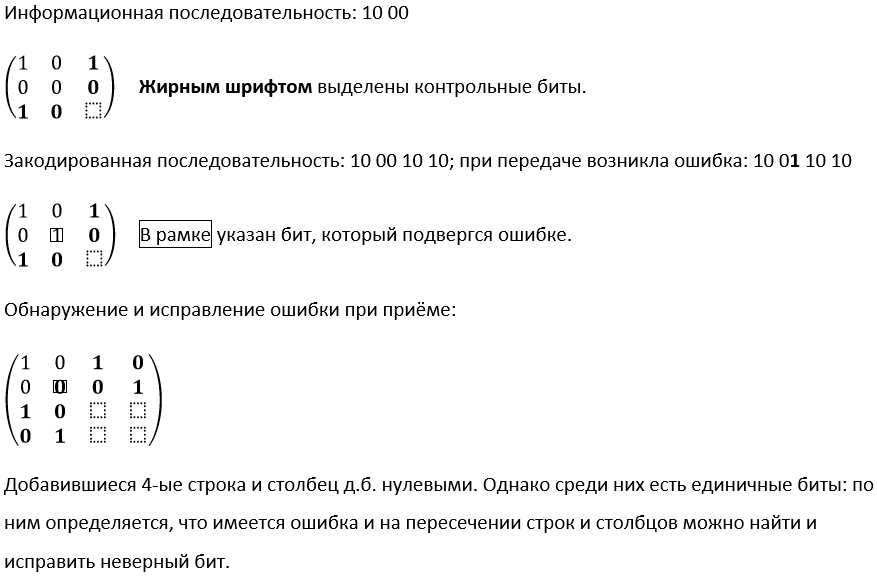


Рис. 1.2. Алгоритм дополнения до чётности, двумерный вариант

Для трёхмерного представления более удобно использовать кодирование бит в виде трёх матриц 3х3. Помимо чётности по столбам и строкам, добавляется матрица проверки «по глубине» - см. Рис. 3. Пример работы с такой матрицей рассмотрен в [1]. На каждые 27 информационных бит мы получаем 27 контрольных, и можем исправить 3 ошибки.

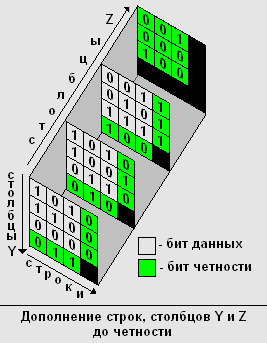


Рис. 1.3. Представление закодированных бит в трёхмерном варианте контроля на чётность

Следует напомнить, что данный вид кодирования не защищает от ошибок в контрольных битах. Однако по сравнению с одномерным вариантом при чётном числе ошибок мы можем узнать, что принятая последовательность – битая.

### 1.1.2. Код Хэмминга

Данный алгоритм весьма популярен. Он позволяет обнаружить и исправить однократные ошибки. В результате исследований было определено, что наиболее оптимальным вариантом будет код Хэмминга (4, 7): 4 информационных бита, 3 контрольных (7 всего). Пример работы с кодом Хэмминга приведён в [2].

Следует отметить, что кодирующий полином м.б. задан разработчиком: как следствие мы получим дополнительную защиту от прослушки. Кроме того, данный алгоритм не ломается при наличии ошибки в контрольном бите.

К минусам данного алгоритма можно отнести сложность алгоритма декодирования: он требует работы с матрицами (что часто бывает ресурсоёмкой процедурой). Однако при использовании частного варианта кода Хэмминга, при котором код ошибки совпадает с позицией сломанного бита, работа с матрицами не требуется (однако при этом мы теряем дополнительную защиту от прослушки).

### 1.1.3. Циклические коды

Данный вид кодов является наиболее сложным. Однако он позволяет при вводе большой избыточности исправлять более одной ошибки. Тем не менее, в результате исследования было решено, что наиболее оптимальным вариантом будет применением циклического кода (4, 7). Примеры работы с данным видом кодов можно найти в [3].

Несомненный плюс предложенного варианта алгоритма – ввод своей таблицы кодирования. Также при вводе большой избыточности мы получаем возможность исправлять более чем одну ошибку.

К минусам данного алгоритма стоит отнести декодирование: он требует многократные циклические сдвиги и деление полиномов.

## 1.2. Канальный уровень: свёрточные кодеки

Канальное кодирование применяется сразу ко всему передаваемому пакету, без разделения его на блоки. По результатам исследований было выделено два наиболее популярных свёрточных кодека, а также предложено улучшение к одному из них.

### 1.2.1. Ввод контрольных бит

Данный метод является наиболее простым. Пример работы с данным методом описан в [1], однако сам предложенный там алгоритм требует доработки.

При кодировании между информационными битами вставляются проверочные (они представляют собой сумму по модуль два соседних бит). Декодирование выглядит следующим образом: суммируются соседние биты исходных данных и сравниваются с их проверочным битом.

* Если для двух соседних проверочных битов была зафиксирована ошибка, то общий информационный бит для этих двух проверочных битов - неверен. Для исправления ошибки необходимо заменить его на противоположный.
* Если для одного проверочного символа была зафиксирована ошибка, а два соседних проверочных символа ошибку не показали, это означает, что сбой произошел в проверочном символе, а информационные биты корректны.

