

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

УДК 37-1174

Цык
Антон Владимирович

**АНАЛИЗ ХАРАКТЕРИСТИК РАСПОЗНАВАНИЯ РЕЧИ
КОМПЬЮТЕРНЫХ ПРИЛОЖЕНИЙ**

Автореферат
на соискание степени магистра
по специальности 1–45 80 01 Системы и сети инфокоммуникаций

Научный руководитель
доцент кафедры ИТ,
Митюхин Анатолий Иванович

Минск 2022

ВВЕДЕНИЕ

В белорусском обществе постепенно осознается, что интеграция (включение) людей с физическими и другими отклонениями в жизнь общества возможна только путем непрерывного образования и профессиональной подготовки на базе концепции безбарьерного доступа к образованию и профессии. Это понимание согласуется со статьей 24 Конвенции ООН о правах инвалидов (2008) относительно общего образования: «Никто не должен быть исключен из системы общего образования». Естественным продолжением процесса инклюзии людей с особыми потребностями в современное общество является включение этих людей и в систему высшего образования.

В условиях объективного ограничения коммутативных возможностей у студентов с ослабленным слухом особенно трудно идет изучение специальных технических дисциплин на начальном этапе. Возникают большие сложности при передаче, записи технической, научной терминологии, и пр. Основная проблема здесь – правильный и однозначный перевод специальной терминологии изучаемой дисциплины.

На сегодняшний день существует большая потребность в усовершенствовании методов образования для людей с ограниченными возможностями.

Система образования в Беларуси должна быть адаптируемой. Интеграция означает, что на всех этапах получения образования глухие и не глухие люди учатся вместе.

В связи с наличием в Беларуси кодекса социального обеспечения инвалидов и принятием Конвенции ООН о правах лиц с инвалидностью потребность в усовершенствовании процесса обучения возрастает с каждым годом.

По данным Всемирной организацией здравоохранения (ВОЗ), более 5% населения мира – 466 миллионов человек. 1,1 миллиарда молодых людей в возрасте 12–35 лет подвергаются риску потери слуха из-за воздействия шума. Для общения такие люди используют язык жестов.

Как следует из публикуемого сегодня Всемирной организацией здравоохранения первого Всемирного доклада по проблемам слуха, к 2050 г. почти 2,5 миллиарда человек во всем мире – то есть каждый четвертый – будут страдать той или иной степенью потери слуха. Если не будут приняты меры, по меньшей мере 700 миллионов из них будут нуждаться в услугах по охране здоровья уха и органов слуха и других реабилитационных услугах.

Несмотря на то, что вычислительные мощности растут с каждым днем, создание систем распознавания речи остается сложной задачей. Это

обуславливается как ее междисциплинарным характером (филологии, лингвистике, цифровой обработке сигналов, акустике, статистике, распознавании образов и т.д.), так и высокой вычислительной сложностью алгоритмов. Существенные ограничения на системы автоматического распознавания речи оказывает объем обрабатываемого словаря, скорость получения ответа и его точность.

В диссертационной работе рассматриваются наиболее известные системы и методы распознавания речи. Производится анализ характеристик систем распознавания речи и рассматриваются существующие системы распознавания речи для использования в процессе обучения людей с ограниченными возможностями, в частности, для людей с нарушением слуха. Чтобы оценить возможность внедрения таких систем в учебный процесс проводятся экспериментальные исследования, на основе которых оцениваются точностные характеристики. Предложена методика экспериментальных исследований оценки точностных характеристик и распознавания речевого сигнала. В диссертации рассмотрен вопрос технического обеспечения образовательного процесса для людей с нарушением слуха.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с крупными научными программами

Тема диссертационной работы соответствует пункту 1 приоритетных направлений научной, научно-технической и инновационной деятельности Республики Беларусь на 2021–2025 гг., утвержденных Указом Президента Республики Беларусь №156 от 7 мая 2020 г. «Цифровые информационно-коммуникационные и междисциплинарные технологии, основанные на них производства». Работа выполнялась в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники».

Цель и задачи исследования

Целью диссертационной работы является анализ существующих систем распознавания речи для использования в процессе обучения людей с ограниченными возможностями, в частности, для людей с нарушением слуха. Актуальность темы подтверждается резолюциями ЮНЕСКО в области образовании ООН.

Задачи диссертационной работы сводилось к анализу современных систем распознавания речи, их возможностей, оценки точностных характеристик.

Для достижения поставленной цели в диссертации решены следующие задачи:

- 1 Рассмотрены наиболее известные системы и методы распознавания речи.
- 2 Разработана методика оценки систем распознавания речи.
- 3 Сделан анализ системных характеристик систем распознавания речи.
- 4 Произведена оценка точностных характеристик известных систем для распознавания речи

Личный вклад соискателя ученой степени

Личный вклад автора заключается исследовательской работе на тему «Анализ характеристик распознавания речи компьютерных приложений», постановке и проведении экспериментов по исследованию точностных характеристик современных систем распознавания речи, а также в предложенной методике по внедрению современных систем распознавания речи в учебный процесс для людей с ограниченными возможностями слуха.

