

характер. Полученные эмпирическим методом значения параметров степени влияния усталости и степени воздействия стрессового фактора в виде звукового сигнала, могут быть использованы в дальнейшем при оценке стрессоустойчивости испытуемых. Предварительные результаты применения предложенной методики оценки влияния стрессовых факторов на параметры внимания свидетельствуют о простоте ее применения и обработке данных. Анализ видов стрессовых факторов показал возможность применения с минимальными изменениями разработанной методики для оценки ухудшения параметров внимания при воздействии стрессогенов другого типа: мерцания, движения и повышенного темпа выполнения работ.

INFORMATION SUPPORT RESEARCH OF PERSON ATTENTION PARAMETERS

R.Y. YAKOVENKO, V.S. OSIPOVICH, N.V. SCHERBINA, K.D. YASHIN

Abstract

There are given the research results of attention parameters under the influence stress sound using modified proofreading test «Landolt Rings». There are accomplished noise influence evaluation of subject error amount, nature and degree.

Keywords: stress factor, proofreading test "Landolt Rings", attention parameters, stress resistance.

Список литературы

1. *Куприянов Р.В., Кузьмина Ю.М.* Психодиагностика стресса: практикум. Казань, 2012.
2. *Шабалин А.В.* // Бюллетень СО РАМН. 2003. № 4 (110). С. 25–29.
3. *Покровский В.М.* Сердечно-дыхательный синхронизм в оценке регуляторно-адаптивного статуса организма. Краснодар, 2010.
4. *Алешин С.* Информационный стресс: практические рекомендации. М., 2000.

УДК 616-009.5; 004.934

ЧАСТОТНО-ВРЕМЕННОЙ АНАЛИЗ РЕЧЕВЫХ СИГНАЛОВ В ДИАГНОСТИКЕ БУЛЬБАРНЫХ НАРУШЕНИЙ

М.М. МЕЖЕННАЯ, Ю.Н. РУШКЕВИЧ*, А.А. БОРИСКЕВИЧ

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь*

**РНПЦ Неврологии и нейрохирургии МЗ РБ
Ф. Скорины 24, Минск, 220114, Беларусь*

Поступила в редакцию 15 ноября 2016

Предложен метод качественной и количественной диагностики бульбарных нарушений на основе цифровой обработки речевых сигналов. Для реализации данного метода авторами разработано программное обеспечение с графическим интерфейсом, которое позволяет повысить точность и скорость постановки диагноза.

Ключевые слова: речевой сигнал, частотно-временной анализ, частота основного тона, бульбарный синдром.

Введение

Бульбарные нарушения (БН) представляют собой симптомокомплекс, который включает изменение звучности голоса (дисфонию), затруднения при глотании – дисфагию и замедленность речи, нарушение артикуляции – дизартрию, т.е. симптомы, связанные с вовлечением мускулатуры языка, глотки, гортани и мягкого неба. Причинами БН является непосредственное поражение ядер языкоглоточного, блуждающего и подъязычного черепных нервов, расположенных в каудальных отделах ствола головного мозга (бульбарный синдром), а также поражение вышеописанных мышц, нервно-мышечного аппарата, патологические процессы в области ствола мозга и задней черепной ямки [1–3].

К ранним проявлениям бульбарного синдрома относится дисфония: голос больных становится слабым, глухим, истощающимся вплоть до полной афонии. Возникает гнусавость. Звуки при этом произносятся невнятно, «смазанно». Гласные звуки становятся трудноотличимыми друг от друга, согласные звуки, разные по способу образования (твердые, мягкие, смычные, щелевые) и месту артикуляции (губные, переднеязычные, заднеязычные), произносятся однотипно с неопределенным местом артикуляции. Речь оказывается резко замедленной и утомляет больных. Развивающаяся дисфагия из-за невозможности сглатывать слюну и приводит к слюнотечению. При бульбарном параличе наступает атрофия мышц языка и выпадают глоточный и небный рефлекс. У тяжелобольных с бульбарным синдромом, как правило, развиваются расстройства ритма дыхания и сердечной деятельности, что нередко приводит к смерти [2, 3].

Эффективная дифференциальная диагностика позволяет своевременно оказать медицинскую помощь пациентам с БН. К достоверным методам диагностики бульбарного синдрома относятся данные электромиографии и прямого осмотра ротоглотки. Однако в настоящее время имеются сложности постановки диагноза на ранней стадии, сопровождающейся, как отмечалось выше, нарушениями речевой функции. Для проведения своевременной и объективной диагностики бульбарных нарушений предлагается использовать методы частотно-временного анализа нестационарного по своей природе речевого сигнала.

