

Система детектирования речевых нарушений у пациентов с БАС на основе теста на протяжное произнесение звука [а]¹

Резюме

В статье рассмотрен подход к выявлению речевых нарушений у пациентов с боковым амиотрофическим склерозом (БАС) на основе акустического анализа голосового сигнала. Целью работы являлась проверка пригодности речевого задания на протяжное произнесение гласного звука /а/ для автоматического выявления пациентов с БАС. Рассмотрены четыре группы информационных признаков голосового сигнала, получающиеся в результате анализа контура частоты основного тона (ЧОТ): джиттер, шиммер, статистические параметры ЧОТ и частотно-временные параметры ЧОТ. Проведенные эксперименты показали, что использование предложенных методов акустического анализа голоса и классификации по методу к ближайших соседей позволило получить систему детектирования речевых нарушений у пациентов с БАС, обладающую точностью на уровне 95,7% (при чувствительности 91,5% и специфичности 97,4%).

Ключевые слова: боковой амиотрофический склероз, акустический анализ голоса, классификация

Введение

Боковой амиотрофический склероз (БАС) является нейродегенеративным неизлечимым заболеванием, приблизительно 50% пациентов с БАС умирают в течение 30 месяцев с момента появления симптомов [1]. В среднем на постановку диагноза БАС уходит более года [2]. Использование акустического анализа голоса и речи является перспективным способом совершенствования процесса выявления БАС и мониторинга прогрессирования данного заболевания [3]. Это становится возможным, так как бульбарные моторные изменения (т. е., трудности с речью или глотанием) являются первыми симптомами примерно у 30% лиц с БАС [4], а на более поздних стадиях бульбарные симптомы появляются почти у всех пациентов с БАС. Важно, что атипичные значения акустических параметров голоса были выявлены у пациентов БАС при помощи теста на протяжное произнесение гласного звука, в то время как в их разговорной речи не чувствовалось ощутимых изменений [5].

Необходимость разработки средств диагностики голоса и речи пробудила интерес к разработке специальных методов акустического анализа. Одним из важных аспектов современных исследований, связанных с выявлением неврологических заболеваний (таких как болезнь Паркинсона и БАС), является то, что они нацелены на использование смартфонов и планшетов для записи голоса с помощью стандартного микрофона в различных домашних условиях [6, 7]. Продвижение в этом направлении может привести к созданию простых в использовании средств для детектирования, отслеживания и телемониторинга заболеваний.

В предыдущих исследованиях для детектирования БАС использовались различные подходы и речевые задания. Одним из общих подходов, стремящихся к решению проблемы дифференциальной диагностики, является классификация типа дизартрии по речевому/го-

¹ Вашкевич М.И., Рушкевич Ю.Н. Система детектирования речевых нарушений у пациентов с боковым амиотрофическим склерозом на основе теста на протяжное произнесение звука [а] // Актуальные проблемы неврологии и нейрохирургии. Рецензируемый сборник научных трудов. Выпуск 22. Под ред. д.м.н. Р.Р. Сидорова и д.м.н., проф. С.А. Лихачева. – С. 278–290.

лосовому сигналу [8, 9]. Основная трудность этого подхода заключается в том, что он требует сбора репрезентативной выборки аннотированных данных для всех типов дизартрии. Более простой подход к обнаружению БАС основан на использовании речевой базы, содержащей патологические и нормальные образцы голоса/речи [4, 7, 10-12]. В ряде исследований обнаружение БАС осуществляется с использованием кинематических датчиков, размещаемых на артикуляционных органах для последующего моделирования артикуляции и измерения просодических элементов, таких как длительность гласных или скорость речи [2, 4, 13-15].

Цели исследования:

- 1) проверить пригодность простого теста на протяжное произнесение гласного звука /a/ для обнаружения в голосе признаков БАС;
- 2) проверка различных подходов к решению задачи детектирования речевых нарушений при БАС – метода линейного дискриминантного анализа и метода к ближайших соседей;
- 3) выявления набора наиболее информативных акустических признаков для решения задачи детектирования речевых нарушений при БАС.

1 Методы

1.1 Анализ голосового сигнала: базовые понятия

Голосовой сигнал, образующийся в результате прохождения потока воздуха через колеблющиеся голосовые складки, относится к классу *почти периодических сигналов* (англ. *nearly periodic*). Это означает, что в отличие от строго периодических сигналов, для которых один цикл (т.е. форма колебания) неотличим от другого, у почти периодических сигналов повторяющиеся циклы могут иметь незначительные отличия.

