

## СЕКЦИЯ VIII ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И МОДЕЛИРОВАНИЕ

---

### ИССЛЕДОВАНИЕ РЕЧЕВЫХ РАССТРОЙСТВ ПРИ ПОМОЩИ ГОЛОСОВОЙ ИДЕНТИФИКАЦИИ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

**П. А. Меньшаков**

*Учреждение образования «Гомельский государственный технический университет имени П. О. Сухого», Республика Беларусь*

Научный руководитель И. А. Мурашко

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.

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 выделяет статистические свойства обучающего множества и группирует сходные векторы в классы. Предъявление на вход вектора из данного класса даст определенный выходной вектор.

**Аппаратная реализация.** Первоначальным этапом является получение эталонного голоса диктора. Для этого необходим микрофон, фильтр и аналого-цифровой преобразователь для дальнейшей работы с цифровой записью голоса [2].

С выхода микрофона сигнал подается на вход блока фильтрации. Следующим этапом является прохождение АЦП.

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.

Также полученный вектор заносится в энергонезависимую память. Это необходимо для последующего сравнения с полученным отпечатком.

Далее при работе с устройством диктор вводит текущую голосовую запись и при помощи микропроцессора и нейросетевого обработчика производится анализ отклонений голоса и вывод информации на экран. Общая схема устройства представлена на рис. 1.

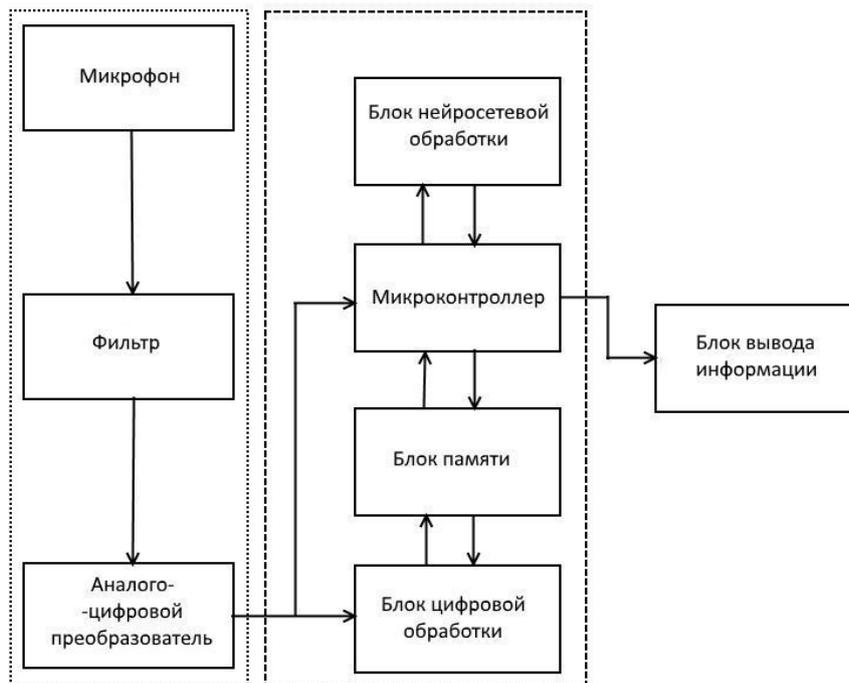


Рис. 1. Схема устройства

**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 рис. 2.

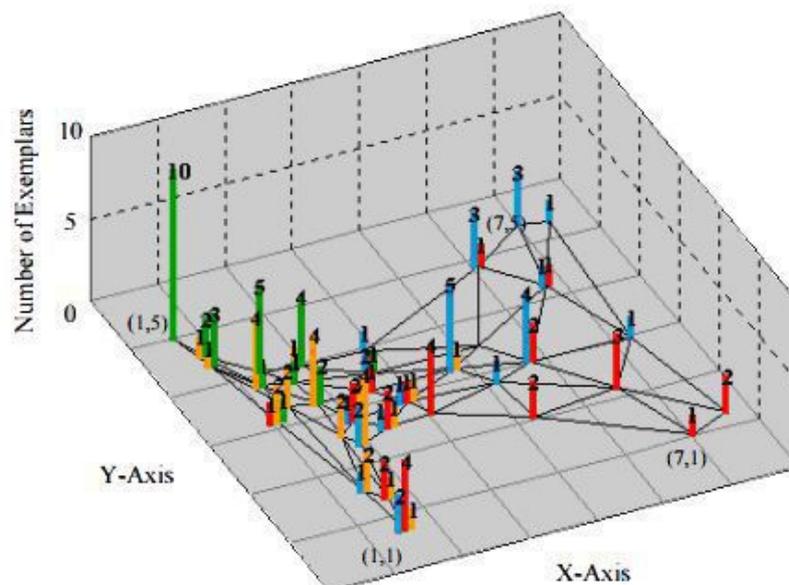


Рис. 2. Результаты определения дисфонии голоса (зеленый – нормальный голос, желтый – гипотонусная, красный – гипертонусная, синий – спазматическая)

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, необходимо произвести следующий ряд действий:

– получить голосовую запись диктора;

– коррекция голоса в соответствии с таблицей оценок [3], [4];

– получение голосовых отпечатков из скорректированных голосовых записей;

– обучение нейронной сети с использованием модифицированных голосовых отпечатков [5];

– ввод текущего голоса и нейросетевое сравнение с имеющимися отпечатками.

**Заключение.** 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 возможности удаленной диагностики, включая использование электронных средств связи. Также данная процедура имеет возможность полной автоматизации и нетребовательна к ресурсам.

#### Литература

1. Introduction to digital audio coding and standards / M. Bosi, R.E. Goldberg – Springer Science+Business, Media USA. – 2003. – 434 p.
2. AudioCoding: Theory and Applications / Y. You. – NY: Springer, 2010. – 349 p.
3. In Raymond D. Kent and Martin J. Ball Editors: Voice Quality Measurement, Singular: San Diego (2000). – P. 103–116.
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.