Методика регистрации и обработки речевых сигналов

Диагностические исследования бульбарных нарушений были проведены на базе РНПЦ неврологии и нейрохирургии МЗ РБ. В группе пациентов с боковым амиотрофическим склерозом с бульбарным синдромом, а также в контрольной группе здоровых лиц были зарегистрированы тестовые речевые сигналы. Тест представлял собой счет от одного до десяти. Последующая обработка речевых сигналов выполнялась в среде MatLab с помощью специально разработанного авторами статьи программного обеспечения с графическим интерфейсом. Обработка включала: автоматическое выделение в зарегистрированном сигнале речевых фрагментов; подсчет количества выделенных речевых фрагментов; построение спектрограммы зарегистрированного сигнала; построение кепстрограмм для выделенных речевых фрагментов; определение с помощью кепстральной функции частоты основного тона для каждого речевого фрагмента; расчет средней величины частоты основного тона; расчет коэффициента вариации частоты основного тона. Далее приведена подробная методика реализации вышеописанных этапов.

Исходный зарегистрированный сигнал характеризовался частотой дискретизации 44,1 кГц, разрядностью 16 бит. Предварительно производилось усреднение зарегистрированного сигнала в окне без перекрытия для снижения исходной частоты

дискретизации: $A[i] = \frac{1}{M} \sum_{j=0}^{M-1} x[i \cdot M + j]$, где $A[i]$ – отсчеты, полученные из исходного

сигнала x посредством усреднения; $i = 0 \dots \frac{N}{M} - 1$ – номер окна; N – число отсчетов исходного

речевого сигнала x ; M – длина окна (число усредняемых точек); j – номер временного отсчета внутри окна.

В результате усреднения при $M = 5$ частота дискретизации была понижена до 8,82 кГц. Это позволило впоследствии увеличить скорость обработки данных без потери полезной информации в сигнале.

Речь человека содержит паузы между словами. Традиционно для решения задачи разделения речевого сигнала на голосовые и неголосовые участки исходный сигнал разделяется на фрагменты длиной 5–100 мс. Для классификации принадлежности фрагмента к сигналу или паузе рассчитывалась кратковременная энергия сигнала в данном фрагменте:

$$E_m = \sum_{j=0}^{L_{fr}-1} A[m \cdot L_{fr} + j]^2, \text{ где } L_{fr} - \text{длина фрагмента; } m = 0 \dots \frac{N}{M \cdot L_{fr}} - 1 - \text{количество}$$

фрагментов; j – номер временного отсчета усредненного сигнала внутри фрагмента. В качестве L_{fr} авторами выбраны 400 отсчетов, что соответствует временной реализации сигнала в 45,4 мс.

На основе экспериментальных исследований речевых сигналов в норме было сформировано условие, при выполнении которого принималось решение о принадлежности m -ого фрагмента к речи: $E_m \geq level_E \cdot \langle E \rangle$, где $\langle E \rangle$ – средняя кратковременная энергия всех фрагментов, $level_E$ – пороговый уровень кратковременной энергии. Авторами установлено, что при $level_E = 0.2$ происходит автоматическое выделение слов и/или отдельных фонем в сигнале. Далее выполнялось построение спектрограммы сигнала. Для этого речевой сигнал разделялся на временные отрезки, в пределах которых его можно считать стационарным (5–100 мс). Исходный сигнал A на выбранном отрезке умножался на оконную функцию w и подвергался быстрому преобразованию Фурье в соответствии с выражением:

$$STFT_A^w[f_k, \tau] = \sum_{i=0}^{L-1} A[i] \cdot w[i] \cdot e^{-\frac{j2\pi ki}{L}}, f_k = \frac{k \cdot f_d}{L}, k = 0 \dots (L-1)/2, \text{ где } L - \text{длина окна, } \tau -$$

величина перекрытия окон, f_d – частота дискретизации.

После данной операции путем возведения в квадрат амплитудной части оконного преобразования Фурье получали спектрограмму мощности для анализируемого окна:

$$\text{Спектрограмма } A[f, t] = |STFT_A^w[f_k, \tau]|^2. \text{ Далее производилось смещение окна на величину}$$

τ и процедура повторялась. Подобным образом анализировались все подинтервалы сигнала и строилась результирующая спектрограмма, представляющая собой двумерную матрицу, строки которой соответствуют временным отсчетам t от 0 секунд до окончания времени регистрации речевого сигнала, столбцы – частотам f от 0 до 4,41 кГц, а в ячейках рассчитана амплитуда сигнала [4]. В качестве основных параметров частотно-временной обработки выбраны следующие: окно Хэмминга, размер окна L в 512 отсчетов, частота дискретизации f_d в 8,82 кГц, перекрытие окон τ в 50 %. Указанные характеристики обеспечивают качественное частотно-временное представление речевого сигнала, высокое разрешение по частоте $\Delta f = 17,2$ Гц и по времени $\Delta t = 29,0$ мс.