Очевидная проблема данного алгоритма – ошибка в крайних битах. Для её устранения предложен следующий подход: добавлять в начало и конец последовательности по два дополнительных бита, значение которых нам всегда известно. Независимо от принятых данных, мы будем декодировать код исходя из того, что первые и последние три бита – это наши константы и проверочный бит, образованный ими. Таким образом, даже при наличии ошибки в первых/последних двух переданных битах мы сможем без проблем декодировать последовательность.

К плюсам данного алгоритма можно отнести следующее: простота реализации, ошибка в контрольных битах не приводит к краху алгоритма.

Серьёзным минусом данного алгоритма можно считать следующее: если расстояние между ошибочными битами меньше 2 (они стоят рядом или через один), то алгоритм породит дополнительные ошибки.

### 1.2.2. Кодирование полиномом с задержкой

Данный метод подразумевает использование полинома с задержкой. Кодеры такого типа имеют скорость ½ или ¼ (т.е. на каждый входной бит приходится 2 или 4 закодированных бита). Алгоритм кодирования описан в [4] и [5], алгоритм декодирования – в [6].

К плюсам данного метода можно отнести высокую надёжность: по результатам опытов декодер справлялся примерно с 40% ошибок. Кроме того, несколько подряд идущих ошибок (не более 3) не ломали алгоритм декодирования.

К минусам данного метода можно отнести большую избыточность, высокую сложность алгоритма декодирования и его ресурсоёмкость.

Однако несмотря на указанные минусы, именно этот метод является самым распространённым и используемым в сетях связи.

## 1.3. Канальный уровень: перемежение (interleaving)

Ранее отмечалось, что не все кодеки могут справиться с несколькими ошибками в переданном пакете/блоке информации. Такие ошибки называются пачечными: они возникают из-за того, что длительность воздействующего вредного сигнала достаточно высока для того, чтобы повредить несколько рядом стоящих бит.

Для борьбы с пачечными ошибками было введено перемежение (интерливинг – interleaving). Перед отправкой в канал связи биты перестраиваются таким образом, чтобы соседние биты оказались как можно дальше друг от друга. При приёме происходит деинтерливинг: мы получаем исходную последовательность; в результате повреждённые биты оказываются на довольно большом расстоянии друг от друга. Примеры интерливинга можно найти в [4].

## 1.4. Речевые кодеки

Можно выделить два основных вида речевых кодеров, которые в свою очередь порождают гибридный вариант: это вокодеры и липридеры. Нередко бывает так, что одно устройство поддерживает сразу оба варианта.

### 1.4.1. Вокодер

Принцип действия вокодеров основан на том, что они сопоставляют принятый голосовой сигнал с имеющимся фиксированным словарём, достают из него номер фонемы, который соответствует сигналу, и передают его. Такой подход значительно снижает количество передаваемой информации, но сам голосовой сигнал теряет характерное для человека звучание. Более подробно с принципом действия вокодеров можно ознакомиться в [7].

Сейчас почти не используется в чистом виде, однако применяется в сочетании с липридером.

### 1.4.2. Липридер

Данный вид речевых кодеров является наиболее распространённым в наши дни. Липридеры снимают характеристики принятого голосового сигнала и передают их. Данный подход обеспечивает более человеческое звучание, но количество передаваемой информации заметно возрастает. Более подробно с принципом действия липридеров можно ознакомиться в [7] и [8].

Следует заметить: из-за использования линейного предсказания в липридерах речь собеседника не всегда удаётся передать более точно (возможно искажение звуков).

### 1.4.3. Гибрид из вокодера и липридера

При сочетании вокодера и липридера удаётся избежать минусов устройств, в которых данные подходы реализованы отдельно (однако в таком случае повышается количество передаваемой информации: помимо номера фонемы передаются параметры речевого сигнала).

Основой аналитической части таких устройств служат фиксированный и адаптивный кодовые словари, которые позволяют более точно подобрать параметры фильтра для воспроизводящего устройства. Более подробно можно ознакомиться с гибридным вариантом в [8] и [9].

# 2. Проектирование библиотек

С учётом того, что обработка речевого сигнала м.б. разделена на два основных этапа, мы можем выделить библиотеки для работы с канальным уровнем и работы с самим речевым сигналом.

Библиотека для канального уровня будет содержать четыре модуля: кодирования, декодирования, интерливинга и деинтерливинга.

Для обработки речевого сигнала требуется четыре модуля: два кодирования и два декодирования (по одному на каждый вид кодека).

Было решено написать библиотеки на языке Си. Принятию данного решения способствовали следующие факторы:

* язык Си наравне с ассемблерами используется для программирования микроконтроллеров – таким образом область применения написанных библиотек значительно расширяется;
* написанное на классическом Си приложение будет кроссплатформенным (на уровне исходных кодов): данные библиотеки получат возможность встраиваться в приложения, написанные на разных ОС;
* согласно разным источникам [10][11], язык программирования Си довольно долгое время занимает верхние позиции по популярности (таким образом, найти специалиста для поддержки и сопровождения библиотек будет просто).

Данные библиотеки разрабатывались по такой технологии проектирования ПО, как прототипирование. В пользу данной технологии были следующие факторы:

* нам известны не все требования (формат, в котором мы будем передавать данные в канал связи; скорость передачи данных);
* данная технология позволила нам быстро увидеть некоторые свойства продукта (применимость, удобство);

Подробнее с данной моделью можно ознакомиться в [12].

