

УДК 004.934

## СИСТЕМА МЕДИЦИНСКОЙ ДИАГНОСТИКИ ГОЛОСОВЫХ СИГНАЛОВ НА ОСНОВЕ МЕТОДОВ ОБРАБОТКИ, СИНХРОНИЗИРОВАННЫХ С ЧАСТОТОЙ ОСНОВНОГО ТОНА

Пастернак В. В., студент гр.950702

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

Вашкевич М.И. – канд. техн. наук

**Аннотация.** В работе описывается алгоритм классификации пациентов с боковым амиотрофическим склерозом (ALS) на основе анализа звуковых сигналов, полученных при произнесении продолжительных гласных звуков. Для этого был использован метод вычисления формантных параметров звуковых сигналов, основанных на дискретном преобразовании Фурье и анализе спектральных пиков, соответствующих формантам. Алгоритм был проверен на базе голосовых записей здоровых людей и пациентов с боковым амиотрофическим склерозом. Полученные значения частот и амплитуд формантных пиков были использованы в качестве признаков для классификации пациентов.

**Ключевые слова.** Боковой амиотрофический склероз, БАС, акустический анализ, форманты, спектральные пики, дискретное преобразование Фурье.

### Введение

В настоящее время боковой амиотрофический склероз (ALS) является серьезной проблемой в области неврологии. Он характеризуется постепенной дегенерацией моторных нейронов, что приводит к прогрессирующей слабости мышц и нарушению речи. В связи с этим, разработка эффективных методов диагностики и мониторинга ALS является актуальной задачей.

В последние годы акустический анализ звуковых сигналов, получаемых при произнесении голосовых звуков, стал широко используемым инструментом для оценки функционального состояния голосового аппарата и диагностики ряда заболеваний, включая ALS. Одним из методов анализа является вычисление формантных параметров звуковых сигналов.

### Анализ гармонической структуры гласных

Гармоническая структура продолжительных гласных звуков была признана важной и информативной характеристикой для идентификации патологий голоса. Неполное закрытие голосовой щели во время фонации, позволяющее воздуху выходить, является одним из факторов, делающих голос более шумным. В частности, для гласной /a/ это приводит к нарушению гармонической структуры: амплитуда первой гармоники (H1) становится выше второй (H2).

Одним из важных аспектов качества голоса является стабильность гармонической структуры в процессе фонации. Оценка гармонической структуры может рассматриваться как характеристика для описания возбуждающего источника (силы, приводящей к производству звука). Сложность оценки гармонических параметров заключается в том, что они зависят от основной частоты  $F_0$ . В данном исследовании проводился анализ голоса на основе фиксированного количества основных периодов. Было обращено внимание на извлечение среднего значения и стандартного отклонения (SD) первых десяти гармоник гласных. Для заданного голосового сигнала  $s(n)$  процесс анализа может быть описан следующими шагами [1]:

1. Заданный голосовой сигнал  $s(n)$  разбивается на периоды основного тона с помощью метода сопоставления волновых форм с ограничением фазы (рисунок 1).

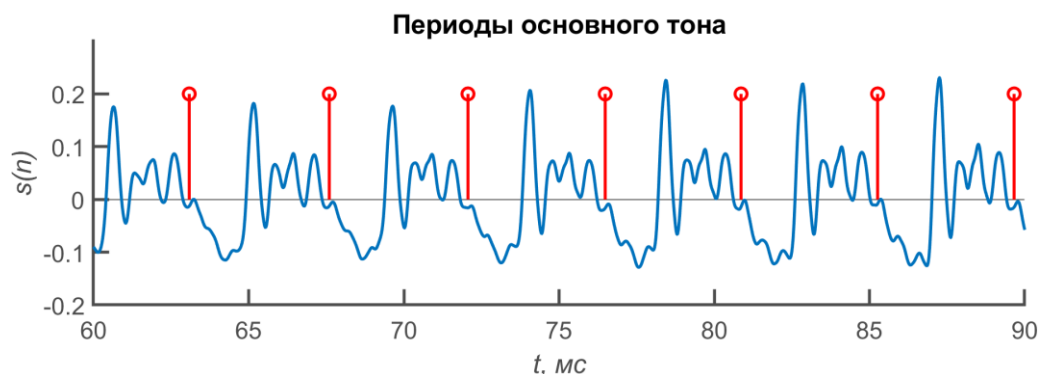


Рисунок 1 – Заданный сигнал с разметкой на периоды основного тона

Метод сопоставления волновых форм использует тот факт, что текущий период связан с предыдущим периодом. Недостатком такого подхода является то, что если возникнет ошибка при обнаружении текущего периода, то это повлияет на все последующие периоды. На практике это приводит к сдвигу фаз, если мы сравниваем циклы, находящиеся друг от друга на расстоянии. Метод сопоставления волновых форм с ограничением фазы избавился от этого недостатка [2].

