

АЛГОРИТМЫ АНАЛИЗА РЕЧЕВОГО СИГНАЛА ДЛЯ ОПРЕДЕЛЕНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ СУБЪЕКТА

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

I. ВВЕДЕНИЕ

Распознавание эмоционального состояния субъекта является актуальной темой в данный момент и может быть использована во многих отраслях, таких как медицина, психология, обеспечение безопасности и т.д.

Одним из способов распознавания речи является использование нейронных сетей. Нейронные сети – это одно из направлений исследований в области искусственного интеллекта, основанное на попытках воспроизвести нервную систему человека (способность нервной системы обучаться и исправлять ошибки).

II. ОПИСАНИЕ РАБОТЫ СО ЗВУКОМ

Чтобы сохранить звуковой сигнал на цифровом носителе, его необходимо разбить на множество промежутков – фреймов - и взять отсчеты (мгновенные значения сигнала) через равные промежутки времени. Информация об уровне сигнала между взятыми отсчетами может теряться, поэтому, воспользуясь теоремой Котельникова, сигнал может быть полностью восстановлен. Согласно теореме Котельникова, для воссоздания исходной звуковой волны, почти идеально приближенной к исходной, достаточно использовать частоту дискретизации, вдвое превышающую самую высокую частоту записываемого звука. Звук представляет собой волну, распространяющуюся в пространстве. Рассмотрим пример графика звуковой волны слова «Hello» на небольшом отрезке времени (см.рис.1)



Рис. 1 – Звуковая волна

III. ОПИСАНИЕ НЕЙРОННОЙ СЕТИ

Для преобразования звуковой волны в цифровой формат, необходимо использовать значе-

ния амплитуды волны в равностоящих точках – звуковой сигнал необходимо дискретизировать (см.рис.2), затем использовать дискретное преобразование Фурье для подачи на нейронную сеть. Также к полученным значениям рекомендуется применить оконную функцию Хэмминга, что бы “сгладить” значения на границах фреймов и уменьшить утечку фреймов.

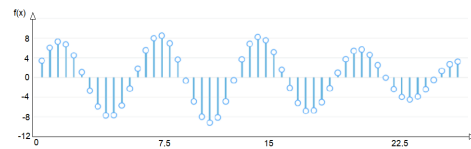


Рис. 2 – Дискретный сигнал звуковой волны

На нейронную сеть подается сигнал, на выходе которого получаем значение буквы. Для распознавания используется рекуррентная нейронная сеть, так как каждая распознанная буква влияет на следующие, например, если у нас есть «Hel’», то следующие буквы будут, с большой вероятностью «lo’».

IV. ВЫВОДЫ

Таким образом, запоминая предыдущие результаты, рекуррентная нейронная сеть сможет давать некоторые прогнозы о будущих выходных сигналах. Обработывая полученный сигнал фразы с самого начала, сеть не только распознает каждый звук или фонему, но и может с определенной точностью предсказывать конец фразы.

Список литературы

1. Портал искусственного интеллекта AIportal [Электронный ресурс] / Нейронные сети - Москва, 2016. – Режим доступа: <http://www.aiportal.ru>. – Дата доступа: 15.03.2019.
2. Neurohive [Электронный ресурс] / Нейронные сети - Санкт-Петербург, 2015. –Режим доступа: <https://neurohive.io/ru/>. – Дата доступа: 15.03.2019.

Цвирко Александра Игоревна, магистрантка 1 курса ФИТуУ БГУИР, tsvirko8@gmail.com.

Научный руководитель: Свито Игорь Леонтьевич, доцент кафедры ТОЭ БГУИР, к.т.н., доцент.