Также было решено определить такой тип данных, как Voice\_type: он описывает речевой сигнал. В данном типе предлагается хранить принятый речевой сигнал, а также содержимое словарей (фиксированного и адаптивного).

## 2.1. Модули для обработки речевого сигнала

Данные модули можно использовать отдельно от предлагаемых модулей для работы на канальном уровне. В таком случае, если взаимодействие предлагаемых кодеков для обработки речи с канальными кодеками (отличными от предложенных в данной работе) будет невозможно из-за различающихся форматов представления данных, предлагается использовать такой шаблон проектирования, как адаптер (adapter/wrapper).

Все модули для обработки речевого сигнала представлены в виде интерфейсов. Это было сделано по следующим причинам:

* существует множество способов реализации речевых кодеков, каждый из которых требует довольно большой теоретической базы (математическое описание, особенности речи). Ввод хотя бы в один способ реализации займёт крайне много времени.
* используя интерфейс, мы можем без проблем рассказать, какие действия должен уметь выполнять данный алгоритм и, следовательно, что в данной функции должно быть реализовано.

### 2.1.1. Модуль для кодирования речевого сигнала

Предполагается, что речевой сигнал уже получен, и мы имеем его в цифровом виде. Пользователю требуется указать, откуда считать полученный сигнал и вид кодера.

Создав объект типа Voice\_type, в него следует считать полученный речевой сигнал. Далее идёт вызов функций, которые переводят речевой сигнал в формат кодека. Это отдельные функции для вокодера и липридера (подробнее см. 3.3.1); для использования гибрида рекомендуется последовательно применить данные функции последовательно, или модифицировать функцию липридера – в таком случае можно будет обойтись вызовом только функций липридера.

После, используя функцию вызова канального кодирования, пользователь подвергает полученный речевой сигнал помехоустойчивому кодированию.

### 2.1.2. Модуль для декодирования речевого сигнала

При декодировании пользователь вызывает функцию синтезатора речи, в которую передаёт следующие параметры: тип кодека и массив, в который нужно поместить полученный с выхода канального декодера результат (номер фонемы, число нулей интенсивности и прочие параметры речевого сигнала), а также ID свёрточного и блочного кодера (см. 2.2.1).

Далее к указанному массиву обращается модуль, который воспроизводит полученную информацию (в данной работе этот модуль не рассматривается).

## 2.2. Модули для работы на канальном уровне

Часть модулей для работы на канальном уровне представлены в виде интерфейсов. В данном случае это сделано по следующим причинам:

* алгоритм хорошо известен и/или его написание не составляет особого труда;
* пользователю предлагается самому подобрать нужный кодирующий полином/образующую матрицу.

### 2.2.1. Модуль для канальных кодеков

Данный модуль служит для помехоустойчивого кодирования и декодирования речевого сигнала. По-умолчанию он предоставляет пользователю 3 описанных в данной работе кодека для блочного кодирования и 2 свёрточных кодека (см. 1.1 и 1.2).

За вызов функций данного модуля отвечает функция, которая принимает на вход тип речевого кодека, ID блочного кодека и ID свёрточного кодека.

Ниже приведены таблицы 2.1 и 2.2, в которых указано соответствие между ID кодеков и реализуемых ими алгоритмами.

Таблица 2.1.

**Соответствие между ID блочного кодека и его алгоритмом**

|  |  |
| --- | --- |
| **ID кодека** | **Реализованный алгоритм** |
| **1** | **2** |
| 1 | Проверка на чётность (двумерная) |
| 2 | Код Хэмминга (4, 7) |
| 3 | Циклический код (4, 7) |

Таблица 2.2.

**Соответствие между ID свёрточного кодека и его алгоритмом**

|  |  |
| --- | --- |
| **ID кодека** | **Реализованный алгоритм** |
| **1** | **2** |
| 1 | Ввод контрольных бит |
| 2 | Кодирование полиномом с задержкой |

При желании пользователь может добавить в библиотеку свои кодеки.

### 2.2.2. Модуль интерливинга и деинтерливинга

Данный модуль предназначен для перемежения речевого сигнала, который уже подвергся помехоустойчивому кодированию.

Функции данного модуля вызываются из модулей канальных кодеков: применение интерливинга без помехоустойчивого кодирования не имеет смысла! Данный модуль содержит две функции: интерливинга (на вход подаются биты, подвергнутые помехоустойчивому кодированию), и деинтерливинга (на вход подаются принятые из канала передачи данных биты).

После интерливинга данные отправляются в канал передачи. Из канала передачи принятые биты поступают в деинтерливер.

# 3. Разработка библиотек

Перед описанием процесса разработки стоит отметить, что подключение библиотек к проекту выполняется путём добавления .h и .c файлов в папку проекта.

## 3.1. Структура библиотек и описание

Можно выделить четыре основных заголовочных файла разработанных библиотек: это Channel\_coding.h, Analytic.h, Synthetic.h и Speech\_decoder.h

Первый файл содержит прототипы функций для канальных кодеков, а последующие три содержат прототипы функций для речевого кодека.

