

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

УДК \_\_\_\_\_

Кручок  
Денис Николаевич

Психоакустически мотивированный алгоритм экстракции  
характеристического вектора для идентификации диктора в акустических  
шумах

**АВТОРЕФЕРАТ**

на соискание степени магистра технических наук  
по специальности 1-40 80 01 Элементы и устройства вычислительной  
техники и систем управления

Научный руководитель  
Петровский Александр Александрович  
доктор технических наук, профессор

Минск 2015

## ВВЕДЕНИЕ

Разработка и использование алгоритмов экстракции характеристического вектора для идентификации диктора, основанных на психоакустике, позволяет повысить точность систем распознавания и их взломоустойчивость, использовать системы распознавания в средах, содержащих посторонние звуки (акустические шумы) и при этом обеспечивать эффективность их работы на требуемом уровне.

Целью магистерской работы является разработка алгоритма экстракции характеристического вектора, который обеспечивает идентификацию диктора в акустических шумах.

Задачи исследования:

- Изучить существующие подходы и методы экстракции характеристического вектора, сравнить их между собой.
- Разработать психоакустически мотивированный алгоритм экстракции характеристического вектора для идентификации диктора в акустических шумах.
- Спроектировать и реализовать программный модуль идентификации диктора с использованием разработанного алгоритма.
- Сравнить эффективность работы алгоритма с эффективностью существующих методов в различных окружающих обстановках (без шумов и с акустическими шумами).
- Провести анализ полученных результатов.

Объектом исследования являются системы идентификации диктора в акустических шумах. Предметом исследования выступают алгоритмы и методы экстракции характеристического вектора для идентификации диктора в акустических шумах.

На качество работы систем распознавания влияют характеристические признаки и решающие правила. В большинстве работ по распознаванию диктора используются параметры в виде коэффициентов кепстра. В дополнение к коэффициентам кепстра используются также их первые и вторые разности по времени. Преимущество такого подхода заключается в вычислительной простоте, а также в том, что в кепстре отражаются индивидуальные характеристики голосового источника и анатомия речевого тракта. Как показывает практика, системы распознавания на основе кепстральных коэффициентов демонстрируют высокие результаты при работе в местах с хорошим акустическим окружением и когда акустическая среда обучаемого и тестируемого множества совпадает. Но в случае неблагоприятной обстановки (наличие акустических шумов) или различий

обстановки окружающего пространства при обучении и тестировании, эффективность работы систем распознавания сразу начинает падать, и их дальнейшее использование становится невозможным.

Поэтому разработка психоакустически мотивированных алгоритмов и методов экстракции характеристического вектора для идентификации диктора в акустических шумах является актуальной задачей в области систем распознавания.

Библиотека БГУИР

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

На сегодняшний день, особое внимание уделяется разработке и применению систем распознавания по голосу. Речевые технологии находят все большее распространение во сферах жизни человека: охранные системы и системы безопасности, банковские операции, концепция «умный дом», робототехника, системы управления оборудованием, средствах телекоммуникаций. Хотя и эффективность работы систем распознавания в закрытых или помещениях с благоприятной акустической обстановкой находится на довольно высоком уровне, результаты распознавания данных систем на улице или в местах с наличием акустических шумов ещё далеки от требуемых. Например, пожилому человеку сложно разглядеть и набрать код домофона при входе в подъезд, поэтому его идентификация по голосу предпочтительнее. Или при снятии наличных в банкомате, подтверждение личности пользователя происходило не только с использованием однофакторной методологии аутентификации личности (ПИН-код или буквенно-цифровой пароль), но еще и по голосу, что увеличит безопасность снятия денежных средств. Все эти случаи предполагают, что процесс распознавания будет происходить в обстановке с акустическими шумами.

