

Министерство образования Республики Беларусь
Учреждение образования
«Полоцкий государственный университет»

**ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ:
ДОСТИЖЕНИЯ, ПРОБЛЕМЫ, ИННОВАЦИИ
(ИКТ-2018)**

Электронный сборник статей

I Международной научно-практической конференции,
посвященной 50-летию Полоцкого государственного университета

(Новополоцк, 14–15 июня 2018 г.)

Новополоцк
Полоцкий государственный университет
2018

Информационно-коммуникационные технологии: достижения, проблемы, инновации (ИКТ-2018) [Электронный ресурс] : электронный сборник статей I международной научно-практической конференции, посвященной 50-летию Полоцкого государственного университета, Новополоцк, 14–15 июня 2018 г. / Полоцкий государственный университет. – Новополоцк, 2018. – 1 электрон. опт. диск (CD-ROM).

Представлены результаты новейших научных исследований, в области информационно-коммуникационных и интернет-технологий, а именно: методы и технологии математического и имитационного моделирования систем; автоматизация и управление производственными процессами; программная инженерия; тестирование и верификация программ; обработка сигналов, изображений и видео; защита информации и технологии информационной безопасности; электронный маркетинг; проблемы и инновационные технологии подготовки специалистов в данной области.

Сборник включен в Государственный регистр информационного ресурса. Регистрационное свидетельство № 3201815009 от 28.03.2018.

Компьютерный дизайн М. Э. Дистанова.

Технические редакторы: Т. А. Дарьянова, О. П. Михайлова.

Компьютерная верстка Д. М. Севастьяновой.

211440, ул. Блохина, 29, г. Новополоцк, Беларусь
тел. 8 (0214) 53-21-23, e-mail: irina.psu@gmail.com

ДИАГНОСТИКА НЕВРОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ НА ОСНОВЕ АНАЛИЗА РЕЧЕВОГО СИГНАЛА

*студент А.Д. ГВОЗДОВИЧ, канд. техн. наук М.И. ВАШКЕВИЧ
(Белорусский государственный университет
информатики и радиоэлектроники, Минск)*

Введение. Неврологические заболевания – одна из самых распространенных категорий болезней в мире. Эффективность лечения любого неврологического заболевания значительно повышается является при условии его ранней диагностики. В работе рассматривается вопрос построения системы диагностики и выявления неврологических заболеваний на основе анализа речевого сигнала.

В неврологии существует несколько способов получения информации о болезни:

- неврологические тесты на наличие рефлексов;
- аппаратные исследования (МРТ, ЭЭГ, УЗИ, ЭКГ);
- лабораторные исследования: крови, мазков, цитология тканей.

Применение указанных способов требует значительных затрат времени и материальных средств. Не во всех случаях есть возможность оперативно посетить специалиста и пройти необходимые обследования. Для таких людей полезно было бы иметь средство, с помощью которого можно выявить заболевание в домашних условиях.

Проблемы, связанные с речью и глотанием, являются первичными симптомами у приблизительно 30% пациентов с диагнозом боковой амиотрофической склероз [6], наличие проблем с речью также имеет место при болезнях Паркинсона и Альцгеймера. Поэтому одним из возможных путей к диагностике и выявлению неврологических заболеваний является построение системы на основе анализа речевого сигнала.

Неврологические заболевания влияют на работу органов, отвечающих за формирование речи. В зависимости от недуга, его степени и возраста пациента возникают различные речевые патологии. Во всех неврологических заболеваниях есть такая патология, как невозможность держать одинаковый темп речи. Характерны также такие признаки, как слабость в голосе, хриплость, дисфония, дизартрия и т.д. Данные патологии можно выявить путем анализа речевого сигнала. Чтобы сделать вывод о наличии неврологического заболевания требуется комплексный анализ многих параметров речевого сигнала.

Большинство подходов к автоматической диагностики по речевому сигналу основаны на использовании большого набора признаков, извлеченных из речевого сигнала (дискретное преобразование Фурье, автокорреляция, джиттер, шиммер, отношение "гармоники/шум" и т.д.). На основе этих данных обучается нейронная сеть, которая учится классифицирует речевой сигнал как нормальный или патологический. Однако такой подход имеет два недостатка. Во-первых, существует риск, что система будет слишком сильно оптимизирована под обучающую выборку, что ограничит ее способность к обобщению. Во-вторых, данный подход требует больших вычислительных ресурсов, что не позволяет его использовать в портативных устройствах [2]. Цель данной работы состоит в том, чтобы сделать достаточно точную систему, которая не будет требовать больших вычислительных ресурсов.