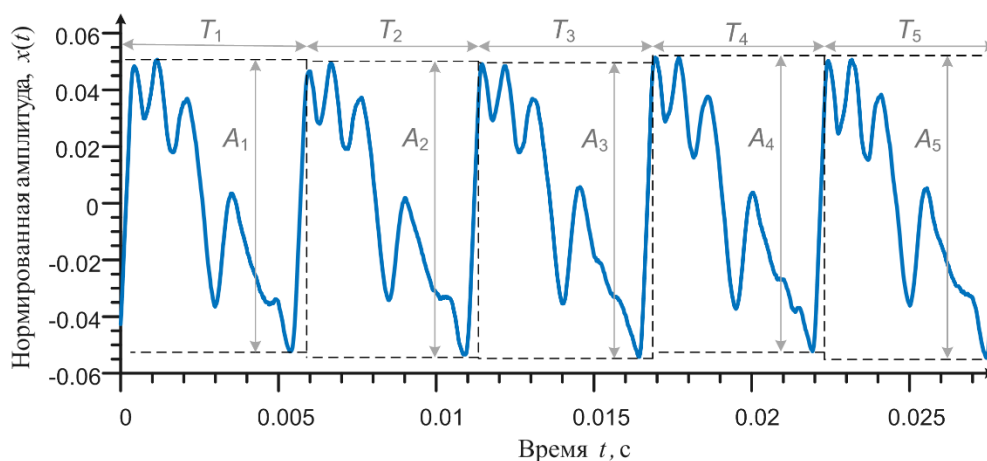


Рисунок 1 – Голосовой сигнал здорового человека с разметкой на периоды основного тона

На рисунке 1 приведен короткий участок записи голоса здорового человека. Видно, что форма колебания повторяется раз за разом с небольшими вариациями. Частоту данных колебаний называют *частотой основного тона* (ЧОТ). Путем разбиения голосового сигнала на периоды основного тона можно получить последовательность значений периодов T_1, T_2, \dots, T_n , а также соответствующую последовательность значений амплитуд A_1, A_2, \dots, A_n

(см. рисунок 1). Анализируя данные последовательности можно получить ряд важных параметров голосового сигнала, о которых будет сказано ниже.

Если вместо периода основного тона перейти к значению частоты основного тона $f_0 = 1/T$ и рассматривать f_0 как функцию зависящую от времени, то мы получим контур частоты основного тона $f_0(t)$. Поскольку единицей измерения периода T является секунда, то значения f_0 измеряются в Герца. Пример контура ЧОТ приведен на рисунке 2. Анализируя функцию $f_0(t)$ также можно получить ряд важных акустических признаков, позволяющих проводить разделение голосов на норму и патологию.

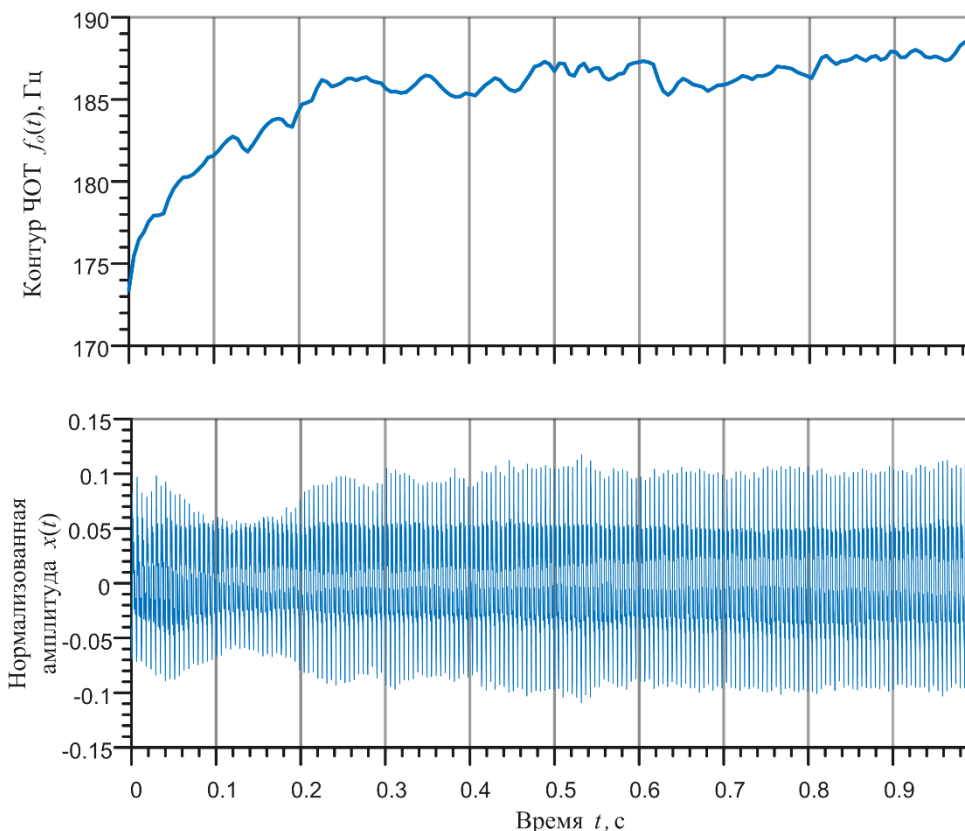


Рисунок 2 – Пример контура частоты основного тона, полученный для голоса здорового человека