В библиотеке имеется одна сущность: Voice\_type. Она описывает свойства речевого сигнала. Выглядит она следующим образом:

**Листинг 1. Содержимое Voice\_type.h (описание сущности Voice\_type)**

#ifndef VOICE\_TYPE

#define VOICE\_TYPE

typedef struct

{

int amplitude;

int frequency;

int phoneme;

/\*

Иные характерные для речевого сигнала параметры:

-кратковременная энергия речевого сигнала;

-число нулей интенсивности (мгновенная частота);

-форманты речевого сигнала (концентрация энергии в ограниченной частотной области);

-коэффициенты линейного предсказания (ковариационный и автокорреляционный методы);

-распределение энергии сигнала по частотным группам;

-длительность пауз.

\*/

} Voice\_type;

#endif

## 3.2. Разработка канальной библиотеки

Разработка библиотеки проходит по следующему плану:

1. разработка одного из описанных в 1.1 и 1.2 методов;
2. сравнение полученной закодированной последовательности с результатом эталонной модели. В случае успеха переход к следующему пункту, в случае неудачи – приступаем к отладке;
3. разработка декодера для созданного метода кодирования;
4. тестирование декодера; в случае успеха переходим к реализации следующего метода, в случае неудачи – приступаем к отладке.

Пример листинга алгоритма на языке C, реализующий один из методов канального кодека (Хэмминг (4, 7), код ошибки совпадает с номером бита), приведён в приложении 1.

Следует заметить, что данная библиотека имеет 2 заголовочных файла: Channel\_coding.h и Interleaving.h. Их содержимое приведено в приложении 2.

### 3.2.1. Описание функций канальной библиотеки

Ниже приведена таблица 3.1, в которой описаны функции предлагаемой канальной библиотеки. Следует заметить: ID – идентификатор метода канального кодирования; подробнее см. 2.2.1.

Таблица 3.1.

**Соответствие между ID блочного кодека и его алгоритмом**

|  |  |  |
| --- | --- | --- |
| **Функция** | **Аргументы** | **Описание** |
| **1** | **2** | **3** |
| void bm\_ID | int to\_encode[], int out[] | Функция bm\_ID реализовывает блочное кодирование речевого сигнала согласно заданному алгоритму. |

Продолжение таблицы 3.1.

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
|  | int to\_encode[], int out[] | Данная функция не возвращает никаких результатов.  to\_encode – указатель на массив данных, которые нужно закодировать.  out – указатель на массив, куда нужно сложить результат.  Входит в заголовочный файл "Channel\_coding.h" |
| void cm\_ID | int to\_encode[], int out[] | Функция cm\_ID реализовывает свёрточное кодирование принятого речевого сигнала согласно заданному алгоритму.  Данная функция не возвращает никаких результатов.  to\_encode – указатель на массив данных, которые нужно закодировать.  out – указатель на массив, куда нужно сложить результат.  Входит в заголовочный файл "Channel\_coding.h" |
| void De\_bm\_ID | int to\_decode[], int out[] | Функция De\_bm\_ID реализовывает декодирование принятого блочного кода согласно заданному алгоритму.  Данная функция не возвращает никаких результатов.  to\_decode – указатель на массив данных, которые нужно декодировать.  out – указатель на массив, куда нужно сложить результат.  Входит в заголовочный файл "Channel\_coding.h" |

Продолжение таблицы 3.1.

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
| void De\_cm\_ID | int to\_decode[], int out[] | Функция De\_cm\_ID реализовывает декодирование принятого свёрточного кода согласно заданному алгоритму.  Данная функция не возвращает никаких результатов.  to\_decode – указатель на массив данных, которые нужно декодировать.  out – указатель на массив, куда нужно сложить результат.  Входит в заголовочный файл "Channel\_coding.h" |
| Interleaving | int to\_interleaving[] | Функция Interleaving реализовывает перемежение бит перед отправкой в канал связи.  Данная функция не возвращает никаких результатов.  to\_interleaving – указатель на массив данных, которые нужно переставить.  Входит в заголовочный файл "Interleaving.h" |
| DE\_interleaving | int to\_DEinterleaving[] | Функция DE\_interleaving реализовывает перемежение бит перед отправкой в канал связи.  Данная функция не возвращает никаких результатов.  to\_ DEinterleaving – указатель на массив данных, которые нужно переставить.  Входит в заголовочный файл "Interleaving.h" |

## 3.3. Разработка библиотеки для речевого сигнала

Данная библиотека представлена в виде интерфейсов и состоит из 6 файлов: 3 заголовочных и 3 – исходных кодов (Analytic, Synthetic, Speech\_decoder).

Рассмотрим процедуру кодирования речи. Она состоит из следующих шагов:

1. Создаём объект VT\_obj типа Voice\_type.
2. Получаем речевой сигнал (сторонний модуль, в данной работе не рассмотрен) и записываем его в VT\_obj.
3. Определяем тип кодека, который нам нужен.
4. В зависимости от типа кодека вызываем методы Phoneme\_number\_generator(2) или LPC(2), на вход которых подаём тип кодера и VT\_obj.
5. Вызывается метод get\_signal\_parameter(3). На его вход подаётся тип кодера, а также VT\_obj (третий параметр – значение VT\_obj на предыдущем шаге (если выбран липридер) или текущее значение VT\_obj(если выбран вокодер)).
6. В зависимости от типа кодека вызываются методы Compare\_with\_fixed\_dictionary(1)/Compare\_with\_adaptive\_dictionary(2), на вход которых подаётся VT\_obj (и его значение на предыдущем шаге – если выбран липридер).

