

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

УДК

Дубровский
Павел Сергеевич

**АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ
СУБЪЕКТА НА ОСНОВЕ РЕЧЕВОГО СИГНАЛА**

АВТОРЕФЕРАТ

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

Научный руководитель

В.А. Захарьев,
кандидат технических наук,
доцент

Минск 2018

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.

Методы распознавания эмоций используют различные типы входных данных, т.е. выражение лица, речь, жесты и языка тела, а также физические сигналы, такие как электрокардиограммы (ЭКГ), электромиографии (ЭМГ), электродермальная активность, температура кожи, гальваническое сопротивление, пульс и дыхание.

Цель: исследование существующих и разработка новых алгоритмов, которые позволят увеличить точность определения эмоционального состояния субъекта по речевому сигналу.

Объект исследования: эмоциональное состояние субъекта.

Предмет: процесс определения эмоционального состояния по речевому сигналу с использованием исходных средств.

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 алгоритм, 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 алгоритма, на его основе.

Новизна полученных результатов

Разработан двухэтапный алгоритм распознавания эмоционального состояния человека по речевому сигналу.

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 систем управления 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:

– The International Conference on Information Technologies and Systems ITS 2017 (Минск 2017);

СОДЕРЖАНИЕ РАБОТЫ

В первой главе диссертационной работы производится обзор существующих разработок в области распознавания эмоционального состояния по речевому сигналу.

Системы распознавания эмоционального состояния по речевому сигналу делятся на 2 типа: одноканальные и многоканальные.

У первого типа преимущество в мобильности использования, экономичности, удобстве использования, быстродействии, безопасности. Однако одноканальные системы сложны в реализации и ресурсоемки.

У второго типа преимущество в чувствительности, надежности, эффективности, функциональности, достоверности результатов.

В разрабатываемой системе требуются быстродействие, удобство использования и безопасность, поэтому для дальнейшего исследования были выбраны одноканальные системы.

Во второй главе диссертационной работы производится анализ существующих моделей и методов определения эмоционального состояния человека по речевому сигналу. Приводится подробный анализ: марковских моделей, метода опорных векторов и методов нелинейной динамики. Результаты анализа моделей и методов определения эмоционального состояния по речевому сигналу показали, что выбор метода, зависит от начальных условий. В настоящее время широко развиваются методы, основанные на нейронных сетях. Данный подход был выбран в качестве основы для построения собственного алгоритма.

В третьей главе диссертационной работы производится разработка двухэтапного алгоритма распознавания эмоционального состояния, на основе нейронной сети. Схема данного алгоритма приведена на рисунке 1.

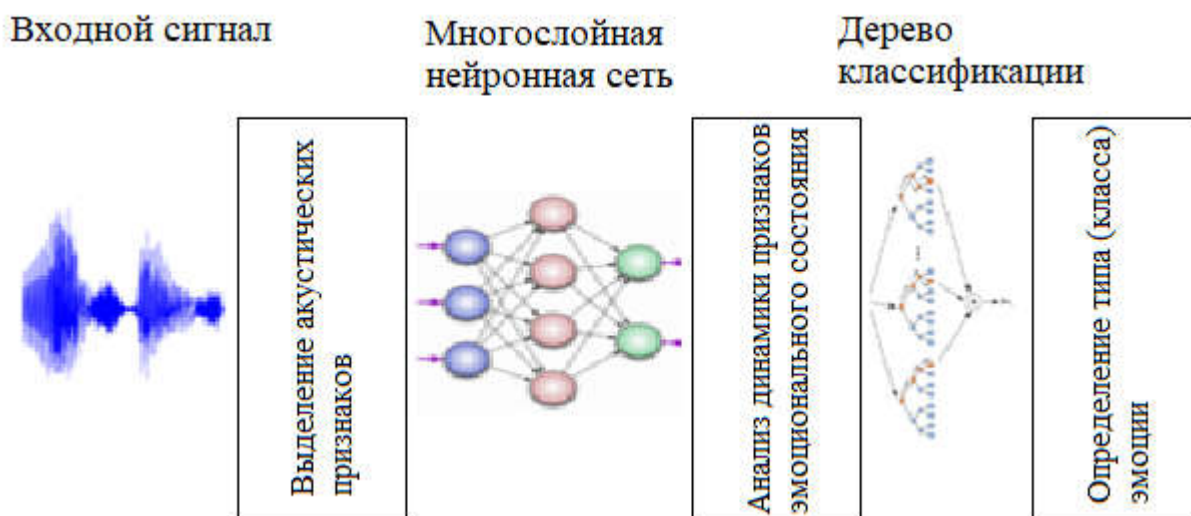


Рисунок 1 – Схема двухэтапного алгоритма

Данный алгоритм состоит включает в себя последовательность шагов, объединенных в два основных блока. На первом этапе мы разбиваем высказывания на частично перекрывающиеся фреймы, для каждого из которых вычисляем акустические признаки. Затем с помощью обученной нейронной сети получаем вероятность каждого фрейма быть в одном из заданных эмоциональных состояний. Для одного интервала физически это означает проецирование вектора признаков на некоторое пространство, имеющее размерность, совпадающую с числом возможных эмоциональных состояний. Получили временной ряд из вероятностей время дискретизации, которого равно длительности фрейма. Классификатор на втором этапе использует эту динамику для принятия решения об эмоции всего высказывания. Для этого временной ряд разбивается на несколько частей, и в каждой из них считаются новые признаки. В данной работе предлагается делить ряд на три части, и считать среднее, медианное и стандартное отклонения вероятностей всех фреймов в участке. Такой подход позволит частично учесть факт того, что эмоциональное состояние фрагмента зависит от его положения в высказывании.

Выбор классификатора на втором этапе обучения обуславливается его способностями в обучении. Размер обучающей выборки будет совпадать с количеством высказываний, что существенно меньше, чем число фреймов. Поэтому использование нейронных сетей в данном случае не оправдано. Одним из подходящих методов является случайный лес (Random Forest Classifier).

ЗАКЛЮЧЕНИЕ

В процессе диссертации был разработан и исследован алгоритм распознавания эмоционального состояния человека по речевому сигналу, основанный на нейронной сети, 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 метод на основе нейронной сети.

В результате разработки и исследования алгоритма, основанного на нейронной сети, были выявлены его недостатки и предложены решения, данных недостатков.

Разработанный алгоритм позволяет распознавать эмоции с более высокой точностью, относительно других подобных методов, примерно на 5-8%. Точность распознавания уменьшается с ростом количества распознаваемых эмоциональных состояний. Самой важной проблемой в распознавании эмоций по речевому сигналу, является извлечение дискриминационных признаков, которые эффективно характеризуют различное эмоциональное состояние.

Для более точной оценки эмоционального состояния человека, наряду с анализом лицевой мимики следует также анализировать речь, интонацию и движения. Полноценный анализ состояния человека невозможен без комплексной системы определения эмоций. Создание подобной системы является целью дальнейшей работы. Разработка такой технологии позволит качественно изменить форму коммуникации между человеком и машиной.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Дубровский, П. С. Обзор методов распознавания эмоционального состояния субъекта / П. С. Дубровский, В.Э. Севрюк, В.А. Захарьев// Материалы международной научной конференции, ITS-2017, Минск, 25 октября 2017 г. / – с. 68-69