

# ВЫДЕЛЕНИЕ ПРИЗНАКОВ НЕВРОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ ИЗ РЕЧЕВОГО СИГНАЛА

доц. Вашкевич М.И., доц. Рушкевич Ю.Н.\* , доц. Азаров И.С., проф. Петровский А.А.

Белорусский государственный университет информатики и радиоэлектроники  
ул. П.Бровки, 6, БГУИР, каф. ЭВС, 220013, Минск, Беларусь, e-mail: {vashkevich, azarov,  
palex}@bsuir.by

\*РНПЦ Неврологии и нейрохирургии Министерства здравоохранения Республики Беларусь,  
ул. Ф. Скорины, 24, Минск, 220114, Беларусь

**Аннотация.** В работе предложен способ оценки патологических изменений в речевом сигнале для диагностики неврологических заболеваний. Выявление отклонений в речи предлагается выполнять путем совместного анализа различных гласных звуков тестируемого сигнала. Приведенные результаты экспериментов показывают практическую применимость предлагаемого способа оценки патологии в голосе.

**Введение.** При прогрессирующих неврологических заболеваниях таких как боковой амиотрофический склероз (БАС) часто одним из первых симптомов болезни является изменение речи. Проблемы, связанные с речью и глотанием, являются первичными симптомами у приблизительно 30 % пациентов с диагнозом БАС [1]. В настоящее время определение патологических изменений в речи пациентов, страдающих БАС или другими неврологическими заболеваниями, основано на субъективной оценке врача, осуществляющего диагностику. В ряде случаев такая оценка является субъективной и зависит от опыта и навыков специалиста [2]. В свою очередь неверная оценка изменений в работе речевого аппарата приводит к значительной задержке при диагностике БАС.

Бульбарные нарушения, появляющиеся в результате БАС, представляют собой набор симптомов, включающий дисфонию (изменение звучности голоса), дизартрию (нарушение артикуляции), замедленность речи. Наличие указанных симптомов говорит о нарушении в работе мускулатуры языка, глотки, гортани и мягкого неба, обусловленных нарушением их иннервации [3]. Акустически дисфония проявляется тем, что голос больного становится глухим, вялым, появляется гнусавость. Вследствие прогрессирующей атрофии мышц речевого аппарата произнесение звуков становится смазанным, теряется разборчивость. Практически пропадают различия между разными гласными звуками, при произнесении согласных звуков чувствуется отсутствие слаженности в движении артикуляционных органов.

## Особенности анализа речевого сигнала.

Поскольку проявление признаков неврологических заболеваний влияют на тембр голоса и разборчивость речи, то для их выявления целесообразно анализировать спектральные огибающие гласных звуков. Ниже дается обоснование выбора конкретных звуков. Для анализа выбирались участки речевого сигнала длительностью 150–200 мс, содержащие гласные звуки, которые обрабатывались по схеме на рисунке 1.

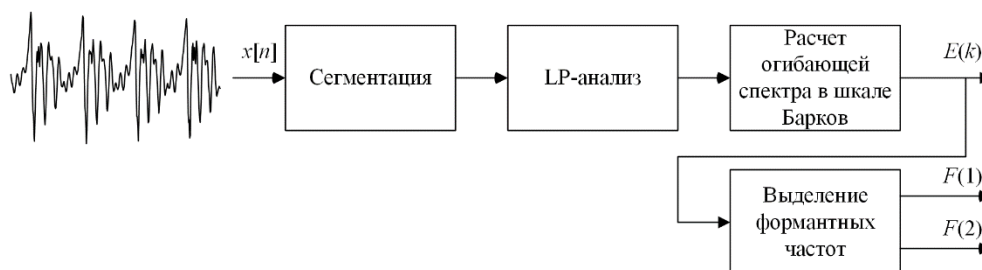


Рисунок 1– Схема анализа речевого сигнала

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*). Полученные в блоке «LP-анализ» коэффициенты линейного предсказания используются для получения огибающей спектра  $E(f)$ , которая затем

переводится в психоакустическую частотную шкалу барков  $E(k)$ . Выделение формантных частот при дисфонических расстройствах речи является нетривиальной задачей, поскольку форманта часто перестает быть точкой максимума на графике огибающей спектра (рисунок 2, а), а является лишь локальным отклонением формы огибающей от общего спадающего тренда. Поэтому для поиска формант предлагается анализировать огибающую после удаления тренда, рассчитанного по значениям огибающей в диапазоне [1; 12] Барк (рисунок 2, б).