Листинг заголовочных файлов, реализующих речевой кодек, приведён в приложении 3.

### 3.3.1. Описание функций библиотеки для речевой обработки сигнала

Далее приведена таблица 3.2, в которой описаны функции предлагаемой библиотеки для речевой обработки сигнала.

Таблица 3.2.

**Соответствие между ID блочного кодека и его алгоритмом**

|  |  |  |
| --- | --- | --- |
| **Функция** | **Аргументы** | **Описание** |
| **1** | **2** | **3** |
| Search\_pause | Voice\_type VT | Функция Search\_pause отвечает за поиск пауз в речи человека.  Данная функция возвращает true, если пауза найдена, и false в противном случае.  VT – полученный речевой сигнал.  Входит в заголовочный файл "Analytic.h" |
| LPC | char coder\_type,  Voice\_type VT | Функция LPC отвечает за реализацию алгоритма линейного предсказания.  Данная функция ничего не возвращает.  coder\_type – тип речевого кодека.  VT – полученный речевой сигнал.  Входит в заголовочный файл "Analytic.h" |
| Set\_current\_sound | Voice\_type VT | Функция Set\_current\_sound заполняет линию задержки.  Данная функция ничего не возвращает.  VT – полученный речевой сигнал.  Входит в заголовочный файл "Analytic.h" |

Продолжение таблицы 3.2.

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
| Get\_previous\_sound | - | Функция Get\_previous\_sound извлекает из линии задержки предыдущий речевой сигнал.  Данная функция возвращает предыдущий речевой сигнал (тип – Voice\_type).  Входит в заголовочный файл "Analytic.h" |
| Compare\_with\_adaptive\_dictionary | Voice\_type VT,  Voice\_type VT\_old | Функция Compare\_with\_adaptive\_dictionary сравнивает полученный речевой сигнал с адаптивным словарём.  Данная функция ничего не возвращает. Результат сравнения заносится в глобальную переменную NEW\_VOICE (тип – Voice\_type).  VT – полученный речевой сигнал.  VT\_old – предыдущий речевой сигнал.  Входит в заголовочный файл "Analytic.h" |

Продолжение таблицы 3.2.

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
| get\_signal\_parameter | char coder\_type,  Voice\_type VT1,  Voice\_type VT2 | Функция get\_signal\_parameter вытаскивает из полученного сигнала его параметры.  Данная функция ничего не возвращает.  coder\_type – вид речевого кодера  VT1, VT2 –речевые сигналы (определяются видом кодека).  Входит в заголовочный файл "Analytic.h" |
| Phoneme\_number\_generator | char coder\_type,  Voice\_type VT | Функция Phoneme\_number\_generator определяет номер фонемы по полученному сигналу.  Данная функция возвращает номер фонемы.  coder\_type – вид речевого кодера  VT – полученный речевой сигнал.  Входит в заголовочный файл "Analytic.h" |
| Save\_best\_phoneme\_id | int id | Функция Save\_best\_phoneme\_id сохраняет идентификатор фонемы лучшей подобранной фонемы на данный момен.  Данная функция ничего не возвращает.  id – идентификатор фонемы, который надо сохранить. Входит в заголовочный файл "Analytic.h" |

Продолжение таблицы 3.2.

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
| Compare\_with\_fixed\_dictionary | Voice\_type VT | Функция Compare\_with\_fixed\_dictionary сравнивает полученный речевой сигнал с фиксированным словарём.  Данная функция ничего не возвращает. Результат сравнения заносится в глобальную переменную BEST\_PHONEM\_ID (тип – int).  VT – полученный речевой сигнал.  Входит в заголовочный файл "Analytic.h" |
| Set\_filter\_parametres | Voice\_type VT | Функция Set\_filter\_parametres устанавливает текущие значения фильтра (используется для линейного предсказания).  Данная функция ничего не возвращает.  VT – полученный речевой сигнал.  Входит в заголовочный файл "Synthetic.h" |
| Set\_phoneme\_num | int phoneme\_num | Функция Set\_phoneme\_num устанавливает текущее значение идентификатора фонемы (используется для линейного предсказания).  Данная функция ничего не возвращает.  VT – полученный речевой сигнал.  Входит в заголовочный файл "Synthetic.h" |

Продолжение таблицы 3.2.

|  |  |  |
| --- | --- | --- |
| **1** | **2** | **3** |
| Code\_generator | char coder\_type,  int block\_code\_id,  int conv\_code\_id | Функция Code\_generator вызывает функции библиотеки канального кодирования.  Данная функция ничего не возвращает. Результат обработки данных кладётся в глобальную переменную CONV\_CODE[500]  block\_code\_id – идентификатор блочного кодера.  conv\_code\_id – идентификатор свёрточного кодера.  Входит в заголовочный файл "Synthetic.h" |
| get\_filter\_parameters | int input\_data[],  int block\_code\_id,  int conv\_code\_id | Функция get\_filter\_parameters определяет параметры речевого фильтра.  Данная функция возвращает полученные параметры речевого фильтра (тип – Voice\_type).  input\_data[] – массив принятых бит.  block\_code\_id – идентификатор блочного кодера.  conv\_code\_id – идентификатор свёрточного кодера.  Входит в заголовочный файл "Speech\_decoder.h" |