1.2 Амплитудные и частотные пертурбационные параметры голоса

Джиттер - это мера вариативности периода основного тона. Поскольку джиттер оценивает краткосрочную вариацию, она не может быть отнесена к контролируемым (произвольным) изменениям частоты основного тона (ЧОТ). Таким образом, джиттер является показателем показателя стабильности работы фонаторной системы. Высокий уровень джиттера возникает в результате нарушения нейромоторной функции [16]. В простейшем случае, джиттер определяется как средняя разница между последовательными периодами, деленная на среднее значение периода основного тона.

$$J_{loc} = \frac{\frac{1}{N-1} \sum_{i=2}^N |T_i - T_{i-1}|}{\frac{1}{N} \sum_{i=1}^N |T_i|}, \quad (1)$$

где T_i – длительность i -го периода основного тона, а N – количество периодов основного тона в сигнале.

В более общем случае, для оценки джиттера используют среднюю разницу между текущим значением периода и усредненным значением периода на нескольких смежных циклах. Такая мера называется коэффициентом возмущений периодов (англ. PPQ – period perturbation quotient):

$$J_L = \frac{\frac{1}{N-L+1} \sum_{i=1+(L-1)/2}^{N-(L-1)/2} \left| T_i - \frac{1}{L} \sum_{n=i-(L-1)/2}^{i+(L-1)/2} T_n \right|}{\frac{1}{N} \sum_{i=1}^N |T_i|}, \quad (2)$$

где L – параметр, определяющий, сколько периодов используется для вычисления «скользящего» среднего. Чаще всего используют значения $L = 3, 5$ и 55 [16,17].

Шиммер – это мера, характеризующая степень вариативности амплитуды акустических колебаний в процессе фонации. В простейшем случае шиммер (S_{loc}) определяется как средняя абсолютная разница между амплитудами последовательных периодов основного тона, деленная на среднюю амплитуду, по аналогии с выражением (1) [16]. Однако, на S_{loc} влияет постепенный равномерный (естественный) спад интенсивности голоса (см. рисунок 3).

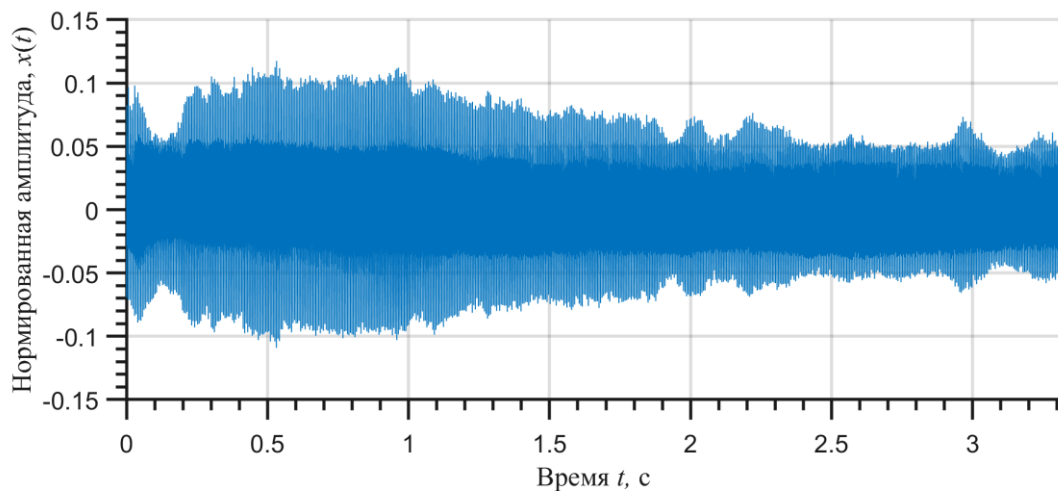


Рисунок 3 – Пример записи здорового человека с естественным спадом интенсивности голоса. Значение $S_{loc} = 3,3\%$, а коэффициент амплитудных возмущений $S_3 = 1,8\%$

Для устранения эффекта "дрейфа" амплитуды сигнала и с целью получения более точной оценки шиммера обычно измеряют коэффициент амплитудных возмущений (англ. APQ – amplitude perturbation quotient) [16]:

$$S_L = \frac{\frac{1}{N-L+1} \sum_{i=1+(L-1)/2}^{N-(L-1)/2} \left| A_i - \frac{1}{L} \sum_{n=i-(L-1)/2}^{i+(L-1)/2} A_n \right|}{\frac{1}{N} \sum_{i=1}^N |A_i|}, \quad (3)$$

где A_i – амплитуда i -го периода основного тона, L – параметр, определяющий, сколько периодов используется для вычисления «скользящего» среднего. Чаще всего используют значения $L = 3, 5, 11$ и 55 [16,17].

