

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 004.422.81

Райкевич  
Алёна Сергеевна

Программное средство классификации речи на основе кепстрального анализа

**АВТОРЕФЕРАТ**

на соискание степени магистра технических наук

по специальности 1-98 80 01 «Методы и системы защиты информации,  
информационная безопасность»

---

Научный руководитель  
Зельманский Олег Борисович  
кандидат технических наук, доцент

---

Минск 2020

## ВВЕДЕНИЕ

Устная речь сегодня остается самым оперативным и распространенным способом передачи информации в любой сфере человеческой деятельности, являясь основной формой выражения – намерений, целей, желаний. Факт того, насколько быстро развивается вычислительная техника и как стремительно она внедряется в жизнь современного человека, заставляет искать наиболее удобные формы общения между компьютером и непосредственным его пользователем. Бесспорно, самой простой и удобной формой для человека является речевой диалог на естественном языке. Использование речевых технологий в прикладных программах в качестве альтернативного средства взаимодействия в системе «человек-компьютер» приобретает все больший размах.

Одной из задач, которую необходимо решить при реализации речевого взаимодействия пользователя с системой является задача классификации речи. Классификация речи представляет собой процесс анализа фонетических единиц, выделенных в ходе сегментации речи, и определения соответствующих фонетических групп к которым они относятся.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Цели и задачи исследования**

Целью магистерской диссертации является разработка программного модуля классификации речи на основе кепстрального анализа.

В соответствии с поставленной целью, в работе сформулированы и решены следующие основные задачи:

- проведен анализ литературно-патентных источников информации;
- исследованы основные признаки, характеризующие речевой сигнал;
- рассмотрены существующие методы классификации речи;
- разработан программный модуль классификации речевых сигналов.

### **Связь с приоритетными направлениями научных исследований и запросами реального сектора экономики**

Тема диссертационной работы соответствует приоритетным направлениям фундаментальных и прикладных исследований Республики Беларусь в области информационной и инженерно-технической безопасности, создания современных систем защиты информации.

### **Личный вклад соискателя**

Содержание диссертации отображает личный вклад автора. Основные научные и практические результаты работы получены лично автором.

Определение целей и задач исследований, интерпретация и обобщение полученных результатов проводились с научным руководителем, кандидатом технических наук, доцентом О.Б. Зельманским.

### **Апробация результатов диссертации**

Основные положения и результаты, изложенные в диссертационной работе, докладывались и обсуждались на XVII Белорусско-российской научно – технической конференции «Технические средства защиты информации», Минск, 11 июня 2019 г., 55-ой юбилейной конференции аспирантов, магистрантов и студентов БГУИР, Минск, 22 – 26 апреля 2019 г., XVIII Белорусско-российской научно-технической конференции «Технические средства защиты информации», Минск, 9 июня 2020 г.

## **Опубликованность результатов диссертации**

По результатам исследований, представленных в диссертации, опубликовано 3 печатные работы в сборниках «Инфокоммуникации: 55-я юбилейная конференция аспирантов, магистрантов и студентов», «ТСЗИ 2019», «ТСЗИ 2020».

## **Структура и объём диссертации**

Диссертационная работа состоит из введения, общей характеристики работы, основной части из трёх разделов, заключения, списка использованных источников, списка собственных источников, семи приложений. Полный объём диссертационной работы составляет 76 страниц, включая 19 иллюстраций, список использованных источников из 32 наименований, список собственных источников из 3 наименований.

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

Во введении обоснована актуальность темы, сформулирована цель диссертации, изложены основные положения, выносимые на защиту.

В общей характеристике работы сформулированы ее цель и задачи, показана связь с научными программами и проектами, даны сведения об объекте исследования и обоснован его выбор, представлены положения, выносимые на защиту, приведены сведения о личном вкладе соискателя, апробации результатов диссертации и их опубликованность, а также, структура и объем диссертации.