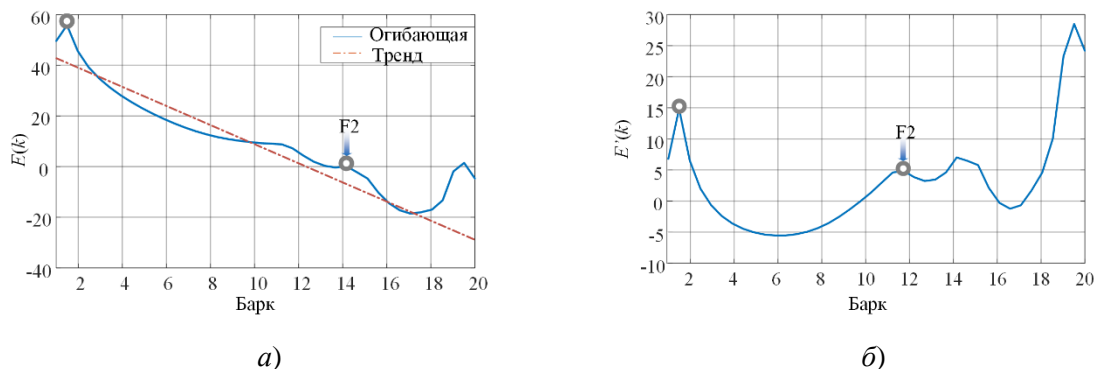


Рисунок 2– Выделение формантных частот:

- а) огибающая спектра гласного звука /а/, положение второй форманты определено с ошибкой;
- б) огибающая после удаления тренда, вторая форманта определена правильно

**Выбор гласных звуков для анализа.**

В результате БАС происходят изменения в стимуляции мускулатуры вообще и мышц языка в частности. С точки зрения звукообразования важным является положение языка по горизонтали (тело языка находится «вперед» или «сзади») и по вертикали («высокий» подъем или «низкий» подъем), см. рисунок 3 [4].

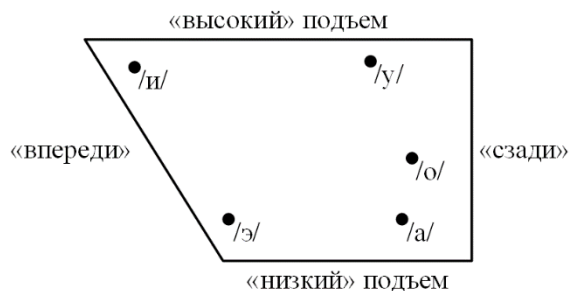


Рисунок 3– Относительное положение тела языка при произнесении гласных звуков /и/, /у/, /о/, /а/, /э/

Для выявления признаков неврологических заболеваний из речевого сигнала целесообразно выбрать звуки /а/ и /и/, поскольку для их формирования требуется значительная работа мышц языка. Для практического сравнения из тестового речевого сигнала с записью счета от 1 до 10 выбирались близкие по времени участки, содержащие звуки /а/ и /и/, после чего выполнялось сравнение взаимной формантной структуры этих звуков (как правило, звуки выбирались из слов «один», «два», «три»). Пример формантной структуры звуков /а/ и /и/, произнесенных здоровым человеком показан на рисунке 4.

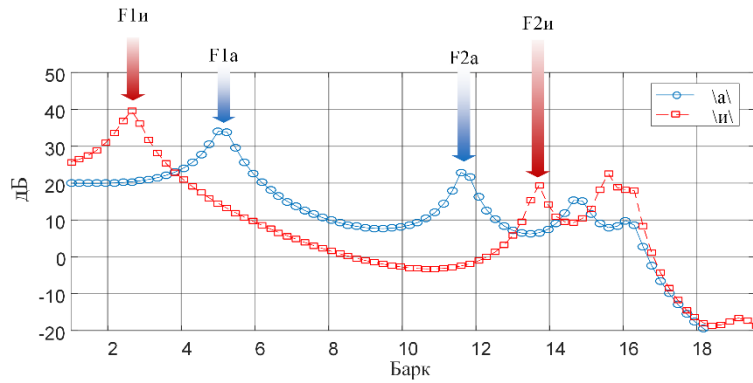


Рисунок 4– Взаимное расположение формантных частот звуков /а/ и /и/ у здорового человека

На основании визуального анализа огибающих звуков /а/ и /и/ на рисунке 4 можно сделать наблюдение, что форманты имеют высокую добротность, значительно разнесены на частотной оси и расположены в следующей последовательности  $F_i(1) < F_a(1) < F_a(2) < F_i(2)$ .

На рисунке 5 показан пример огибающих звуков /а/ и /и/, произнесенных пациентом с диагнозом БАС, имеющим проявления болезни в голосе. Визуальный анализ показывает изменения в порядке взаимного расположения формантных частот (вначале следует форманта  $F_a(1)$ , а затем  $F_i(1)$ ).