Как правило пертурбационные параметры (1)-(3) измеряются в процентах, поскольку абсолютные их значения невелики.

1.3 Энтропия периодов основного тона

В работе [18] предложен информационный признак – энтропия периодов основного тона (англ. PPE – pitch period entropy), который применялся в задаче классификации голосов

пациентов с болезнью Паркинсона. В работе указывалось, что здоровые голоса обладают естественным плавным вибрато или микротремором, которые детектируются при помощи параметра джиттера. Однако, общим симптомом дисфонии при болезни Паркинсона является нарушение способности контролировать стабильность частоты основного тона при протяжной фонации. Поэтому, используя традиционные параметры джиттера трудно отличить естественные, здоровые вариации периода основного тона от дисфонических вариаций, появляющихся в следствии болезни Паркинсона. Поскольку при БАС также нарушается способность контролировать стабильность частоты основного тона при протяжном произнесении звуков, то в данной работе предлагается также использовать этропию периодов основного тона в качестве информационного признака.

1.4 Анализ вибрато

Вибрато – это быстрое и регулярное колебание ЧОТ, возникающее во время продолжительной стабильной фонации. Непосредственный анализ спектрограмм голосов здоровых людей и пациентов БАС показал, что существует значительная разница в частоте у данных групп. Оценка степени патологических изменений вибрато основана на наблюдении, что для здоровых голосов вибрато лежит в диапазоне 5-8 Гц [19], в то время как при БАС характерно наличие высокочастотных составляющих в диапазоне 9-14 Гц [20]. В данном исследовании мы использовали метод оценки степени патологии вибрато (англ. PVI – *pathology vibrato index*), предложенный в [21].

1.5 Дополнительные признаки, получаемые путем анализа ЧОТ

В качестве информационного признака может использоваться параметр среднеквадратичного отклонения (СКО) частоты основного тона – SD_{f_o} [16].

Также использовался частотный диапазон фонации (англ. PFR – *phonatory frequency range*), определяемый как

$$PFR = \frac{\log_{10} \frac{\max f_o}{\min f_o}}{\log_{10} 2},$$

который характеризует отношение максимального значения ЧОТ к минимальному во время фонации.

В работе анализировался направленный коэффициент возмущений (англ. DPF – *directional perturbation factor*) [17]:

$$DPF = \frac{N_{\Delta\pm}}{N} \times 100,$$

где $N_{\Delta\pm}$ – число периодов на которых разница между последовательными циклами основного тона изменяла свой знак.

1.6 Классификация

Для различения двух классов нормальных и патологических голосов использовались два распространенных подхода машинного обучения: метод линейного дискриминантного анализа (ЛДА) и метод k -ближайших соседей [22]. Для оценки качества классификаторов использовался метод перекрестной проверки по K -блокам (англ. *K-fold cross-validation*), которая заключается в следующем. Исходный набор данных перемешивается случайным образом и разбивается на K блоков. Далее выполняется обучение классификатора, причем один из блоков выступает, как тестовый набор, а оставшиеся $K - 1$ в совокупности составляют обучающий набор. Эта процедура повторяется K раз, так, чтобы каждый блок один

раз выступил в роли тестового набора. Метки, присвоенные классификаторами, для тестовых наборов сохраняются и по ним выполняется оценка производительности классификатора [22]. В качестве основных характеристик классификатора вычислялись точность (Acc), чувствительность (Sens), специфичность (Spec) и средняя полнота (R_{ave}). Ниже приводятся выражения для их вычисления:

$$\text{Acc} = \frac{TP + TN}{TP + FP + TN + FN}, \quad \text{Sens} = \frac{TP}{TP + FN}, \quad \text{Spec} = \frac{TN}{TN + FP},$$
$$R_{ave} = \frac{1}{2} (\text{Sens} + \text{Spec}),$$

где TP, TN, FP, FN – истинно положительные, истинно отрицательные, ложно положительные, ложно отрицательные результаты классификации (под положительным результатом понимают верную постановку диагноза БАС). Точность характеризует частоту правильных решений, выносимых классификатором. Однако данный параметр не дает полного представления о качестве работы классификатора, особенно если в исходном наборе данных классы здоровых голосов и голосов с патологией не сбалансированы. Параметр чувствительности показывает способность классификатора детектировать патологию, если она есть. Специфичность характеризует способность классификатора определять отсутствие патологии, когда она действительно отсутствует. Средняя полнота – показывает общую способность классификатора относить голос к правильному классу. Процедура кросс-проверки повторялась 40 раз, чтобы обеспечить независимость оценок точности классификатора от разбиения исходной выборки на блоки. Для параметров точности, чувствительности и специфичности вычислялись средние значения и СКО.

