

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

УДК _____

Цвирко
Александра Игоревна

Алгоритмы анализа речевого сигнала для определения эмоционального
состояния субъекта

АВТОРЕФЕРАТ

на соискание степени магистра технических наук
по специальности 1-40 80 02 «Системный анализ, управление и обработка
информации»

Научный руководитель
Свито Игорь Леонтьевич
кандидат технических наук,
доцент

Минск 2020

ВВЕДЕНИЕ

Распознавание эмоционального состояния человека в данный момент является актуальной темой и может быть использовано во многих отраслях таких как медицина, психология, маркетинг, обеспечение безопасности. Более того, разрабатываемые здесь подходы находят свое применение не только в сфере анализа эмоционального состояния, но и при распознавании других состояний, например - алкогольной интоксикации, усталости, подавленности и т. п.

Распознавание эмоций по голосу находит применение во множестве областей: разработке социальных вспомогательных роботов, автономных транспортных средств, оборудовании для нейро-обратной связи и т. д. Однако в настоящий момент не существует достаточно эффективного решения данной задачи, хотя эмоции универсальны, их понимание и интерпретация являются специфическими и частично культурными.

Целью диссертации является исследование существующих и разработка новых алгоритмов, которые позволят увеличить точность определения эмоционального состояния субъекта по речевому сигналу.

Задачи:

1. Рассмотреть существующие системы распознавания эмоционального состояния по речевому сигналу;
2. Проанализировать существующие модели и методы определения эмоционального состояния по речевому сигналу и выбрать один из них для дальнейшего исследования;
3. Рассмотреть выбранный метод и разработать подход, на его основе.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования

Распознавание эмоционального состояния человека в данный момент является актуальной темой и может быть использовано во многих отраслях таких как медицина, психология, маркетинг, обеспечение безопасности. Более того, разрабатываемые здесь подходы находят свое применение не только в сфере анализа эмоционального состояния, но и при распознавании других состояний, например - алкогольной интоксикации, усталости, подавленности и т. п. Самой актуальной задачей является, задача распознавания эмоционального состояния по речевому сигналу, так как речевой сигнал занимает значительную часть в распознавании эмоций.

Цель исследования

Целью диссертационной работы является исследование существующих и разработка новых алгоритмов, которые позволят увеличить точность определения эмоционального состояния субъекта по речевому сигналу.

Задачи исследования

1. Обзор существующих систем распознавания эмоционального состояния по речевому сигналу;
2. Анализ существующих моделей и методов определения эмоционального состояния по речевому сигналу и выбор одного из них для дальнейшего исследования;
3. Исследование выбранного метода и разработка подхода, на его основе.

Личный вклад соискателя

Соискателем выполнены все изложенные в работе разработки и исследования. Постановка задач и обсуждение результатов проводились совместно с научным руководителем и сотрудниками кафедры информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники.

Соавторы опубликованных работ принимали участие в обсуждении промежуточных и конечных результатов. Обработка, интерпретация данных, а также выводы сделаны автором самостоятельно.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались на следующих научных конференциях:

- Исследование влияния параметров нейронной сети на процесс обучения однослойного персептрона;
- Алгоритмы анализа речевого сигнала для определения эмоционального состояния субъекта.

Библиотека БГУМР

СОДЕРЖАНИЕ РАБОТЫ

Распознавание эмоций в речевом сигнале относится к типу задач, для которых важно учитывать то, как акустический сигнал воспринимается человеком.

В первой главе диссертационной работы представлен обзор актуальных проблем и методов в задаче определения эмоционального состояния по голосу.

Во второй главе производится обзор параметров и классификация систем распознавания речи. Приводятся основные признаки классификации систем и различия методов распознавания речи.

В третьей главе производится обзор и классификация методов обработки речевых сигналов в системах распознавания речи. Рассмотрены базовые методы обработки речевых сигналов, сверточные нейронные сети, а также скрытые марковские модели для задач распознавания речи.

В четвертой главе представлена разработка комбинированного подхода кластеризации и классификации для задачи распознавания эмоций по речи. Разработан подход по решению задачи распознавания эмоций по акустическим характеристикам.

ЗАКЛЮЧЕНИЕ

В процессе диссертации был разработан и исследован алгоритм распознавания эмоционального состояния человека по речевому сигналу, основанный на комбинированном подходе кластеризации и классификации, а также проведен анализ существующих решений и методов определения эмоционального состояния по голосу.

Результаты анализа существующих решений указывают на то, что для соответствующих типов систем распознавания эмоций человека по голосу существуют свои широко и узко профильные сферы применения.

На основе анализа моделей и методов определения эмоционального состояния по речевому сигналу, был выбран комбинированный подход кластеризации и классификации.

В результате разработки и исследования алгоритма, основанного на комбинированном подходе кластеризации и классификации, были выявлены его недостатки и предложены решения по улучшению системы.

Применение предварительной кластеризации данных для проведения их классификации было проведено с помощью алгоритмов fuzzy c-means и k-means. Было показано, что использование метода кластеризации fuzzy c-means дает лучшие значения средней точности, чем при применении метода k-means, во всех случаях.

Статистически значимое улучшение применения предварительной кластеризации было получено только в одном случае: fuzzy c-means для 3 классов для алгоритма классификации SMO.

Статистически значимое улучшение применения отбора информативных признаков как при алгоритме кластеризации k-means, так и для fuzzy c-means получено только при алгоритме классификации k-NN ($k = 2$) для трех классов.

Разрабатываемый алгоритм позволяет строить гибкие и эффективные модели.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

[1] Цвирко А.И., Минчуков А.Э. Исследование влияния параметров сети на процесс обучения однослойного персептрона / 55-я научная конференция аспирантов, магистрантов и студентов учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» – Минск, 2019. – С. 30.

[2] Цвирко А.И., Минчуков А.Э. Алгоритмы анализа речевого сигнала для определения эмоционального состояния субъекта / 55-я научная конференция аспирантов, магистрантов и студентов учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» – Минск, 2019. – С. 114.

Библиотека БГУИР