Для определения частоты основного тона сигнала использовался метод определения кепстра, заключающийся в применении к модулю спектральной плотности исследуемого сигнала обратного преобразования Фурье. При этом в кепстрограмме локализованного отрезка звука появляется пик на расстоянии основного тона сигнала, что и является основополагающим для последующего вычисления частоты основного тона.

Частота основного тона вычислялась для каждого выделенного локализованного фрагмента сигнала. По итогам расчетов определяли среднее значение частоты основного тона, а также коэффициент вариации данного параметра – относительную меру разброса значений признака в статистической совокупности.

Результаты исследований

Результаты обработки тестовых речевых сигналов в норме, при бульбарном синдроме до лечения и после курса транскраниальной магнитной стимуляции и нейрометаболического лечения представлены на рис. 1, 2, 3 соответственно.

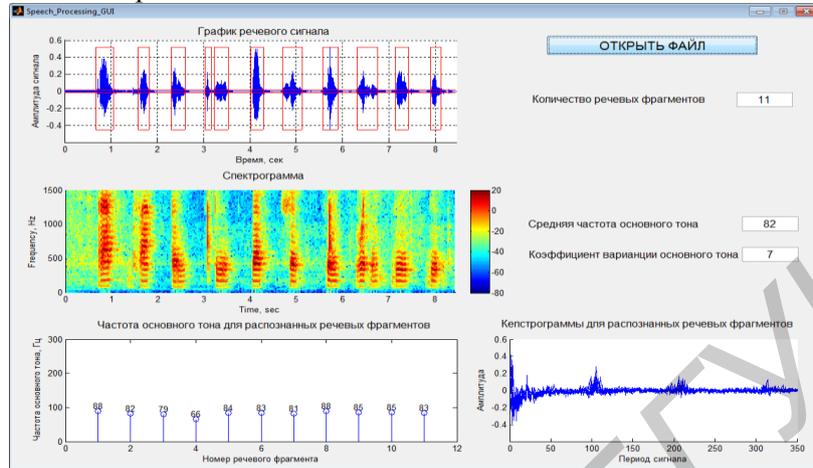


Рис. 1. Результаты обработки речевого сигнала в норме (испытуемый У)

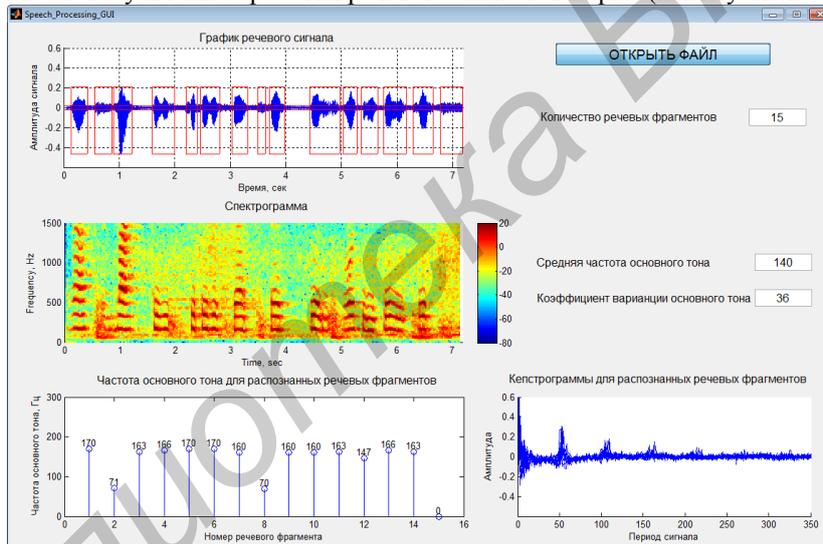


Рис. 2. Результаты обработки речевого сигнала до лечения бульбарного синдрома (пациент К)