2 Материалы

Записи голосов, использованные в данном исследовании, были собраны в Республиканском научно-клиническом центре неврологии и нейрохирургии (Минск, Беларусь). Всего было записано 54 человека, из них 39 здоровых (23 мужчины, 16 женщин) и 15 больных БАС с признаками бульбарных нарушений (6 мужчин, 9 женщин). Средний возраст в здоровой группе составил 41,9 лет (СКО 16,3), а средний возраст в группе БАС – 57,7 лет (СКО 9,0). Всем участникам было предложено произносить протяжный гласный звук [a] как можно дольше с комфортной высотой и громкостью. Фонация производилась на одном дыхании. Голосовые сигналы записывались с помощью смартфона с гарнитурой (частота дискретизации 44,1 кГц) и сохранялись в виде несжатых 16-битных wav-файлов. Средняя продолжительность записей составила 4,1 с. База голосов и Matlab-функции, используемые для анализа голоса, размещены в общедоступном репозитории².

3 Результаты и обсуждение

3.1 Статистический анализ акустических признаков

Для визуализации параметров голоса в контрольной группе и группе пациентов с БАС были рассчитаны статистические графики нескольких признаков. Для описания распределения применялись диаграммы размаха, которые позволяют увидеть медиану распределения и интервал, в который попадает 25% и 75% всех данных, а также максимальное и минимальное значение в выборке; также строился график плотности вероятности, получаемый путем сглаживания с ядром (англ. kernel smoothing).

² <https://github.com/Mak-Sim/Troparion>

На рисунке 4 представлены статистические графики коэффициента возмущений периодов J_3 и коэффициента амплитудных возмущений S_3 . Хотя по-видимому значения данных параметров в группах здоровых и пациентов с БАС сильно перекрываются, тем не менее данные параметры часто отбирались при автоматическом поиске наилучшего вектора информационных признаков. Это может говорить о том, что, хотя сами по себе эти параметры не позволяют выполнить разделение на больных и здоровых, они могут содержать дополнительную информацию, способную улучшить качество классификации.

На рисунке 5 приведены статистические графики направленного коэффициента возмущений и частотного диапазона фонации. В группе БАС медианное значение частотного диапазон фонации значительно выше, чем в группе здоровых. Это может объясняться тем, что пациенту с БАС сложнее сохранять стабильность ЧОТ во время фонации, что приводит к кратковременным отклонениям в меньшую или большую сторону от среднего значения ЧОТ.

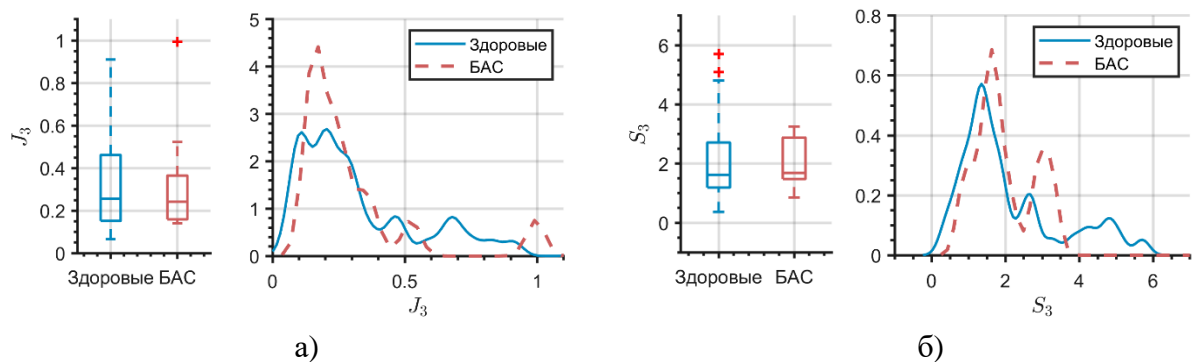


Рисунок 4 – Диаграмма размаха и плотность вероятности: а) коэффициент возмущения периодов ЧОТ J_3 ; б) коэффициент амплитудных возмущений S_3