Система диагностики. Неврологические заболевания влияют на тембр голоса, разборчивость и темп речи, то для их выявления целесообразно анализировать спектральные огибающие гласных звуков и паузы между словами. Для анализа в системе (рис. 1) брались речевые сигналы с записью счета от 1 до 10.

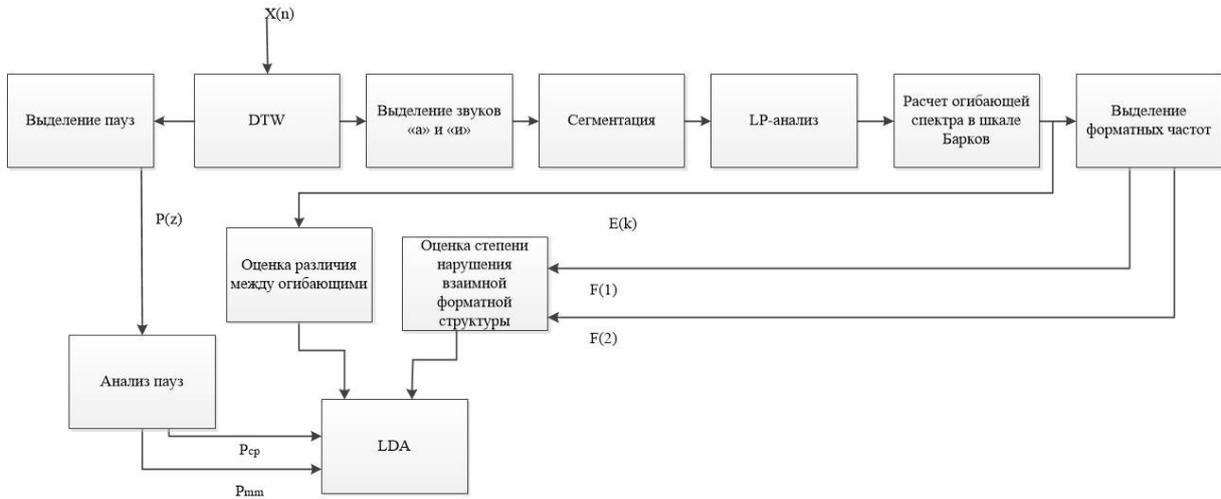


Рисунок 1. – Схема анализа речевого сигнала

В блоке DTW (англ. dynamic time warping – динамическая трансформация временной шкалы) речевой сигнал сравнивается с размеченным эталоном [3]. Далее уже из размеченного речевого сигнала выделяются паузы $P(z)$ и звуки «а» и «и». В блоке сегментации речевой сигнал разбивается на кадры длительностью 27 мс с перекрытием в $1/8$ от длительности кадра. Каждый кадр сигнала анализируется методом линейного предсказания (англ. LP – Linear Prediction), а коэффициенты линейного предсказания используются для получения огибающей спектра $E(f)$, которая затем переводится в психоакустическую частотную шкалу барков $E(k)$. В последнем блоке выделяются формантные частоты $F(1)$ и $F(2)$ [4].

Известно, что форматная структура звуков /а/ и /и/, произнесенных здоровым человеком имеет следующий вид (рис. 2).

