

ИССЛЕДОВАНИЕ РЕЧЕВЫХ РАССТРОЙСТВ ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ

Меньшаков П. А., Мурашко И. А.

Кафедра «Информационные технологии», Гомельский государственный технический университет имени П.О. Сухого

Гомель, Республика Беларусь

E-mail: pmenshakov@gmail.com, iamurashko@tut.by

В настоящее время, голосовая идентификация и анализ голоса являются одним из передовых направлений, как в области систем защиты информации, так и в определении голосовых параметров [1]. Предлагается рассмотреть способ анализа состояния голоса и выявления его проблем, таких как усталость голосовых связок, повреждение или воспаления речевого тракта при помощи ранее записанного голоса диктора и нейросетевого анализа голосовых изменений.

ВВЕДЕНИЕ

Оценка нарушения голоса включает использование оценки множественных характеристик речи (степень дыхания, напряжение, шероховатость). Одна из проблем, связанных с использованием многомерных данных – их сравнение. Для выполнения сравнения и классификации предлагается использовать самоорганизующуюся карту Кохонена. Ввиду возможности обучения без учителя, она не нуждается в целевом векторе для выходов и, следовательно, не требует сравнения с predetermined идеальными ответами, а обучающее множество состоит лишь из входных векторов. Процесс обучения, следовательно, выделяет статистические свойства обучающего множества и группирует сходные векторы в классы. Предъявление на вход вектора из данного класса даст определенный выходной вектор.у

I. АППАРАТНАЯ РЕАЛИЗАЦИЯ

Для реализации устройства необходим микрофон, фильтр и аналого-цифровой преобразователь, для дальнейшей работы с цифровой записью голоса. Схема устройства представлена на рисунке 1.

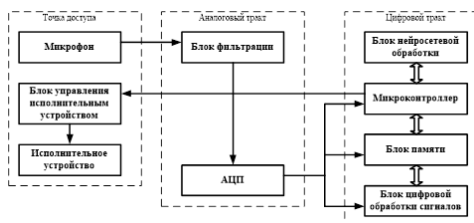


Рис. 1 – Схема устройства

С выхода микрофона сигнал подается на вход блока фильтрации. Следующим этапом является прохождение АЦП [2]. Далее оцифрованный сигнал попадает в блок цифровой обработки. В блоке цифровой обработки сигнал фильтруется и преобразуется в вектор, с которым в дальнейшем будет работать микропроцессор и нейросетевой обработчик. Для последующего

сравнения с сохраненным ранее вектором мелкепстральных коэффициентов, полученный вектор заносится в энергонезависимую память. После сравнения вектора в памяти с полученным вектором, микроконтроллер подает команду на блок управления внешним устройством, к примеру, на магнитный дверной замок. Сам процесс голосовой идентификации не требователен к ресурсам, и состоит из двух этапов. Первым этапом является получение речевых признаков диктора и преобразование к виду, в котором его можно будет сравнить с другими. Вторым шагом является их сравнение при помощи обученной нейронной сети [3].

II. Дисфония

Дисфония четко определяется как основное расстройство фонации, является следствием заболеваний или патологии голосовых связок. Так как данное отклонение работы речевого тракта сопровождается слышимым изменением голоса, то может быть записано и обработано в отношении к образцу голоса диктора, до появления отклонений. Есть два вида дисфонии – функциональная и повреждающая. Повреждающие дисфонии делится на два вида: врожденная и приобретенная дисфония. Повреждающая дисфония в некоторых случаях может появиться после функциональной дисфонии. Возможны функциональные дисфонии:

- хрипота;
- ларингит;
- воспаление гортани;
- гиперкинез;
- слизистая киста или киста оболочки;
- гипокинез;
- инфекция горла;
- глоссоплегия;
- ринофарингит.

Указанные заболевания имеют различные степени тяжести, что существенно отражается на качестве жизни человека. Таким образом, диагностирование нарушений фонации помогут опреде-

литель уровень качества жизни. Шкала голосовых нарушений приведена в таблице 1.

Таблица 1 – Шкала голосовых нарушений

Градация нарушений	Степень Нарушений	Описание	Рекомендации
Нет нарушений	0	-	-
Незначительные нарушения	1	Лингвистическое нарушение едва чувствуется или ощущается только самим пациентом	Рекомендуется речевая терапия
Умеренные нарушения	2	Снижение легкости и скорости говорения	Необходима речевая терапия
Тяжелые нарушения	3	Говорящий нуждается в помощи слушателя. Пациент часто не может быть понятым, но понимает сам	Необходима речевая терапия и помощь со стороны слушателя
Глубокие повреждения	4	Говорение фрагментарными выражениями. Слушатель должен о многом догадываться. Информации мало, и слушатель	Необходима речевая терапия и изучение языка жестов, консультации или синтез голоса.

III. АНАЛИЗ НАРУШЕНИЯ РЕЧИ

В исследовании, проведенном Leinonen Et Al., для замены непосредственного прослушивания голоса, была создана шкала оценок различных степеней и форм дисфонии. Для сравнения критериев использовалась нейронная сеть без учителя, обучение которой было проведено с использованием перцепционной карты оценок

нормального и дисфонического голоса [3]. Результаты эксперимента представлены на рисунке 2.

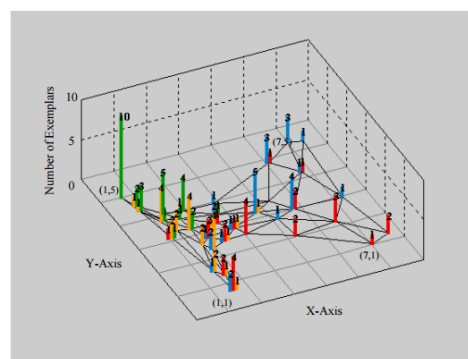


Рис. 2 – Результаты определения дисфонии голоса (зеленый – нормальный голос, желтый – гипотонусная, красный - гипертонусная, синий – спазматическая)

У данного подхода имеется несколько недостатков:

- Отсутствие сравнительной характеристики с предыдущим состоянием голоса. Ввиду данного недостатка невозможно отделить врожденные нарушения от приобретенных.
- Отсутствие возможности диагностирования нескольких расстройств одновременно.

Классификация различных форм и степень дисфонии может быть произведена путем использования не только перцепционных оценок патологии, шероховатости, дыхания, напряжения и астении но и сверкой данных оценок с предыдущим значением, путем включения голосового отпечатка с симуляцией той или иной степени заболевания в качестве входных параметров. Таким образом голосовой отпечаток может быть использован не только для средств контроля доступа, но и для оценки голосовых отклонений диктора. Очевидное преимущество данного подхода заключается в отсутствии непосредственного контакта с диктором и возможность удаленной диагностики, включая использование электронных средств связи. Так же, данная процедура имеет возможность полной автоматизации и нетребовательна к ресурсам.

1. Bosi, M. Introduction to digital audio coding and standards / M. Bosi, R.E. Goldberg – Springer Science+Business Media, USA. – 2003. – 434 p.
2. You, Y. AudioCoding: Theory and Applications / Y. You – NY: Springer, 2010 – 349 p.
3. Загуменнов, А. П. Компьютерная обработка звука./ А. П. Загуменнов - М.: ДМК, 1999. - 384 с
4. Using self-organizing maps to cluster music files based on lyrics and audio features / Research Congress 2013 De La Salle University, Manila, March 7-9, 2013.
5. Kohonen, T. (2001). Self-Organizing Maps. Berlin: Springer.