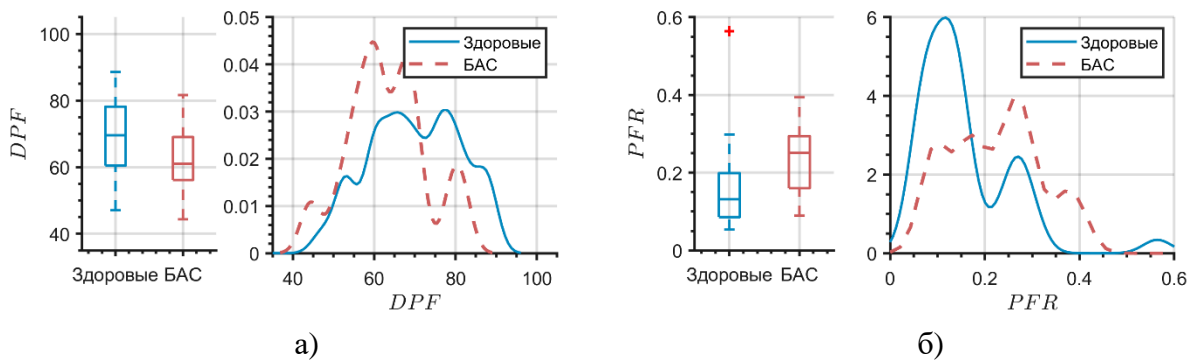


Рисунок 5 – Диаграмма размаха и плотность вероятности: а) направленный коэффициент возмущений DPF ; б) частотный диапазон фонации PFR

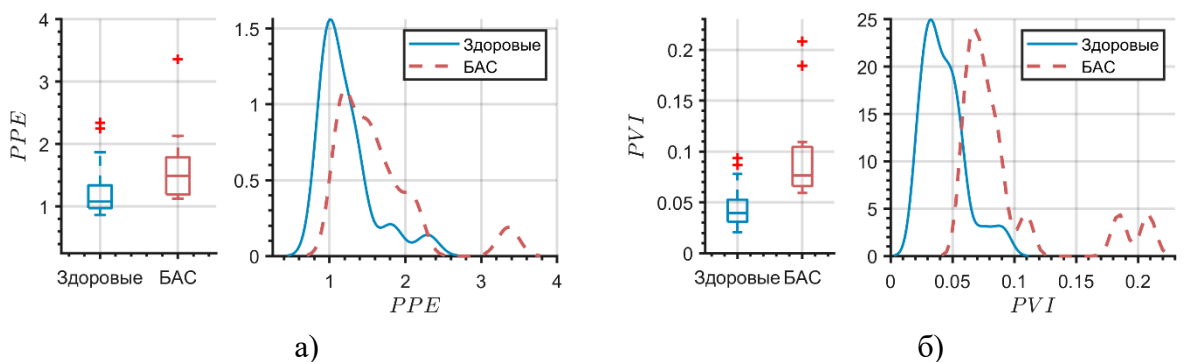


Рисунок 6 – Диаграмма размаха и плотность вероятности: а) энтропии периодов основного тона PPE ; б) степень патологии вибрато PVI

На рисунке 6 приведены статистические графики энтропии периодов основного тона и степени патологии вибрато. В группе пациентов БАС медианное значение энтропии периодов основного тона значительно выше, чем в группе здоровых. Показатель степени патологии вибрато показал наилучшую способность к разделению исходной выборки на здоровых и больных.

3.2 Результаты классификации

Были проведены эксперименты по классификации собранной базы голосов по методу кросс-проверки по K блокам ($K = 6$) с использованием классификатора на основе линейного дискриминантного анализа и метода k ближайших соседей. Вектор информационных признаков состоял из следующих четырнадцати параметров, которые можно разбить на четыре категории:

- джиттеры ($J_{loc}, J_3, J_5, J_{55}$);
- шиммеры ($S_{loc}, S_3, S_5, S_{11}, S_{55}$);
- статистические параметры ЧОТ (SD_{f_0}, PFR);
- параметры, полученные путем анализа контура ЧОТ, т.е. частотно-временные параметры (DFP, PPE, PVI).

Размерность вектора информационных признаков $D = 14$ в решаемой задаче является относительно низкой, что позволило провести полный перебор всех возможных вариантов подмножеств признаков ($2^D - 1$ вариантов), чтобы найти лучший из них.

Таблица 1 – Результаты классификации методом ЛДА

Вектор информационных признаков	Средняя полнота, %	Точность, %	Чувствительность, %	Специфичность, %
$S_{loc} S_3 S_{11} S_{55} PVI$	89,9	91,3±1,9	86,7±1,5	93,1±2,4
$S_{loc} S_3 S_{11} PVI$	89,8	91,0±2,5	87,0±1,4	92,6±3,6
$S_3 S_5 PVI DFP$	89,6	89,7±2,6	89,7±5,0	89,7±2,9
$S_3 S_5 PVI$	89,3	90,2±2,9	87,1±4,8	91,5±3,6
$S_3 PVI$	87,9	86,9±2,4	90,0±4,5	85,8±2,6

В таблице 1 приведены результаты классификации с использованием метода ЛДА. Результаты отсортированы по убыванию средней полноты, потому что наша цель получить классификатор, обладающий одновременно высокой специфичностью и чувствительностью. Первые три строчки в таблице соответствуют трем лучшим найденным комбинациям признаков, четвертая строка – лучшая комбинация, состоящая из трех признаков, а пятая строка – лучшая комбинация из двух признаков. Нужно отметить тот факт, что все наборы признаков, представленные в таблице 1 включают параметр шиммера S_3 и степени патологии вибрато (PVI).

