

# СИСТЕМА ДИАГНОСТИКИ И ВЫЯВЛЕНИЯ НЕВРОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ НА ОСНОВЕ АНАЛИЗА РЕЧЕВОГО СИГНАЛА

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

Гвоздович А.Д.

Вашкевич М.И. – к.т.н., доцент

Неврологические заболевания – одна из самых распространённых категорий болезней в мире. Условием эффективного лечения любого неврологического заболевания является его ранняя диагностика. 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:

- сбор анамнеза и регистрация объективных данных;
- неврологические тесты на наличие рефлексов;
- аппаратные исследования (МРТ, ЭЭГ, УЗИ, ЭКГ);
- лабораторные исследования: крови, мазков, люмбальной жидкости, цитология тканей.

Но далеко не у каждого есть возможность посетить специалиста для диагностики неврологического заболевания. В связи с этим было решено разработать алгоритм, который сможет диагностировать и выявить заболевание в домашних условиях.

В ходе изучения различных возможностей диагностики заболевания замечен самый доступный для обычного человека способ: диагностика и выявление неврологических заболеваний на основе анализа речевого сигнала.

Неврологические заболевания влияют на работу органов, отвечающих формированию за речи. В зависимости от недуга, его степени и возраста пациента возникают различные речевые патологии. Во всех неврологических заболеваниях есть такая патология, как невозможность выдержать одинаковую паузу между словами в речи. Также есть такие патологии, как слабость в голосе, хрипкость, дисфония и т.д. Данные патологии можно выявить путем исследования речевого сигнала и проанализировав найденные патологии можно сделать вывод о наличии или отсутствии неврологического заболевания.

Большинство подходов к автоматической диагностике 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 для учебных данных, что ограничит её 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.

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 брались речевые сигналы с записью счета от 1 до 10, которые обрабатывались по следующей схеме (рис. 1).

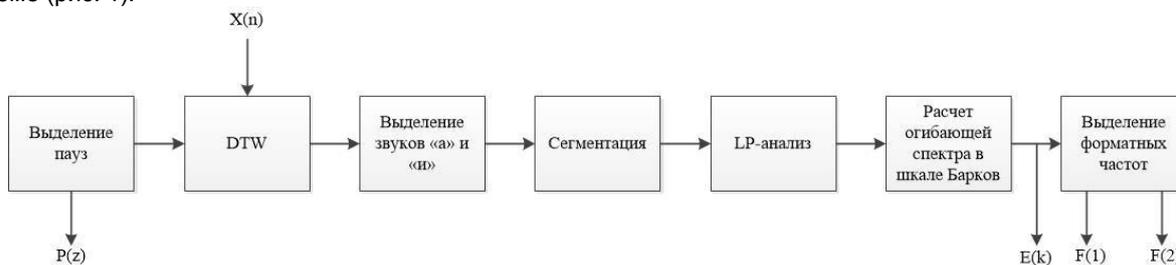


Рисунок 1– Схема анализа речевого сигнала

В блоке DTW (dynamic timewarping – 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].

Для количественной оценки степени нарушения взаимной формантной структуры звуков /а/ и /и/ предлагается использовать меру

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

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» и «и» предлагается использовать  $l_1$ -норму расстояния:

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

где  $E_i$  – огибающая звука «и»,  $E_a$  – огибающая звука «а»,  $K$  – число точек в частотной области барков, в которых определена огибающая.

Для оценки пауз предлагается анализировать среднюю паузу сумму между максимальной и минимальной паузой:

$$P_{cp} = \frac{\sum_{z=1}^9 P(z)}{9}, \quad P_{mm} = \frac{P_{max} + P_{min}}{P_{cp}},$$

где  $P_{max}$  – максимальная по длине пауза,  $P_{min}$  – минимальная пауза.

Так как здоровый человек может выдержать паузу при медленном счете, то значение  $P_{mm}$  у здорового будет стремиться к 2. В то время как у больного это значение необязательно будет близко к 2.

Ниже представлены распределения плотности вероятности для вычисляемых признаков.

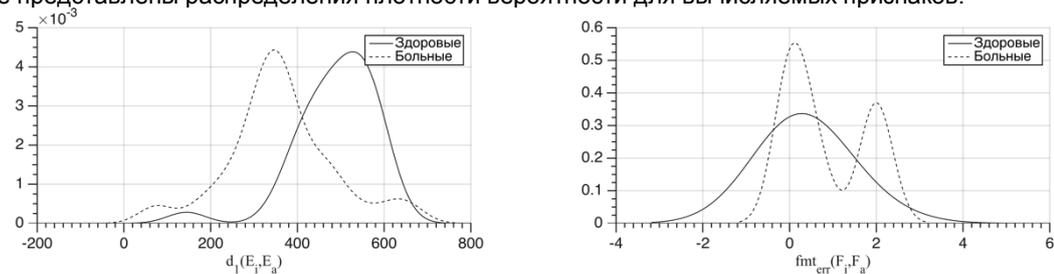


Рисунок 2 – Распределение плотности вероятности: а –  $d_1(E_i, E_a)$ , б –  $f_{mt_{err}}(F_i, F_a)$

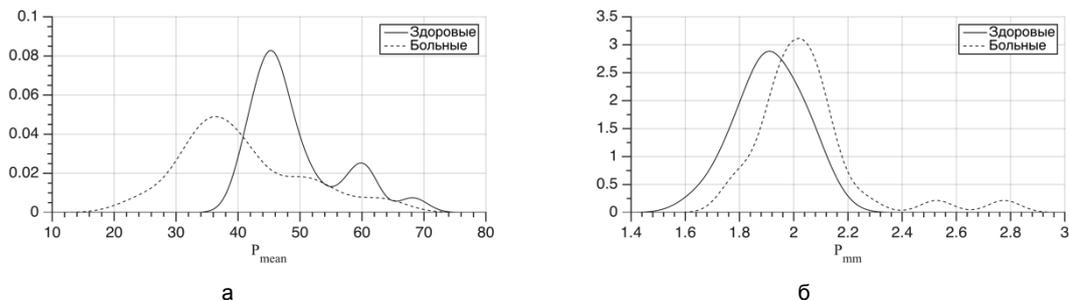


Рисунок 3 – Распределение плотности вероятности: а –  $P_{cp}$ , б –  $P_{mm}$

В работе рассмотрена задача выявления и диагностики неврологических заболеваний на основе анализа речевого сигнала. Предложено для решения данной задачи выполнять совместный анализ огибающих различных гласных звуков и анализ пауз. Анализ графиков показывает предлагаемые признаки позволяют выявлять патологические изменения в речи.

Список использованных источников:

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. Вашкевич М.И. Выделение признаков неврологических заболеваний из речевого сигнала / М.И. Вашкевич, Ю.Н. Рушкевич, И.С. Азаров, А.А. Петровский // Цифровая обработка сигналов и ее применение: труды 20-й междунар. конф., Россия, Москва, 28-30 марта, 2018 г. – Москва, 2018. – Т. 1. – С. 179–184.