Продолжение таблицы 3.2.

|  |  |  |
| --- | --- | --- |
| get\_phoneme\_number | int input\_data[],  int block\_code\_id,  int conv\_code\_id | Функция get\_phoneme\_number определяет идентификатор фонемы.  Данная функция возвращает идентификатор фонемы (тип – int).  input\_data[] – массив принятых бит.  block\_code\_id – идентификатор блочного кодера.  conv\_code\_id – идентификатор свёрточного кодера.  Входит в заголовочный файл "Speech\_decoder.h" |
| get\_ALL | int input\_data[],  int block\_code\_id,  int conv\_code\_id | Функция get\_ALL определяет идентификатор фонемы и параметры речевого фильтра.  Данная функция возвращает речевой сигнал (тип – Voice\_type).  input\_data[] – массив принятых бит.  block\_code\_id – идентификатор блочного кодера.  conv\_code\_id – идентификатор свёрточного кодера.  Входит в заголовочный файл "Speech\_decoder.h" |

Продолжение таблицы 3.2.

|  |  |  |
| --- | --- | --- |
| made\_voice | char coder\_type,  int input\_data[],  int block\_code\_id,  int conv\_code\_id | Функция made\_voice восстанавливает речевой сигнал по полученным данным.  Данная функция возвращает речевой сигнал (тип – Voice\_type).  input\_data[] – массив принятых бит.  block\_code\_id – идентификатор блочного кодера.  conv\_code\_id – идентификатор свёрточного кодера.  Входит в заголовочный файл "Speech\_decoder.h" |
| channel\_decoding | int input\_data[],  int block\_code\_id,  int conv\_code\_id | Функция channel\_decoding декодирует принятые из канала передачи биты.  Данная функция возвращает речевой сигнал (тип – Voice\_type).  input\_data[] – массив принятых бит.  block\_code\_id – идентификатор блочного кодера.  conv\_code\_id – идентификатор свёрточного кодера.  Входит в заголовочный файл "Speech\_decoder.h" |

# Заключение

В результате выполнения работы были исследованы методы обработки речи для передачи по каналу связи.

На основании исследования были спроектированы и разработаны требуемые библиотеки:

1. для канального кодирования с кодеками:
   * блочными,
   * свёрточными.
2. для речевого кодирования.

Стоит отметить, что данная библиотека может считаться уникальной, т.к. аналогов в открытом доступе нет/мало.

# Список использованных источников

1. Избыточное кодирование информации: [Электронный документ]. (<http://all-ht.ru/inf/systems/p_0_11.html>). Проверено: 13.05.2016.
2. Код Хэмминга: [Электронный документ].

(<http://informkod.narod.ru/5_3item.htm>). Проверено: 13.05.2016.

1. Исправление ошибок с помощью циклических кодов: [Электронный документ].

(<http://ido.tsu.ru/iop_res1/kodi/index.php-mod=article&id=196.htm>). Проверено: 13.05.2016.

1. GEO-Mobile Radio Interface Specifications; Part 5: Radio interface physical layer specifications; Sub-part 3: Channel Coding: [Электронный документ].

(<http://www.etsi.org/deliver/etsi_ts/101300_101399/1013770503/01.01.01_60/ts_1013770503v010101p.pdf>). Проверено: 13.05.2016.

1. Способы задания сверточного кода: [Электронный документ].

(<http://sernam.ru/book_tec.php?id=95>). Проверено: 13.05.2016.

1. Алгоритм декодирования Витерби: [Электронный документ].

(<http://sernam.ru/book_tec.php?id=96>). Проверено: 13.05.2016.

1. Что такое вокодер и липридер?: [Электронный документ].

(<http://www.bnti.ru/showart.asp?aid=720&lvl=01.02.09>.). Проверено: 13.05.2016.

1. Кодирование и декодирование речевого сигнала: [Электронный документ].

(<http://www.sbi-telecom.ru/kodirovanie-i-dekodirovanie-rechevogo-signala.html>). Проверено 13.05.2016.

1. Речевой кодер с линейным предсказанием и использованием анализа через синтез: [Электронный документ].

(<http://www.freepatent.ru/patents/2163399>). Проверено 13.05.2016.

1. TIOBE Index: [Электронный документ].

(<http://www.tiobe.com/tiobe_index>). Проверено 16.05.2016.

1. Какой язык программирования будет наилучшим для изучения в 2015 году?: [Электронный документ].

(<https://habrahabr.ru/post/260797/>). Проверено 18.05.2016.

1. Основы программной инженерии. Жизненный цикл ПО. 2015: [Электронный документ].

(<http://kspt.icc.spbstu.ru/media/files/2015/course/se/SE2015_01_LifeCycle.pdf>). Проверено 16.05.2016.