Таблица 2 – Результаты классификации методом k ближайших соседей

Вектор информационных признаков	Средняя полнота, %	Точность, %	Чувствительность, %	Специфичность, %
$J_3 PVI PPE$	94,4	95,7±2,5	91,5±3,0	97,4±2,6
$J_5 PFR PVI$	92,1	94,8±1,2	86,0±2,9	98,2±1,2

J_3 PFR PVI	91,9	94,5±0,8	86,0±2,0	97,7±0,9
J_3 PVI	88,4	91,9±2,1	80,5±5,3	96,2±2,3

В таблице 2 приведены результаты классификации с использованием метода к ближайших соседей, результаты также отсортированы по убыванию средней полноты. Первые три строчки в таблице соответствуют трем лучшим найденным комбинациям признаков, четвертая строка – лучшая комбинация, состоящая из двух признаков. Следует заметить, что, как и в случае с классификатором на основе ЛДА, во всех приведенных комбинациях признаков присутствует параметр PVI, что говорит о его ключевом значении для правильной классификации.

Общей особенностью классификаторов с использованием метода к ближайших соседей является то, что их специфичность значительно превышает чувствительность. В целом, если судить по значению средней полноты, то классификация методом к ближайших соседей более предпочтительна, так как позволяет достигнуть значения $R_{ave} = 94,4\%$ по сравнению с 89,9% в методе ЛДА.

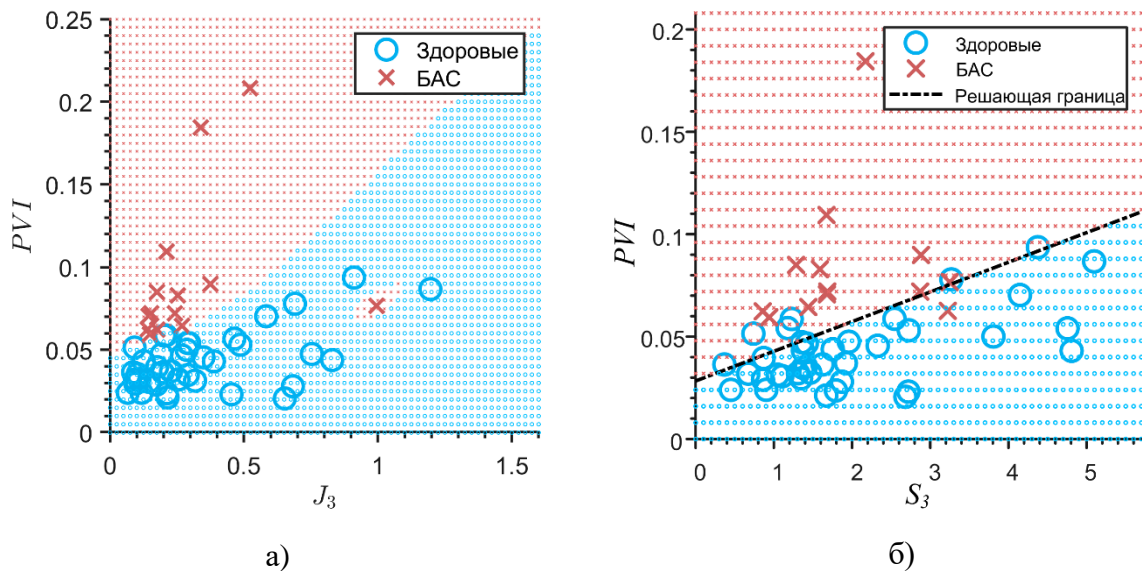


Рисунок 7 – Пример классификации: а) методом к ближайших соседей; б) методом ЛДА

Чтобы объяснить разницу между рассмотренными подходами обратимся к графической интерпретации процесса классификации. На рисунке 7 приведены примеры классификации по методу к ближайших соседей и методу ЛДА для случая, когда характеристический вектор признаков состоит из двух компонент. Из рисунка 7,б видно, что в методе ЛДА все векторное пространство разбивается на две части, после чего решение о том какому классу принадлежит вектор новых данных принимается в зависимости от того в какую область он попадает. В случае метода к ближайших соседей решающая граница имеет нелинейный вид, кроме того, количество областей на которые разбивается все векторное пространство может быть больше двух (см. рисунок 7,а). Именно данной особенностью (т.е. большей гибкостью) объясняется преимущество метода к ближайших соседей над ЛДА.

Заключение

В исследовании представлен подход к анализу голосового сигнала для обнаружения речевых нарушений при БАС. Рассмотрены два подхода к классификации голосовых сигналов – метод линейного дискриминантного анализа и метод к ближайших соседей.