Определение целей и задач исследований, интерпретация и обобщение полученных результатов проводились совместно с научным руководителем, доцент А.И. Митюхин.

Апробация диссертации и информация об использовании ее результатов

Основные положения и результаты диссертационной работы докладывались и обсуждались на конференциях:

- 1 54-й научной конференции аспирантов, магистрантов и студентов, Минск, 23–27 апреля 2018
- 2 XX Белорусско-российская научно-техническая конференция, Минск, 7 июня 2022 г.

Опубликование результатов диссертации

По результатам исследований, представленных в диссертации, опубликовано 2 тезиса конференций.

Структура и объем диссертации

Диссертационная работа состоит из введения, общей характеристики работы, трех глав с выводами по каждой главе, заключения, библиографического списка.

Общий объем диссертационной работы составляет 77 страниц, из них 50 страниц текста, 17 рисунков на 9 страницах, 19 таблиц на 8 страницах, список использованных библиографических источников (41 наименование на

3 страницах), список публикаций автора по теме диссертации (2 наименования на 1 странице), графический материал на 7 страницах.

Проверка на уникальность

Проведена экспертиза диссертации Цыка Антона Владимировича «Анализ характеристик распознавания речи компьютерных приложений» на корректность использования заимствованных материалов с применением сетевого ресурса «Антиплагиат» (адрес доступа: <https://antiplagiat.ru>) в on-line режиме 09.06.2022 г. В результате проверки установлена корректность использования заимствованных материалов (оригинальность диссертационной работы составляет 71 %).

Библиотека БГУИР

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первая глава была посвящена обзору задачи распознавания речи и путей ее решения. В этой главе:

1 рассмотрен объект исследования – речь, а также способы ее представления в цифровой форме с помощью осциллограмм и спектрограмм. Приведена классификация фонем русского языка;

2 поставлена задача распознавания речи и описана общая структура системы автоматического распознавания речи;

3 приведена классификация и описание различных подходов составления векторов признаков, работающих как в частотной области так и во временной;

Во второй главе был сделан анализ задачи предварительной обработки речевого сигнала и выделения признаков. В ней:

1 рассмотрена процедура нормализации, позволяющая снизить разброс громкости речи;

2 рассмотрен метод выделения речи на основе анализа кратковременной энергии и частоты проходов через ноль входного сигнала;

3 описана процедура выделения признаков – нахождение вектора мел-кепстральных коэффициентов;

4 рассмотрен метод проектирования КИХ фильтра с помощью окон.

В третьей главе были проведены экспериментальные исследования точностных характеристик наиболее известных существующих систем распознавания речи. Для эксперимента была предложена методика оценки точности распознавания, использующая простые слова, слова научно-технического характера и текст (отрывок лекции).

Исходя из полученных результатов исследования сделаны следующие выводы:

1 Точность распознавания речи для обычных слов достаточно велика и составила 100%;

2 Точность распознавания для более сложных слов, особенно терминов технического характера, находится в пределах 89 – 93.2%.

3 Результаты научных исследований по диссертации могут найти применения как для обучения людей с ограничениями по слуху, так и в общем для модернизации образовательного процесса студентов.

ЗАКЛЮЧЕНИЕ

В диссертационной работе проведены исследования связанные с возможностью использования современных систем для распознавания речи для обучения людей с ограниченными возможностями слуха. Предложена

методика оценки точности систем распознавания речи на основе простых слов, научно-технических слов и текста.

Результаты научных исследований показывают, что точность распознавания обычных слов составляет около 100%. Для терминов технического характера точность находится в пределах 89–93.2%.

Показано, что современные системы для распознавания речи таких компаний как Yandex, Microsoft и Google позволяют реализовывать учебный процесс на современном уровне для людей с ограничениями по слуху.

На основе анализа литературы по теме диссертации, выбора и обоснования методов распознавания, проведения экспериментальных исследований, получены результаты, подтверждающие возможность внедрять подобные системы для повышения качества учебного процесса.

В диссертации рассмотрен вопрос технического обеспечения образовательного процесса для людей с нарушением слуха.

В результате исследовательской работы возникают и другие вопросы, которые подлежат дальнейшему исследованию.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

Тезисы конференций

1–А. Митюхин, А. И. Алгоритм кодирования в зашумленном широкополосном канале / Митюхин А.И., Цык А.В. // XX Белорусско-российская научно-техническая конференция: тезисы докладов XX Белорусско-российской научно-технической конференции. Минск, 7 июня 2022 г. / редкол. : Т. В. Борботько [и др.]. – Минск : БГУИР, 2022. –С. 72.

2–А. Цык, А. В. Измерительный приемник для контроля каналов связи по линиям электропередачи / А. В. Цык // Инфокоммуникации: материалы 54-й научной конференции аспирантов, магистрантов и студентов, Минск, 23–27 апреля 2018 г. – Минск: БГУИР, 2018. – С. 86.