В работе предлагается психоакустически мотивированный алгоритм извлечения характеристических признаков. Данный алгоритм позволит проводить идентификацию в агрессивных средах, которые содержат посторонние шумы, и при этом точность распознавания такой системы будет находиться на высоком уровне.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В диссертации была рассмотрена проблема эффективной идентификации в условиях акустических шумов. Для её решения был предложен психоакустически мотивированный алгоритм извлечения признаков, который моделирует процесс обработки речевого сигнала человеческим ухом. Также для проверки предложенного алгоритма была спроектирована система идентификации диктора на основе нейронных сетей прямого распространения. Система идентификации и предлагаемый алгоритм реализованы программно в среде моделирования Matlab.

В первой главе представлены общие сведения о распознавании речи и диктора. Подробно описан процесс идентификации диктора, а также приведена общая схема систем идентификации диктора. Дана классификация методов, обеспечивающих идентификацию диктора в акустических шумах.

Во второй главе были рассмотрены методы и подходы, которые используются для получения характеристических признаков речевого сигнала, рассмотрены их ключевые особенности. Подробно рассмотрен алгоритм получения широко используемых мелчастотных кепстральных коэффициентов.

В третьей главе были описаны психоакустически мотивированный алгоритм получения характеристических признаков, основанный на использовании слухового преобразования. Проведён сравнительный анализ признаков, полученных на базе слухового преобразования и преобразования Фурье.

В четвёртой главе выполнен обзор решающих правил, используемых в области распознавания, подробно рассмотрены искусственные нейронные сети. Описан выбор архитектуры нейронной сети и приведена схема системы идентификации диктора. Представлена оценка временных затрат для обучения используемой нейронной сети.

В пятой главе представлены результаты проведённых экспериментов по применению психоакустически мотивированного алгоритма экстракции характеристического вектора и приведён их анализ.

Таким образом, в ходе проделанной работы был предложен психоакустически мотивированный алгоритм извлечения характеристических признаков, была спроектирована система идентификации на основе нейронной сети прямого распространения и была проведена экспериментальная оценка предложенного алгоритма.

## ЗАКЛЮЧЕНИЕ

В данной работе были рассмотрены главные особенности основных методов и алгоритмов, которые используются для извлечения признаков сигнала с целью их дальнейшего использования при распознавании диктора в акустических шумах. Подробно рассмотрены два основных алгоритма экстракции характеристического вектора МЧКК и КККФ. Выполнен сравнительный анализ преобразований, которые используются в этих алгоритмах – преобразование Фурье и слуховое преобразование.

Был реализован психоакустически мотивированный алгоритм, основанном на слуховом преобразовании, а также алгоритм извлечения мелчастотных кепстральных коэффициентов с использованием преобразования Фурье. Для проверки данных алгоритмов была спроектирована система идентификации дикторов (для 4-х дикторов) на основе нейронных сетей.

В результате проведенных экспериментов, алгоритм, основанный на слуховом преобразовании, показал лучшую точность распознавания в акустических шумах, в сравнении с широко используемым алгоритмом извлечения мелчастотных кепстральных коэффициентов (идентификация в акустических шумах). В среднем, результаты составили 80% для признаков КККФ к 72% для признаков МЧКК для обучающего множества и 69% для признаков КККФ и 66% для признаков МЧКК для тестового множества.

Все поставленные задачи были успешно реализованы.

В дальнейшем планируется улучшить алгоритм извлечения признаков использовать различные функции, моделирующие поведение волосковых клеток), а также увеличить число дикторов для системы идентификации.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

[1] Кручок, Д.Н. Слуховое преобразование. Его использование в идентификации диктора / Кручок Д.Н. // Тезисы докладов 51 научно-технической конференции студентов и магистрантов – Минск, УО «БГУИР» – 2015.

[2] Кручок, Д.Н. Психоакустически мотивированный алгоритм экстракции характеристического вектора для идентификации диктора в акустических шумах / Кручок Д.Н. // Тезисы докладов 51 научно-технической конференции студентов и магистрантов – Минск, УО «БГУИР» – 2015.

[3] Кручок, Д.Н. Особенности однопользовательских компьютерных игр и их влияние на мировоззрение современного человека / Кручок Д.Н. // Тезисы докладов 51 научно-технической конференции студентов и магистрантов – Минск, УО «БГУИР» – 2015.

Библиотека БГУИР