Наилучший достигнутый результат - точность 95,7% (чувствительность 91,5% и специфичность 97,4%) получен с применением метода к ближайших соседей. Наиболее релевантным признаком для построения классификаторов явился параметр степени патологии вибрато (PVI). Представленные методы анализа голоса также могут быть полезны при распознавании различных степеней тяжести БАС. В будущем планируется изучить и другие заболевания, сопровождающиеся речевыми нарушениями, чтобы определить насколько показатель PVI специфичен именно для БАС.

Литература

1. Amyotrophic lateral sclerosis / M.C. Kiernan [et al.] // *Lancet*. – 2011. – vol.377 (issue 9769). – P. 942-955.
2. Detection of bulbar ALS using a comprehensive speech assessment battery/ Y. Yunusova [et al.] // *Proceedings of the International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications* – 2013. – P. 217-220.
3. Relation of automatically extracted formant trajectories with intelligibility loss and speaking rate decline in amyotrophic lateral sclerosis / R.L. Horwitz-Martin [et al.] // *Proceedings of Interspeech*. – 2016. – P. 1215–1219.
4. Fractal features for automatic detection of dysarthria / T. Spangler [et al.] // *IEEE EMBS International Conference on Biomedical Health Informatics*. – 2017 – P. 437–440.
5. Tomik, B. Dysarthria in amyotrophic lateral sclerosis: A review / B. Tomik, R.J. Guiloff // *Amyotrophic Lateral Sclerosis* – 2010. – vol. 11, № 1-2. – P. 4-15.
6. Benba, A. Discriminating between patients with Parkinsons and neurological diseases using cepstral analysis / A. Benba, A. Jilbab, A. Hammouch // *IEEE Transactions on Neural Systems and Rehabilitation Engineering* – 2016. – vol. 24, no. 10. – P. 1100–1108.
7. Detection of amyotrophic lateral sclerosis (ALS) via acoustic analysis / R. Norel [et al.] // *Proceedings of Interspeech* – 2018.– P. 377–381.
8. Guerra, C. A modern approach to dysarthria classification / C. Guerra, D. Lovey // *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEMBS)*. – 2003. – vol.~3. – P. 2257–2260.
9. Liss, J. M. Discriminating dysarthria type from envelope modulation spectra / J.M. Liss, S. LeGendre, A. J. Lotto // *Journal of Speech, Language, and Hearing Research*. – 2011. – № 5 (53). – P. 1246-1255.
10. Automatic early detection of amyotrophic lateral sclerosis from intelligible speech using convolutional neural networks/ K. An [et al.] // *Proceedings of Interspeech* – 2018 – P. 1913-1917.
11. Features extraction for the automatic detection of ALS disease from acoustic speech signals/ M. Vashkevich [et al.] // *Proceedings of inter. conf. Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. – 2018. – P. 321-326.
12. Детектирование бульбарных нарушений при боковом амиотрофическом склерозе на основе анализа речевого сигнала / Гвоздович А.Д. [и др.] // *Доклады БГУИР*.– 2018. – № 6 (116) –С. 52-58.
13. Comparison of speech tasks for automatic classification of patients with amyotrophic lateral sclerosis and healthy subjects/ Illa, A. [et al.]// *Proceedings of IEEE Inter. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. – 2018 – P. 6014-6018.
14. Towards Automatic Detection of Amyotrophic Lateral Sclerosis from Speech Acoustic and Articulatory Samples/ J. Wang, [et al.] // *Proceedings of Interspeech*. – 2016 – P. 1195-1199.

15. Classification of bulbar ALS from kinematic features of the jaw and lips: Towards computer-mediated assessment / A. Bandini, [et al.] // Proceedings of Interspeech. – 2017 – P.1819-1823.
16. Baken, R.J., R.F. Orlikoff Clinical measurement of speech and voice, 2nd edition / R.J. Baken, Orlikoff R.F. // Thomson Learning, 2000. – 864 p.
17. Telephony-based voice pathology assessment using automated speech analysis / R. J. Moran [et al.] // IEEE Transactions on Biomedical Engineering – 2006. – vol. 53, no. 3. – P. 468-477.
18. Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease / M. A. Little, [et al.] // IEEE Transactions on Biomedical Engineering. – 2009. – vol. 56, no. 4. – P. 1015-1022.
19. Nakano, T. An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features / T. Nakano, M. Goto, Y. Hiraga // Proceedings of Interspeech – 2006. – P. 1706–1709.
20. Rapid voice tremor, or flutter, in amyotrophic lateral sclerosis / A. E. Aronson [et al.] // Annals of Otology, Rhinology & Laryngology. – 1992. – vol. 101, no. 6. – P. 511–518.
21. Vashkevich, M. Bulbar ALS detection based on analysis of voice perturbation and vibrato / M. Vashkevich, A. Petrovsky, Yu. Rushkevich // Proceedings of inter. conf. Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). – 2019. – P. 267-272.
22. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах : пер. с англ. – М.: ДМК Пресс, 2015. – 400 с.