2. Разделение сигнала  $s(n)$  на  $N_f$  перекрывающихся кадров, содержащих  $N_c$  периодов основного тона с одним периодом перекрытия. Для каждого кадра  $s^i(n)$ ,  $i = 1 \dots N_f$  выполняются шаги 3-5.
3. Интерполяция  $s^i(n)$  в  $I * N_c$  равноудалённых временных точек:  $s^i(n) \rightarrow \hat{s}^i(m)$ .
4. Применение окна Хэмминга  $h(m)$  к интерполированному кадру (рисунок 2) и вычисление дискретного преобразования Фурье (ДПФ) (рисунок 3):

$$\hat{S}^i(k) = \text{ДПФ}[\hat{s}^i(m) h(m)].$$

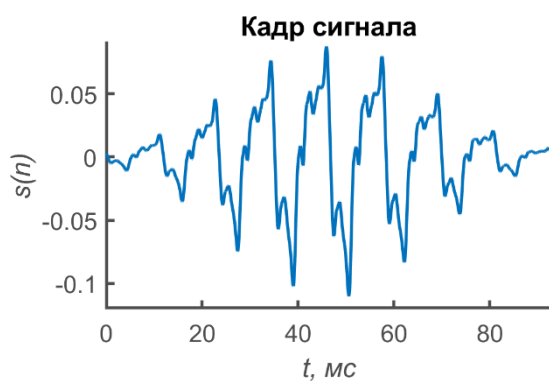


Рисунок 2 – Интерполированный кадр с применённым окном Хэмминга

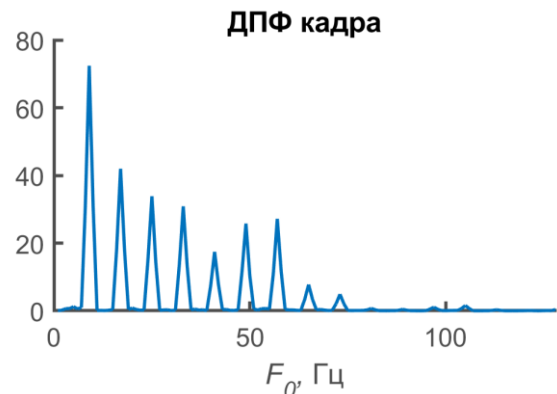


Рисунок 3 – Дискретное преобразование Фурье кадра

5. Извлечение гармонических амплитуд:  $h_p(i) = |\hat{S}^i(p * I)|$   $p = 1, 2 \dots 10$ .
6. Масштабировать гармонические амплитуды следующим образом:

$$\tilde{H}_p(i) = 20 \log_{10} \left( \frac{h_p(i)}{\max_{p \in [1, 10], i \in [1, N_f]} \{h_p(i)\}} \right).$$

7. Рассчитать среднее значение и стандартное отклонение для масштабированных гармонических амплитуд:

$$Hp^\mu = E\{\tilde{H}_p\}, Hp^\sigma = \sqrt{E\{(\tilde{H}_p - Hp^\mu)^2\}}.$$

8. Рассчитать дополнительный признак – обратное от суммы абсолютных значений  $Hp^\mu$  и  $Hp^\sigma$ :

$$RelHp = \frac{1}{|Hp^\mu| + Hp^\sigma}.$$

Причина расчёта этого признака в том, что сильные и стабильные гармоники должны иметь низкие отмасштабированные амплитуды  $Hp^\mu$  и низкие отклонения  $Hp^\sigma$ , а следовательно, высокие значения  $RelHp$ .