# Приложение 1

Листинг 1. Метод кодирования Хэмминга (4, 7), код ошибки совпадает с номером ошибочного бита

|  |
| --- |
| void bm\_2(int to\_encode[], int out[]){  int to\_encode\_size, i,  bit\_1, bit\_2, bit\_3, bit\_4, bit\_5, bit\_6, bit\_7;  …  /\*  Алгоритм кодирования Хэмминга (4,7):  биты 3, 5, 6, 7 - информационные; биты 1, 2, 4 - контрольные.  Контрольные биты образуются через сумму по модуль 2 между:  1: 3, 5, 7 | 2: 3, 6, 7 | 4: 5, 6, 7)  \*/  for (i = 0; i < to\_encode\_size; i++){  printf("Data to ecnode: %d\n", to\_encode[i]);  //Информационные биты:  bit\_7 = (to\_encode[i] & 1);  bit\_6 = (to\_encode[i] & 2) >> 1;  bit\_5 = (to\_encode[i] & 4) >> 2;  bit\_3 = (to\_encode[i] & 8) >> 3;  //Контрольные биты:  bit\_1 = bit\_3 ^ bit\_5 ^ bit\_7;  bit\_2 = bit\_3 ^ bit\_6 ^ bit\_7;  bit\_4 = bit\_5 ^ bit\_6 ^ bit\_7;    //Заполняем результат:  out[i] = bit\_1 | (bit\_2 << 1) | (bit\_3 << 2) | (bit\_4 << 3)  | (bit\_5 << 4) | (bit\_6 << 5) | (bit\_7 << 6);  }  } |

Листинг 2. Метод декодирования Хэмминга (4, 7), код ошибки совпадает с номером ошибочного бита

void De\_bm\_2(int to\_decode[], int out[])

{

int to\_decode\_size, i,

bit\_1, bit\_2, bit\_3, bit\_4, bit\_5, bit\_6, bit\_7,

control\_bit\_1, control\_bit\_2, control\_bit\_3, control\_code = 0;

…

for (i = 0; i < to\_ decode \_size; i++){

//Разбиваем на биты:

bit\_1 = (to\_decode[i] & 1);

bit\_2 = (to to\_decode[i] & 2) >> 1;

bit\_3 = (to to\_decode[i] & 4) >> 2;

bit\_4 = (to to\_decode[i] & 8) >> 3;

bit\_5 = (to to\_decode[i] & 16) >> 4;

bit\_6 = (to to\_decode[i] & 32) >> 5;

bit\_7 = (to to\_decode[i] & 64) >> 6;

//Составляем контрольные биты

control\_bit\_1 = bit\_1 ^ bit\_3 ^ bit\_5 ^ bit\_7;

control\_bit\_2 = bit\_2 ^ bit\_3 ^ bit\_6 ^ bit\_7;

control\_bit\_3 = bit\_4 ^ bit\_5 ^ bit\_6 ^ bit\_7;

control\_code = control\_bit\_1 | (control\_bit\_2 << 1) | (control\_bit\_3 << 2);

if (control\_code == 0)

out[i] = bit\_3 | (bit\_5 << 2) | (bit\_6 << 1) | (bit\_7 << 3);

else{

//Повреждён только контрольный бит

if (control\_code == 1 || control\_code == 2 || control\_code == 4)

out[i] = bit\_3 | (bit\_5 << 2) | (bit\_6 << 1) | (bit\_7 << 3);

else{//Повреждён информационный бит.

switch (control\_code){

case 3:

bit\_3 = (~bit\_3 & 1);

break;

case 5:

bit\_5 = (~bit\_5 & 1);

break;

case 6:

bit\_6 = (~bit\_6 & 1);

break;

case 7:

bit\_7 = (~bit\_7 & 1);

break;

default:

break;

}

out[i] = bit\_3 | (bit\_5 << 2) | (bit\_6 << 1) | (bit\_7 << 3);

}

}

}

}

# Приложение 2

Листинг 3. Файл Channel\_coding.h

|  |
| --- |
| /\*  Данный файл содержит прототипы ф-ций для канальных кодеков.  \*/  #ifndef CHANNEL\_CODING  #define CHANNEL\_CODING  void bm\_1(int to\_encode[], int out[]);  void bm\_2(int to\_encode[], int out[]);  void bm\_2(int to\_encode[], int out[]);  void cm\_1(int to\_encode[], int out[]);  void cm\_2(int to\_encode[], int out[]);  void De\_bm\_1(int to\_decode[], int out[]);  void De\_bm\_2(int to\_decode[], int out[]);  void De\_bm\_3(int to\_decode[], int out[]);  void De\_cm\_1(int to\_decode[], int out[]);  void De\_cm\_2(int to\_decode[], int out[]);  #endif |

Листинг 4. Файл Interleaving.h

|  |
| --- |
| /\*  Данный файл содержит прототипы ф-ций для интерливинга.  \*/  #ifndef INTERLEAVING\_LIB  #define INTERLEAVING\_LIB  void Interleaving(int to\_interleaving[]);  void DE\_interleaving(int to\_DEinterleaving[]);  #endif |

# Приложение 3

Листинг 5. Файл Analytic.h

|  |
| --- |
| /\*  Данный файл содержит прототипы ф-ций аналитической части речевого кодера.  Тип кодера задаётся символом (char):  V - обычный вокодер (определяет и передаёт № фонемы).  L - липредер (вокодер с линейным предсказанием - снимает и передаёт  параметры фильтра).  H - гибридный вариант (определяет как № фонемы, так и параметры фильтра).  \*/  #ifndef ANALYTIC  #define ANALYTIC  #include "Voice\_type.h"  bool Search\_pause(Voice\_type VT);  Voice\_type LPC(char coder\_type, Voice\_type VT); //Linear Prediction  //Next 2 functions - for delay line:  void Set\_current\_sound(Voice\_type VT);  Voice\_type Get\_previous\_sound();  void Compare\_with\_adaptive\_dictionary(Voice\_type VT, Voice\_type VT\_old);  void get\_signal\_parameter(char coder\_type, Voice\_type VT1, Voice\_type VT2);  int Phoneme\_number\_generator(char coder\_type, Voice\_type VT);  void Save\_best\_phoneme\_id(int id);  void Compare\_with\_fixed\_dictionary(Voice\_type VT);  #endif |

