

Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ
Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра дискретной математики
и информационных технологий

**Идентификация человека по голосу с использованием
нейронных сетей**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 5 курса 521 группы
направления 09.03.01 «Информатика и вычислительная техника»
факультета компьютерных наук и информационных технологий
Зиновой Анастасии Игоревны

Научный руководитель

к. ф.-м.н., доцент

подпись, дата

И. Д. Сагаева

Зав. кафедрой

к. ф.-м.н., доцент

подпись, дата

Л.Б. Тяпаев

Саратов 2017

Введение

Попытки научить компьютеры общению с людьми, используя естественный голосовой интерфейс, предпринимались с начала зарождения компьютерной техники. В самом деле, ведь для создания комплекса идентификации пользователя или речи нужно решить множество неоднозначных задач.

Прежде всего, необходимо реформировать колебания воздуха в электрические сигналы, воспользовавшись микрофоном, отфильтровав, в то же время, все помехи и шумы по возможности. Затем, сигнал следует преобразовать в цифровую форму, которая уже будет доступна для обработки при помощи компьютера (оцифровать). Стоит упомянуть, что получение из речевого потока лингвистических структур является собой не легкую задачу. В настоящем мире речь является собой непрерывный поток звуков. В процессе формирования непрерывной речи звуки, соответствующие одним и тем же буквам, могут преобразовываться при соединении с другими звуками.

Звуки непрерывной речи содержат всегда изменяющийся спектр гармонических частот, а также шумы. Для того чтобы выделить из оцифрованного звука лингвистические составляющие, используются всевозможные математические приемы.

Цели моей дипломной работы:

- 1) Изучить способы и методы идентификации голоса;
- 2) Изучить теорию о нейронных сетях, которые будут использованы в комплексе распознавания голоса;
- 3) Разработать комплекс по идентификации голоса.

Данная тема является актуальной как со стороны общения человека с компьютером, так и со стороны биометрических схем защиты.

Основное содержание

Обеспечение безопасности систем с помощью биометрических параметров

Во все времена была проблема осуществлять санкционированный допуск пользователя к несанкционированному объекту или засекреченной информации, распознавать. Одним из самых популярных биометрических параметров, используемых для аутентификации, является человеческий голос.

Методы распознавания речи

Все системы распознавания речи можно разделить на два класса:

- системы, зависящие от диктора;
- системы, не зависящие от диктора

К первому способу относятся системы, работа которых не зависит от обладателя голоса. Такие комплексы не нуждаются в предварительном обучении. Они могут распознавать речь любого пользователя.

Комплексы второго класса настраиваются на речь пользователя в самом процессе обучения. Для работы с другим пользователем такие комплексы требуют полной перенастройки.

Параметры голоса и звука

Главное в построении устройств аутентификации по голосу – выбор параметров, лучше всего описывающих индивидуальность голоса. Эти параметры сигнала называются признаками индивидуальности.

Извлечение признаков

Для аутентификации по такому биометрическому параметру, как голос, зачастую требуется извлечение параметров из выходных сигналов — процесс выделения значимых признаков. Самыми распространенными и успешными методами считаются мел-частотные кепстральные коэффициенты и линейные предсказательные кепстральные коэффициенты.

Обзор методов и подходов идентификации пользователей

Несмотря на то, что способы распознавания человеческого голоса во многом отличаются, вполне возможно обособить следующие основные стадии, которые применяются практически в любом методе:

1. Получение признаков из входного речевого сигнала.
2. Построение модели (шаблона) диктора на основе извлеченных на предыдущем этапе стека признаков.

Нейронные сети. Основные понятия

Нейронная сеть — это последовательность нейронов, объединенных друг с другом синапсами. Нейрон можно охарактеризовать состоянием в текущий момент времени и имеющейся группой **синапсов** — входных связей одного направления, объединенных с выходами остальных нейронов следующих слоев.

Каждый нейрон (за исключением выходного слоя) имеет **аксон**, представляющий из себя выходную связь данного нейрона, с которой сигнал (остановки или же активации) подается на синапсы последующих элементов.

Выбор параметров многослойной сети. Гиперпараметры

Гиперпараметры — это значения, которые нужно подбирать вручную и зачастую методом проб и ошибок. Среди таких значений можно выделить:

- 1) Количество скрытых слоев;
- 2) Количество нейронов в каждом слое.

Идеализация нейронных сетей

Многие вещи, которые людям кажутся трудновыполнимыми, можно с легкостью совершить, обращаясь к компьютерным технологиям. Однако, техника не в состоянии думать так же, как живой биологический мозг.

Система идентификации голоса с помощью нейронных сетей

Технические средства

Первоначальным этапом голосовой идентификации является получение голоса диктора. Для этого необходим микрофон, фильтр и аналого-цифровой преобразователь для дальнейшей работы с цифровой записью голоса.

Принцип получения голосового отпечатка.

Для осуществления процесса преобразования аудиозаписи необходимо совершить некоторую последовательность действий.

Нейросетевое сравнение.

Нейросетевое сравнение — последний и самый важный этап в системах идентификации человеческого голоса. Подавая различные ранее сформированные вектора характеристик, происходит обучение сетей. Сам процесс обучения в основном зависит от вида выбранной сети. В каких то случаях требуется подбирать веса в ручную, в других может понадобится составление базы векторов характеристик, в качестве примера правильных и неправильных ответов для сети (обучение с учителем).

Практическая часть

В данном разделе подробно описана практическая часть, выполненная студенткой Зионовой Анастасией.

Заключение

В ходе выполнения дипломной работы была подробно изучена теория на тему распознавания человеческой речи с помощью нейронных сетей. Описаны методы распознавания речи, параметры, которые могут использоваться в выше озвученных целях, признаки, на основе которых нейронные сети могут делать какие либо умозаключения об обладателе голоса. Так же были изучены основы нейронных сетей, их понятия и классификации.

Опираясь на изученный материал, была разработана практическая часть в виде комплекса по идентификации человеческого голоса. Для нее была самостоятельно собрана база голосов, а так же разработан код на двух языках, реализовывающий задуманную идею. Создана нейронная сеть.