Алгоритм анализа гармонической структуры гласных реализован в *Matlab* следующим образом:

```
function [Hp_mean, Hp_SD, RelHp] =
harmonicAnalysis(fileName)
addpath('Troparion-master/IRAPT/IRAPT_web');
addpath('Troparion-master/Perturbation_analysis');

[s, fs] = audioread(fileName);
[Fo, ~, time_marks] = irapt(s, fs, 'irapt1', 'sustain phonation');

[Fo_periods] = WM_phase_const(s, Fo, time_marks, fs);
```

```

Fo_line = cumsum(Fo_periods);

Nc = 8;
Nf = floor((length(Fo_line) - 1)/(Nc - 1));
l = 512*Nc;

frames = zeros(Nf, l);
ffts = zeros(Nf, l);
window = transpose(hamming(l));
j = 1;
for i = 0:Nc-1:length(Fo_line)-(Nc)
    if i == 0
        temp = s(1:Fo_line(Nc));
    else
        temp = s(Fo_line(i):Fo_line(i+Nc));
    end
    k = linspace(1, length(temp), l);
    temp_interp = interp1(1:length(temp), temp, k);
    filtered = temp_interp .* window;
    frame_fft = fft(filtered);

    ffts(j, :) = abs(frame_fft);
    frames(j, :) = filtered;
    j = j + 1;
end
%% Harmonics
harm = 10;
hp = ffts(:,9:8:(1+8*harm));
Hp = 20*log10(hp/(max(hp, [], 'all')));

%% Rates

Hp_mean = mean(Hp);
Hp_SD = sqrt(mean((Hp - Hp_mean).^2));

RelHp = 1./(abs(Hp_mean) + Hp_SD);

end

```

### Описание эксперимента

Для проведения эксперимента использовалась база голосовых записей здоровых людей и пациентов с боковым амиотрофическим склерозом. Результатом прохождения каждой голосовой записи через алгоритм были 30 значений признаков –  $Hp^\mu$ ,  $Hp^\sigma$  и  $RelHp$  для каждой из 10 гармоник. В эксперименте использовались значения параметров  $N_c = 8$  и  $l = 512$ .

Результаты эксперимента для голосовых записей здоровых людей и пациентов с БАС представлены на рисунках 4-9.