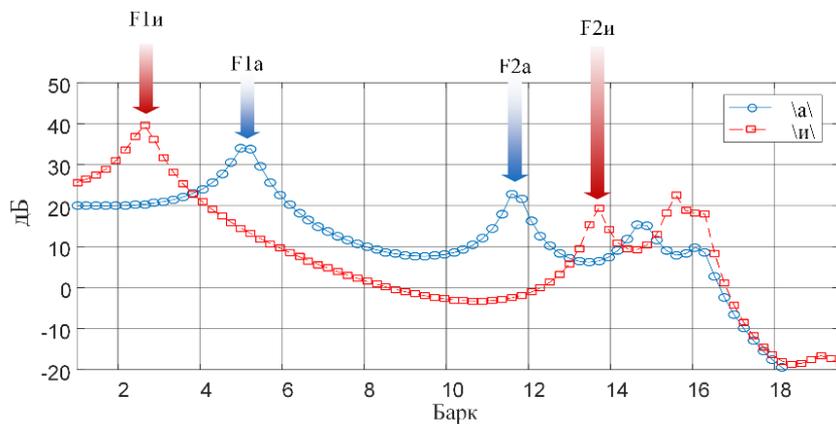


Рисунок 2. – Огибающие звуков /а/ и /и/ у здорового человека

Форманты имеют высокую добротность, значительно разнесены на частотной оси и расположены в следующей последовательности $F_i(1) < F_a(1) < F_a(2) < F_i(2)$. Такой порядок у людей, имеющих неврологические заболевания может быть нарушен. Для количественной оценки степени нарушения взаимной формантной структуры звуков /а/ и /и/ в [4] предложено использовать следующую меру:

$$fnt_{err}(F_i, F_a) = \begin{cases} 2, & \text{если } F_i(1) > F_a(1) \text{ или } F_a(2) > F_i(2) \\ 2 - \frac{F_a(1) - F_i(1)}{2} - \frac{F_a(2) - F_i(2)}{2}, & \text{если } F_i(1) - F_a(1) < 2 \text{ и } F_i(2) - F_a(2) < 2 \\ 1 - \frac{F_a(1) - F_i(1)}{2}, & \text{если } F_a(1) - F_i(1) < 2 \\ 1 - \frac{F_i(2) - F_a(2)}{2}, & \text{если } F_i(2) - F_a(2) < 2 \\ 0, & \text{иначе.} \end{cases}$$

Помимо изменений во взаимной формантной структуре звуков «а» и «и» у пациентов наблюдалось увеличение сходства между формами огибающих этих звуков. Для количественной оценки различия между огибающими звуков «а» (E_a) и «и» (E_i) предлагается использовать l_1 -норму расстояния:

$$d_1(E_i, E_a) = \sum_{k=1}^K |E_i(k) - E_a(k)|,$$

где K – число точек в частотной области барков.

На рис. 3 представлены распределения плотности вероятности для признаков fnt_{err} и d_1 .

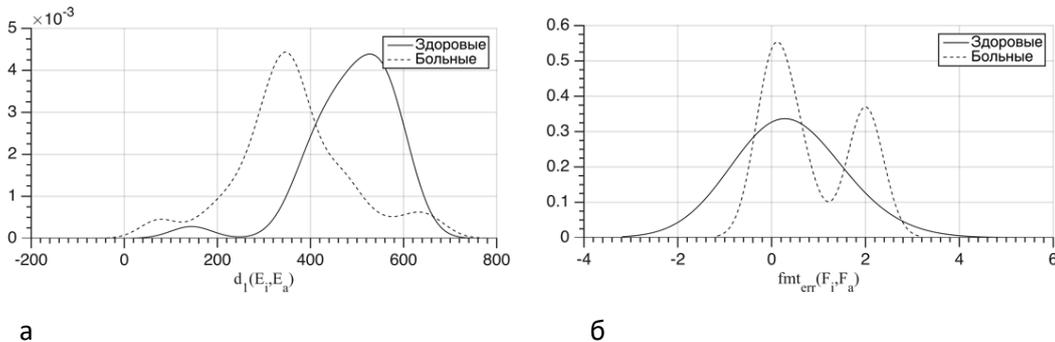


Рисунок 3. – Распределение плотности вероятности: а – $d_1(E_i, E_a)$, б – $fnt_{err}(F_i, F_a)$

Для оценки пауз предлагается анализировать среднюю паузу и сумму между максимальной (P_{max}) и минимальной паузой (P_{min}):

$$P_{cp} = \frac{\sum_{z=1}^N P(z)}{N}, \quad P_{mm} = \frac{P_{max} + P_{min}}{P_{cp}}.$$

Здоровый человек может выдержать паузу при медленном счете, поэтому P_{min} у здорового будет стремиться к 2, а у больного оно не обязательно будет близко к 2.