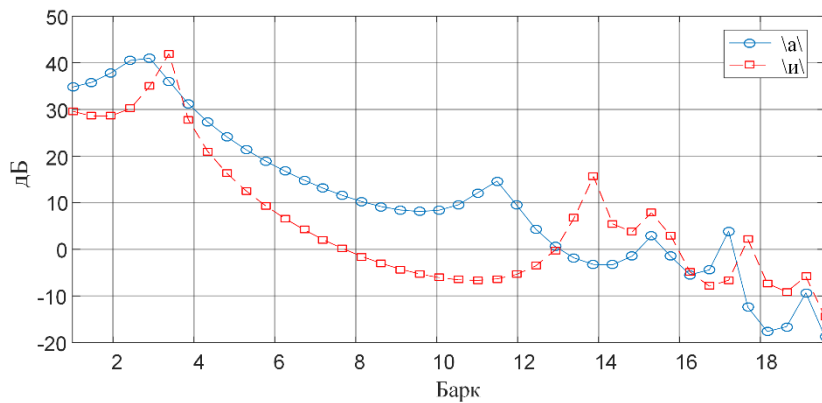


Рисунок 5– Взаимное расположение формантных частот звуков /а/ и /и/ у человека с БАС

В случае, когда правильный порядок не нарушен, то часто имеет место существенное сближение формантных частот. Для количественной оценки степени нарушения взаимной формантной структуры звуков /а/ и /и/ предлагается использовать следующую меру:

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

Функция (1) возвращает значение равное 2 в случае, когда нарушено правильное взаимное расположение формант, т.е. либо  $F_i(1) > F_a(1)$ , либо  $F_a(2) > F_i(2)$ . Замечено, что в норме расстояние между формантами звуков /а/ и /и/ превышает 2 барка, поэтому степень сближения формант оценивается функцией  $f_{mt_{err}}(F_i, F_a)$  при условии, что расстояние между первыми формантами и/или между вторыми формантами звуков /а/ и /и/ составляет менее 2-х барков.

Помимо изменений во взаимной формантной структуре звуков /а/ и /и/ у пациентов с БАС наблюдалось увеличение сходства между формантами огибающих этих звуков. Типичный пример огибающих с высокой степенью похожести дан на рисунке 6.

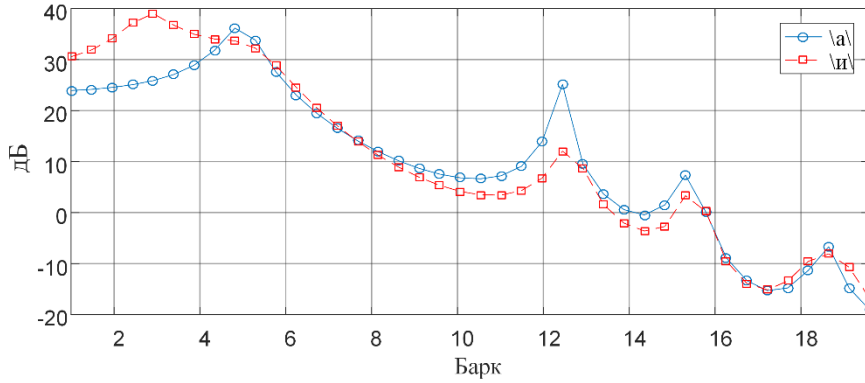


Рисунок 6– Схожесть огибающих звуков /а/ и /и/ у пациента с БАС

Для количественной оценки различия между огибающими звуков /а/ и /и/ предлагается использовать  $l_1$ -норму расстояния:

$$d_1(E_i, E_a) = \sum_{k=1}^P |E_i(k) - E_a(k)|, \quad (2)$$

где  $E_i$  – огибающая звука /и/,  $E_a$  – огибающая звука /а/,  $P$  – число точек в частотной области барков, в которых определена огибающая.

Таким образом, в результате анализа тестовых записей формировались двухкомпонентные характеристические вектора признаков  $\mathbf{x} = [d_1(E_i, E_a); \text{fmt}_{err}(F_i, F_a)]^T$ , на основании которых впоследствии строился классификатор.

**Классификация методом линейный дискриминантного анализа (ЛДА).** Для классификации использовался ЛДА с применением критерия Фишера [5]. Идея линейного дискриминантного анализа заключается в поиске такой гиперплоскости  $\mathbf{w}$  в пространстве признаков, чтобы проекция всех обучающих векторов на неё минимизировала внутриклассовую дисперсию и максимизировало бы межклассовую дисперсию:

$$\mathbf{w} = \arg \max_{\mathbf{w}} \frac{\mathbf{w} \mathbf{S}_B \mathbf{w}^T}{\mathbf{w} \mathbf{S}_W \mathbf{w}^T},$$

где  $\mathbf{S}_B$  и  $\mathbf{S}_W$  – матрицы межклассового и внутриклассового рассеяния, соответственно:

$$\mathbf{S}_B = (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T,$$

$$\mathbf{S}_W = \sum_{j=1}^2 \sum_{\mathbf{x}} (\mathbf{x} - \boldsymbol{\mu}_j)(\mathbf{x} - \boldsymbol{\mu}_j)^T,$$

где  $\boldsymbol{\mu}_1$  – среднее значение характеристических векторов здоровых людей,  $\boldsymbol{\mu}_2$  – среднее значение характеристических векторов людей с речевыми нарушениями. Более подробное описание ЛДА приводится в [5].