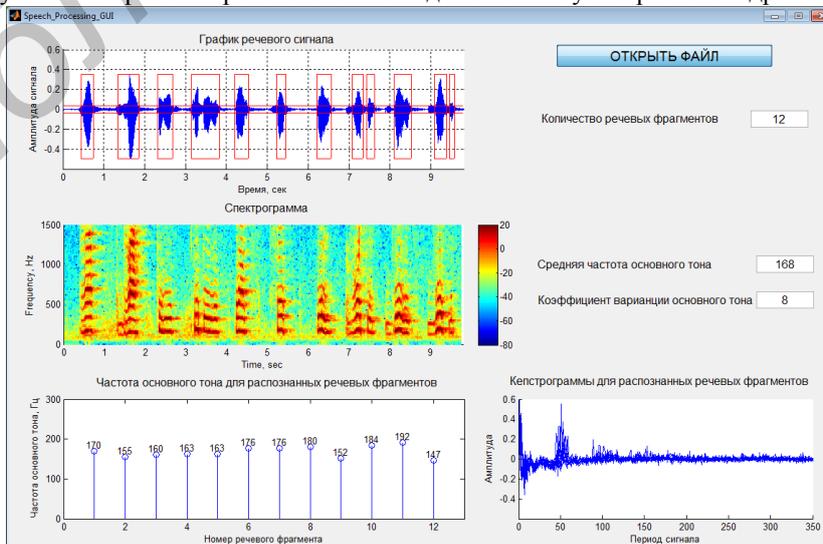


Рис. 3. Результаты обработки речевого сигнала после лечения бульбарного синдрома (пациент К)

Анализ полученных результатов выявил следующие закономерности. В группе здоровых лиц (рис. 1) количество распознанных речевых фрагментов составляет 10–12, что соответствует количеству произносимых слов (10) или фонем («че-тыре», «во-семь»). На спектрограмме отчетливо выделяются равностоящие друг от друга речевые фрагменты, частота основного тона, а также формантные частоты. Кепстрограммы для распознанных речевых фрагментов также имеют характерные пики в области основного тона и кратных ему формантных частот. Коэффициент вариации основного тона невысокий (равен 7 на рис. 1), что свидетельствует о постоянстве данного параметра во время речи.

В группе пациентов с бульбарным синдромом до лечения (рис. 2) количество распознанных речевых фрагментов, как правило, превышает количество произносимых слов (10). Это объясняется характерной для данной патологии невнятностью речи. Наблюдается снижение амплитуды сигнала (голос становится слабым, хриплым). Сигнал на спектрограмме не имеет четкой временной структуры в виде равноотстоящих речевых актов, как у здоровых лиц. Значение частоты основного тона для ряда речевых фрагментов не входит в стандартные диапазоны 70–450 Гц или не определяется вовсе. Показателен коэффициент вариации частоты основного тона (равен 36 на рис. 2), свидетельствующий о сильной степени рассеяния данного параметра относительно среднеарифметического значения.

В группе пациентов с бульбарным синдромом после лечения (рис. 3) количество распознанных речевых фрагментов в целом соответствовало количеству произносимых слов (10) или фонем («че-тыре», «во-семь»). Возросла амплитуда сигнала. На спектрограмме речевые фрагменты приобрели четкие очертания; выделяются паузы, как в группе здоровых лиц. Характерные для основного тона и формантных частот пики демонстрируют кепстрограммы. Вариабельность частоты основного тона вернулась к показателям в норме (равно 8 на рис. 3).

Заключение

Установлено, что речевые сигналы пациентов с бульбарным синдромом содержат число вокализованных фрагментов, превышающее количество произносимых слов (из-за невнятности речи); наблюдается снижение амплитуды сигнала (голос становится слабым, хриплым); сигнал на спектрограмме не имеет четкой временной структуры в виде равноотстоящих речевых актов, как у здоровых лиц; значение частоты основного тона для ряда речевых фрагментов не входит в стандартные диапазоны 70–450 Гц или не определяется вовсе; наблюдается высокая степень рассеяния значений частоты основного тона.

TIME-FREQUENCY ANALYSIS OF SPEECH SIGNALS IN THE DIAGNOSIS OF BULBAR PALSY

M.M. MEZHENNAYA, Yu.N. RUSHKEVICH, A.A. BORISKEVICH

Abstract

The method of qualitative and quantitative differential diagnosis of bulbar palsy has been offered on the basis of digital processing of speech signals. The software with the graphic user interface has been developed by authors for implementation of this method which allows to increase the accuracy and speed of diagnosis.

Keywords: speech signal, time-frequency analysis, main tone frequency, bulbar syndrome, bulbar palsy.

Список литературы

1. Завалишин И.А. Боковой амиотрофический склероз. М., 2009.
2. Andersen P. // Eur J Neurol. 2012. № 19(3). P. 360–375.
3. Miller R. // Neurology. 2009. № 73 (15). P. 1218–1226.
4. Райгайян Р.М. Анализ биомедицинских сигналов. Практический подход. М., 2007.