На рис. 4 представлены распределения плотности вероятности для признаков P_{cp} и P_{mn} .

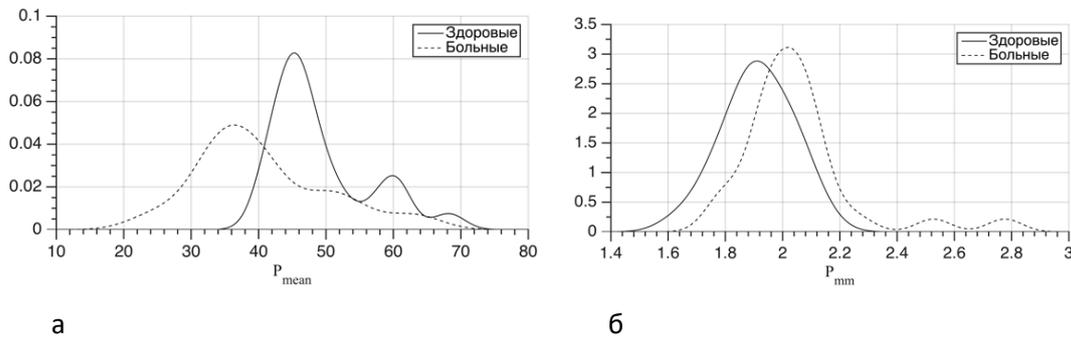


Рисунок 4. – Распределение плотности вероятности: а – P_{cp} , б – P_{mn}

Найденные значения $[d_1(E_i, E_a) \quad f_{mt_err}(F_i, F_a) \quad P_{cp} \quad P_{mn}]$ отправляются в блок LDA (linear discriminant analysis - линейный дискриминантный анализ). В данном блоке происходит разделение между людьми с неврологическим заболеванием и без него. На рис. 5 показано распределение плотности вероятности результатов функции LDA [5].

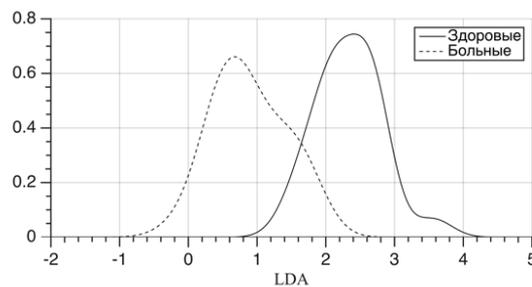


Рисунок 5. – Распределение плотности вероятности LDA

Заключение. В работе рассмотрена задача выявления и диагностики неврологических заболеваний на основе анализа речевого сигнала. Для решения данной задачи выполнять совместный анализ огибающих различных гласных звуков и анализ пауз. Для классификаций предложено использовать метод линейного дискриминантного анализа. Анализ результатов данного алгоритма доказывает возможность реализации задачи выявления и диагностики неврологических заболеваний на основе анализа речевого сигнала.

Литература

1. M. Little Nonlinear, biophysically-informed speech pathology detection / M. Little, P. McSharry, I. Moroz, S. Roberts // IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, May 16 2006, pp. 1080-1083.
2. Thomas Prätzlich, Jonathan Driedger, and Meinard Müller (2016). Memory-Restricted Multiscale Dynamic Time Warping. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 569—573.

3. G. Al-Naymat SparseDTW: A Novel Approach to Speed up Dynamic Time Warping / G. Al-Naymat, S. Chawla, J. Taheri// The 2009 Australasian Data Mining, Melbourne, Australia , 12/2009, С. 117-127
4. Вашкевич, М.И. Выделение признаков неврологических заболеваний из речевого сигнала / М.И. Вашкевич [и др.] // Цифровая обработка сигналов и ее применение : тр. 20-й междунар. конф., Москва, 28–30 марта 2018 г. – М., 2018. – Т. 1. – С. 179–184.
5. Yu, H. A direct LDA algorithm for high-dimensional data — with application to face recognition / H. Yu, J. Yang // Pattern Recognition. – 2001. – С. 2067–2069.
6. Fractal features for automatic detection of dysarthria / T. Spangler [et al.] // Proc. of IEEE EMBS International conference on biomedical & health informatics, Florida, USA, Feb. 16–19, 2017. – P. 437–440.