**Результаты экспериментов.** Запись речи пациентов с БАС проводилась на базе Республиканского научно-практического центра неврологии и нейрохирургии (Минск, Беларусь). Всего была выполнена запись тестовых речевых образцов у 29 лиц, 13 из которых были здоровыми, а 16 имели БАС. Из каждой записи выделялись участки содержащие звуки /а/ и /и/, по которым строились огибающие этих звуков ( $E_a(k)$  и  $E_i(k)$ ) и выделялись форматные частоты ( $F_a$  и  $F_i$ ). Для данных параметров рассчитывались выражения (1) и (2) и формировался характеристический вектор признаков  $\mathbf{x}$ . На рисунке 7 показан график рассеяния характеристических векторов, полученных из всей тестовой базы.

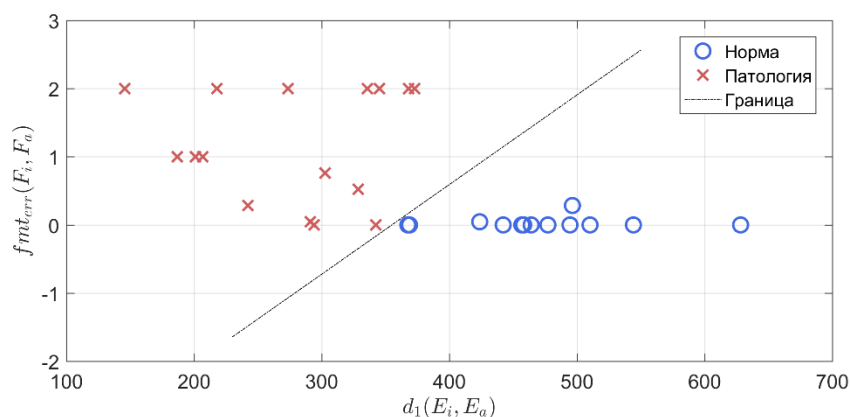


Рисунок 7– Зависимость степени нарушения взаимного расположения формант и различия огибающих звуков /а/ и /и/ для всех 29 субъектов из тестовой базы

Анализ рисунка 7 показывает, что предложенные признаки  $f_{mt\_err}(F_i, F_a)$  и  $d_1(E_i, E_a)$  можно использовать для выявления патологических нарушений в голосе.

#### Вывод.

В работе рассмотрена задача выделения признаков неврологических заболеваний в речевом сигнале. Предложено для решение указанной задачи выполнять совместный анализ огибающих различных гласных звуков. Показано, что огибающие гласных звуков /а/ и /и/ имеют высокую степень схожести между собой при дизартрических нарушениях речи. Выявлено, что часто наличие неврологического заболевания приводит к изменению взаимного расположения формант звуков /а/ и /и/.

#### Литература

1. Spangler T., Fractal features for automatic detection of dysarthria / T. Spangler, N.V. Vinodchandran, S. Samal and J.R. Green // Proc. of IEEE EMBS International conference on biomedical & health informatics, Florida, USA, Feb. 16-19, 2017. – pp. 437-440.
2. Green J.R., Bulbar and speech motor assessment in ALS: challenges and future directions // Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. – vol. 14, No. 7-8, – P. 494–500.
3. Меженная М.М., Рушкевич Ю.Н., Борискевич А.А. Частотно-временной анализ речевых сигналов в диагностике бульбарных нарушений // Доклады БГУИР, 2016. – №7 (101). – С. 260–264.
4. Михайлов В.Г., Златоустова Л.В. Изменение параметров речи/ Под ред. М.А. Сапожкова. – М.: Радио и связь, 1987. – 168 с.
5. Рылов А.С. Анализ речи в распознающих системах. – Мн.: Бестринт, 2003. – 264 с.

## FEATURE EXTRACTION FROM SPEECH SIGNAL FOR NEUROLOGICAL DISEASES DIAGNOSIS

Maxim Vashkevich, Yuliya Rushkevich\*, Elias Azarov, Alexander Petrovsky

Belarusian State University of Informatics and Radioelectronics

\*Republican Research and Clinical Center of Neurology and Neurosurgery

**Abstract.** A method for estimating pathological changes in the speech signal for the diagnosis of neurological diseases is proposed. Detection of deviations is performed by joint analysis of different vowel sounds extracted from continuous-speech test signal. The results of experiments show the practical applicability of proposed method of assessment of pathology in voice.