Листинг 6. Файл Synthetic.h

|  |
| --- |
| /\*  Данный файл содержит прототипы ф-ций синтетической части речевого кодера.  Тип кодера задаётся символом (char):  V - обычный вокодер (определяет и передаёт № фонемы).  L - липредер (вокодер с линейным предсказанием - снимает и передаёт  параметры фильтра).  H - гибридный вариант (определяет как № фонемы, так и параметры фильтра).  \*/  #ifndef SYNTHETIC  #define SYNTHETIC  #include "Voice\_type.h"  #include "Channel\_coding.h"  void Set\_filter\_parametres(Voice\_type VT);  void Set\_phoneme\_num(int phoneme\_num);  //Используем библиотеку для канального кодирования:  void Code\_generator(char coder\_type, int block\_code\_id, int conv\_code\_id);  extern int CONV\_CODE[500];  extern int BLOCK\_CODE[100];  #endif |

Листинг 7. Файл Speech\_decoder.h

|  |
| --- |
| /\*  Данный файл содержит прототипы ф-ций для речевого декодера.  \*/  #include "Voice\_type.h"  #include "Channel\_coding.h"  #include "Interleaving.h"  #ifndef SPEECH\_DECODER  #define SPEECH\_DECODER  Voice\_type get\_filter\_parameters(int input\_data[], int block\_code\_id, int conv\_code\_id);  int get\_phoneme\_number(int input\_data[], int block\_code\_id, int conv\_code\_id);  Voice\_type made\_voice(char coder\_type, int input\_data[], int block\_code\_id, int conv\_code\_id);  Voice\_type get\_ALL(int input\_data[], int block\_code\_id, int conv\_code\_id);  Voice\_type channel\_decoding(int input\_data[], int block\_code\_id, int conv\_code\_id);  #endif |

# Приложение 4

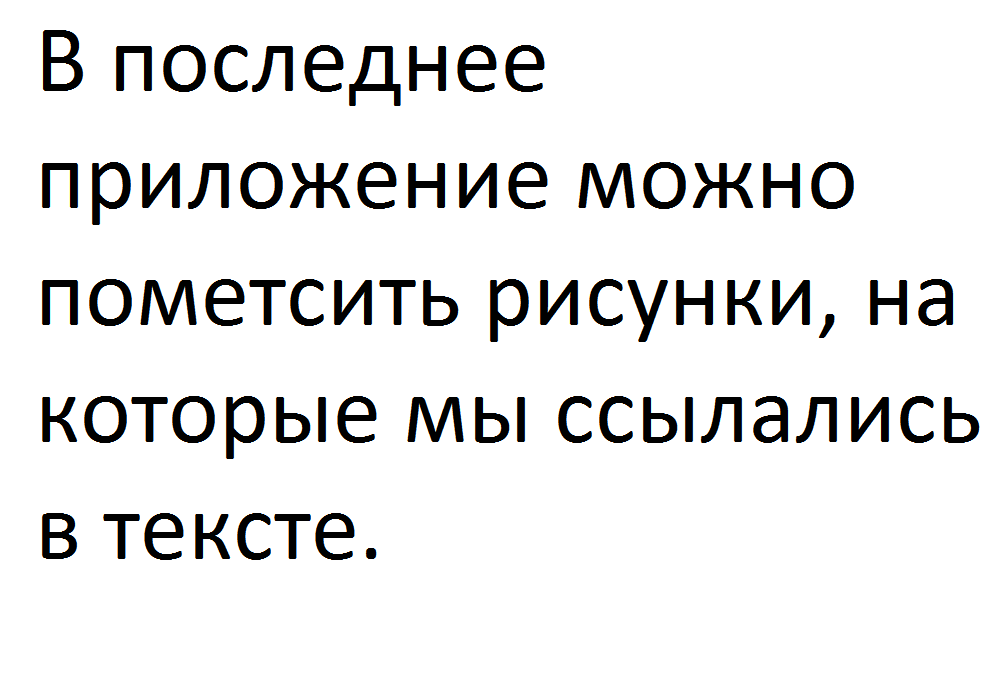
[](Бакалаврсая%20SPEECH.docx)

Рис. 4. Рисунок на который ссылаемся из текста

**Заключительный лист работы**

Выпускная работа выполнена мною самостоятельно. Использованные в работе материалы и концепции из опубликованной научной литературы и других источников имеют ссылки на них.

Список использованных источников содержит \_\_\_\_\_\_\_\_\_\_\_ наименований.

Работа выполнена на\_\_\_\_\_\_\_\_\_\_\_ страницах.

Приложения к работе на \_\_\_\_\_\_\_\_\_ страницах.

Один экземпляр сдан в директорат.

Подпись\_\_\_\_\_\_\_\_\_\_\_\_\_\_/\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_/

(фамилия, инициалы)

Дата «\_\_\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 20\_\_\_\_\_\_г.