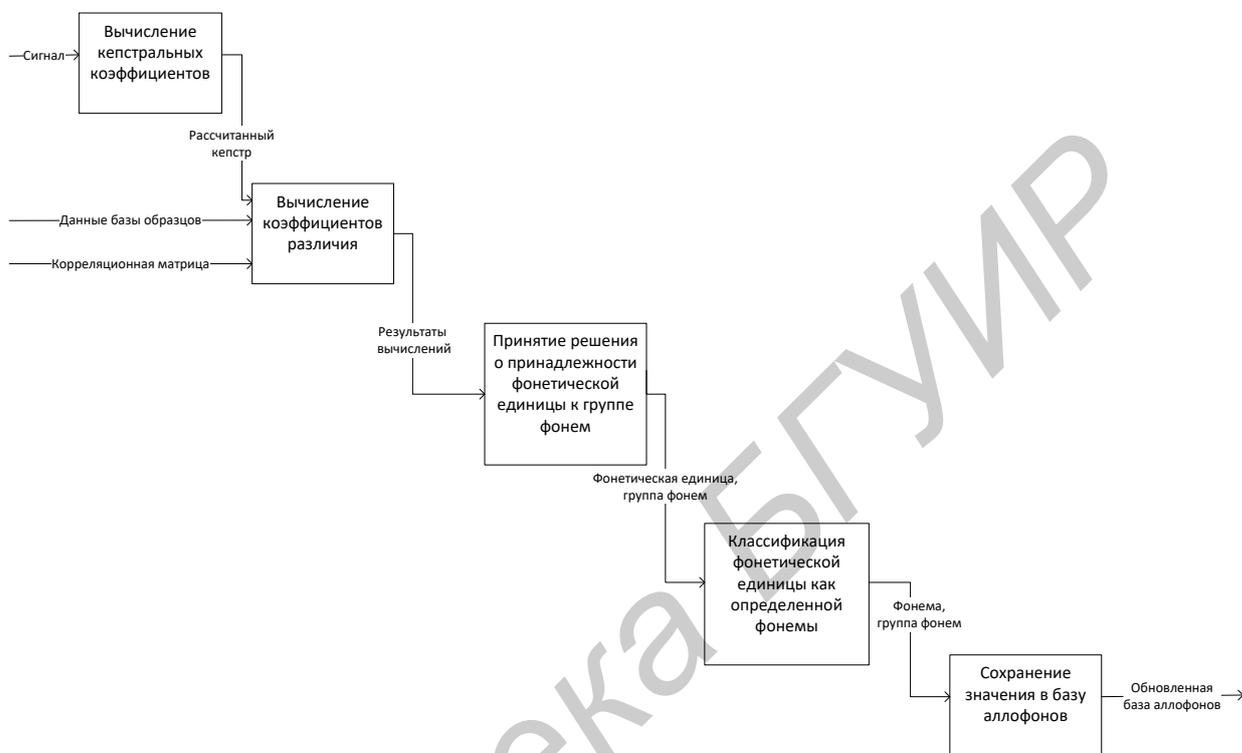
В первом разделе дано описание существующих алгоритмов классификации речи в составе систем распознавания речи. Также в данной главе поставлена цель, которую необходимо решить при рассмотрении модуля классификации речи на основе кепстрального анализа. Рассмотрены принципы классификации звуков алфавита, рассмотрены признаки, характеризующие речевой сигнал.

Во втором разделе проанализированы базовые подходы к классификации речи, основанные на скрытых марковских моделях и нейронных сетях. В ходе анализа выявлены преимущества и недостатки способов. Определены важнейшие характеристики текущего состояния речевого тракта человека.

Выделены основные методы кепстрального анализа речевых сигналов, применяемых для выделения из них векторов параметров.

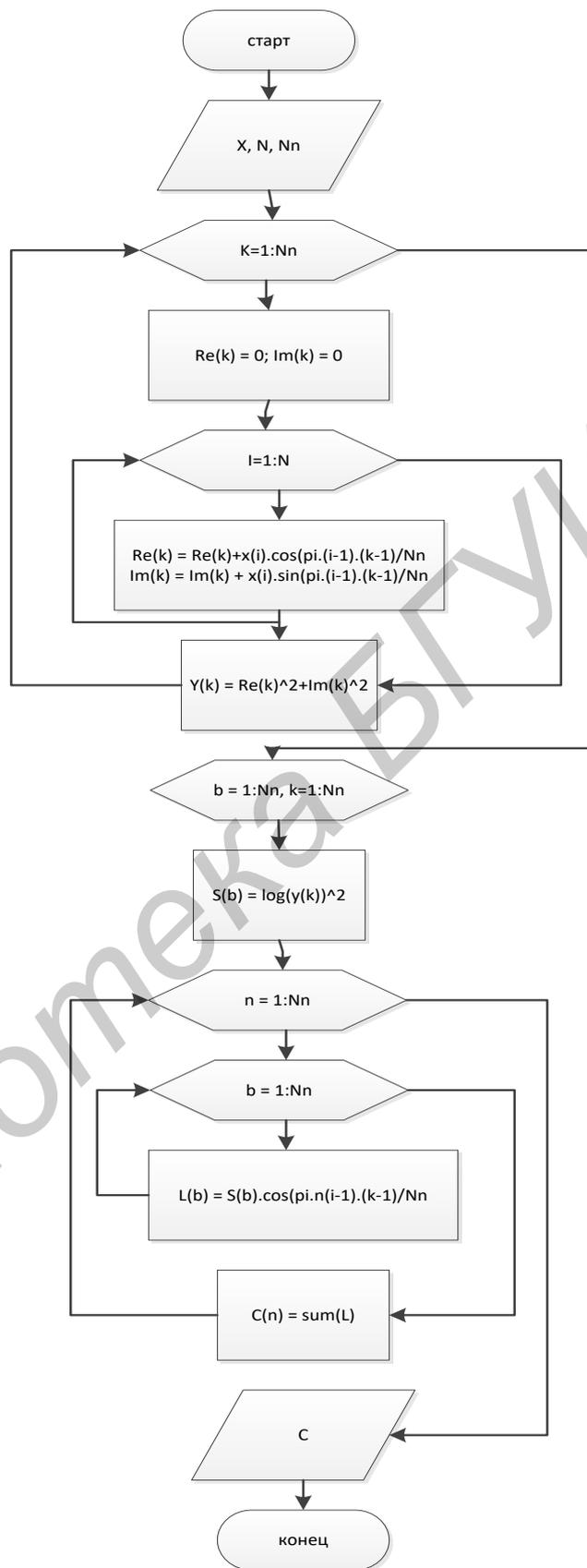
Третья глава посвящена описанию программного модуля и характеристике используемых методов работы с сигналами.