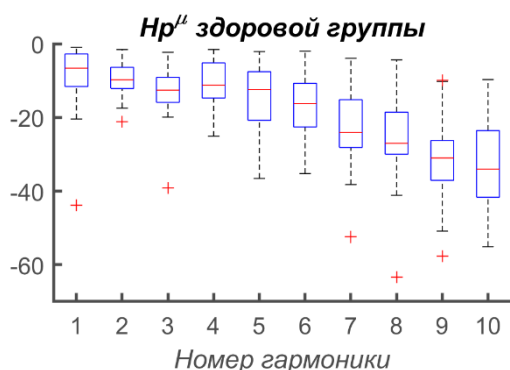


Рисунок 4 – Значения  $Hr^\mu$  группы здоровых людей

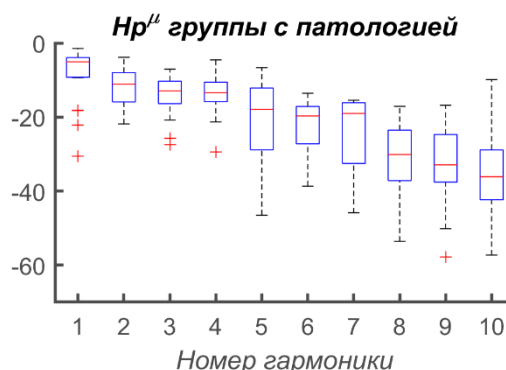


Рисунок 5 – Значения  $Hr^\mu$  группы пациентов с БАС

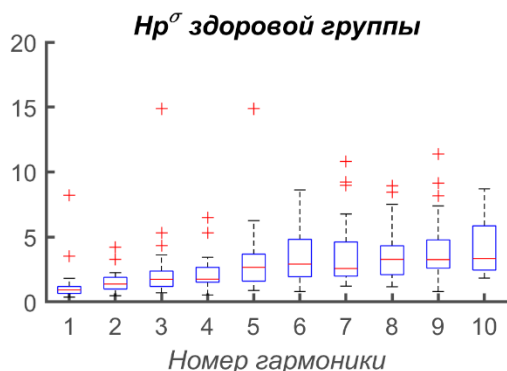


Рисунок 6 – Значения  $Hr^\sigma$  группы здоровых людей

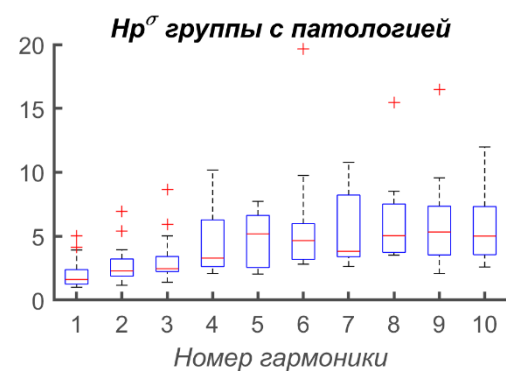


Рисунок 7 – Значения  $Hr^\sigma$  группы пациентов с БАС

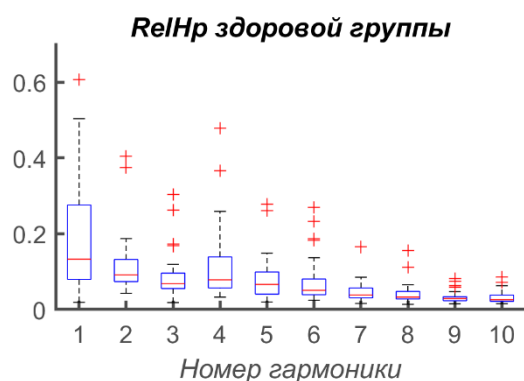


Рисунок 8 – Значения RelHr группы здоровых людей

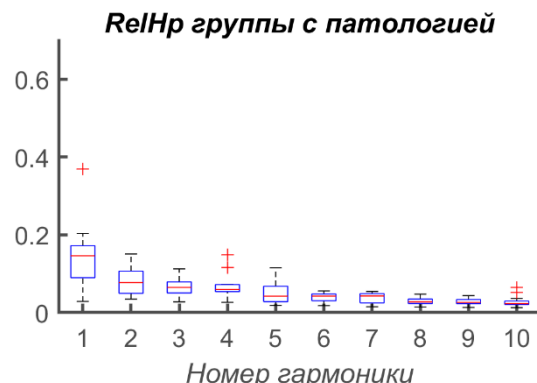


Рисунок 9 – Значения RelHr группы пациентов с БАС

Таким образом, в работе рассмотрена система диагностики голосовых сигналов на основе метода анализа гармонической структуры гласных. Из полученных графиков видно, что полученные признаки гармонических структур гласных у здоровых людей и пациентов с боковым амиотрофическим склерозом несколько отличаются. Это значит, что в дальнейшем эти признаки могут быть использованы методами машинного обучения для диагностики различных заболеваний. Работоспособность метода показана путём MATLAB-моделирования.

**Список использованных источников:**

1. Classification of ALS patients based on acoustic analysis of sustained vowel phonations / Vashkevich M., Rushkevich Y. // *Biomedical Signal Processing and Control*, 2021. – P. 65, 1-14.
2. M. Vashkevich, A. Petrovsky, Y. Rushkevich, Bulbar ALS detection based on analysis of voice perturbation and vibrato, in: *Proc. of Signal Processing: Algorithms, Architectures, Arrangements, and Applications, SPA, 2019*, pp. 267–272

# MEDICAL DIAGNOSTIC SYSTEM BASED ON VOICE SIGNAL PROCESSING METHODS SYNCHRONIZED WITH FUNDAMENTAL FREQUENCY

*Pasternak V.V.*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*Vashkevich M.I. – PhD*

**Annotation.** The paper describes an algorithm for classifying patients with amyotrophic lateral sclerosis (ALS) based on the analysis of sound signals obtained from the pronunciation of sustained vowel sounds. The method used for this was the computation of formant parameters of sound signals based on discrete Fourier transform and analysis of spectral peaks corresponding to formants. The algorithm was tested on a database of voice recordings from healthy individuals and patients with ALS. The obtained values of formant peak frequencies and amplitudes were used as features for patient classification.

**Keywords.** Amyotrophic lateral sclerosis, ALS, acoustic analysis, formants, spectral peaks, discrete Fourier transform.