В ней описан программный модуль, включающий этапы обработки сигнала и классификации речи, схема которого представлена на рисунке 1.



**Рисунок 1 – Функциональная схема модуля классификации речи на основе кепстрального анализа**

Описана процедура выделения признаков – нахождение вектора мел-кепстральных коэффициентов, схема которого представлена на рисунке 2



**Рисунок 2 – Блок-схема вычисления кепстра речевых сигналов**

Кроме этого описаны цели и задачи, возникающие при составлении речевой базы, а также ее основные характеристики. В соответствии с целью

работы, с помощью разработанного программного модуля была сформирована речевая база.

В четвертой главе выполняется моделирование алгоритма классификации речи.

В ней приводится описание используемых функций среды MATLAB для моделирования алгоритма классификации речевых сигналов, а также приведена наглядная демонстрация результатов моделирования для признаков звуков русского языка.

## ЗАКЛЮЧЕНИЕ

В ходе выполнения работы было выявлено, что наиболее подходящими для задач распознавания речи признаками являются такие признаки, которые отражают свойства концентрации энергии и учитывают особенности восприятия слуха человека для получения адекватного природе речевого воздействия частотного разбиения. Данными признаками являются: мел-кепстральные коэффициенты и распределение информационных интервалов речевого сигнала. Проведен анализ особенностей обработки речевых сигналов в задачах распознавания речи: приведены концептуальные схемы распознавания, даны сведения о восприятии и воспроизведении звука человеком.

Проведен сравнительный анализ применения различных мер близости и векторов признаков который показал, что наиболее пригодными, для задач распознавания, векторами признаков можно считать: Мел-кепстральные коэффициенты; распределение информационных интервалов речевого сигнала.

В результате выполнения работы были изучены особенности обработки звукового сигнала и классификации речи. Данные знания помогли выявить сложности, с которыми приходится сталкиваться при разработке систем распознавания речи.

При разработке приложения было важно не просто перевести процесс в автоматизированную среду. Важно, что это приложение упрощает работу и повышает производительность системы. Кроме этого приложение является легко дополняемым и легко расширяемым, а это позволит адаптировать его и под другие задачи.

В результате разработки проекта была достигнута цель, т.е. был разработан программный модуль классификации речи на основе кепстрального анализа.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1–А. Райкевич, А.С. Программный модуль классификации речи на основе кепстрального анализа / А. С. Райкевич, О. Б. Зельманский // Технические средства защиты информации : тезисы докладов XVII Белорусско-российской научно – технической конференции, Минск, 11 июня 2019 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол. : Т. В. Борботько [и др.]. – Минск, 2019. – С. 59.

2–А. Райкевич, А.С. Программное средство классификации речи на основе кепстрального анализа / А. С. Райкевич, Ю.Н. Никитенко, О. Б. Зельманский // Инфокоммуникации: 55-я юбилейная конференция аспирантов, магистрантов и студентов, Минск, 22 – 26 апреля 2019 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2019. – С. 85.

3–А. Райкевич, А.С. Программный модуль классификации речи / А. С. Райкевич // Технические средства защиты информации : тезисы докладов XVIII Белорусско-российской научно – технической конференции, Минск, 9 июня 2020 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол. : Т. В. Борботько [и др.]. – Минск, 2020. – С. 65.