

УДК 612.789.4

*А. К. Алимуратов¹, А. Ю. Тычков¹, А. П. Зарецкий², А. П. Кулешов³*¹ФГБОУ ВО «Пензенский государственный университет», Научно-исследовательский институт фундаментальных и прикладных исследований, 440026, г. Пенза²Московский физико-технический институт (государственный университет)³ФГБУ «НМИЦ ТИО им. ак. В. И. Шумакова» Минздрава России

Способ определения кепстральных маркеров речевых сигналов при психогенных расстройствах

На сегодняшний день для обнаружения психогенных расстройств применяются различные дистанционные экспериментально-статистические методы, наиболее адаптивными из которых являются способы на основе анализа речевых сигналов. Низкая точность обнаружения является одной из основных проблем практической реализации систем дистанционного мониторинга психогенных расстройств. Основная причина низкой точности и больших погрешностей связана с использованием неэффективных и неадаптивных методов обработки нестационарных речевых сигналов. В данной статье предлагается автоматизированный способ определения кепстральных маркеров речевых сигналов при психогенных расстройствах на основе метода улучшенной полной множественной декомпозиции на эмпирические моды с адаптивным шумом (ПМДЭМАШ). Суть способа заключается в разложении речевого сигнала с помощью улучшенной ПМДЭМАШ на частотные составляющие с последующим формированием набора информативных компонент (концентрации информации о психогенных расстройствах) и определением их кепстральных маркеров. Представлена блок-схема разработанного способа и подробное математическое описание. Проведено исследование с использованием сформированной верифицированной базы сигналов здоровых пациентов и пациентов с психогенными расстройствами мужского и женского пола, в возрасте от 18 до 60 лет. В соответствии с результатами исследования, следует, что психогенные расстройства в большей степени влияют на вокализованные характеристики речевого тракта и достаточно полно отображаются в кепстральных маркерах. Предложенный автоматизированный способ может быть использован в системах дистанционного мониторинга психогенных расстройств и внедрен в клиническую практику врача-психиатра для ускорения процесса лечения.

Ключевые слова: речевой сигнал, психогенные расстройства, кепстральные характеристики, мел-частотные кепстральные коэффициенты (МЧКК), улучшенная полная множественная декомпозиция на эмпирические моды с адаптивным шумом.

*А. К. Alimuradov¹, A. Yu. Tychkov¹, A. P. Zaretskiy², A. P. Kuleshov³*¹Penza state University, Research Institute for Basic and Applied Studies, 440026, Penza²Moscow Institute of Physics and Technology (State University)³Federal State Budgetary Institute «Academician V.I.Shumakov Federal Research Center of Transplantology and Artificial Organs», Ministry of Health of the Russian Federation

Method for determining cepstral markers of speech signals under psychogenic disorders

At present, various remote experimental and statistical methods are used for detection of psychogenic disorders, the most adaptive of which are methods based on the analysis of speech signals. Low detection accuracy is one of the main problems for practical implementation of remote monitoring systems of psychogenic disorders. The main reason for the low accuracy and large errors is associated with the use of inefficient and nonadaptive

methods for processing of nonstationary speech signals. This article proposes an automated method for detection of cepstral markers of speech signals under psychogenic disorders based on the method of improved complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN). The method consists in decomposition of a speech signal with the help of the improved CEEMDAN into frequency components with the subsequent formation of a set of informative components (concentration of information on psychogenic disorders), and the determination of their cepstral markers. A block diagram for the developed method and a detailed mathematical description are presented. The research is conducted using the generated verified signal base of healthy male and female patients, and patients with psychogenic disorders, aged 18 to 60 years. It follows in accordance with the results of the study that the psychogenic disorders affect the vocalized characteristics of the vocal tract and are sufficiently fully displayed in cepstral markers. The proposed automated method can be used in remote monitoring systems of psychogenic disorders, and introduced in clinical practice of a psychiatrist to accelerate the treatment process.

Key words: speech signal, psychogenic disorders, cepstral characteristics, mel-frequency cepstral coefficients (MFCC), improved complete ensemble empirical mode decomposition with adaptive noise.

1. Введение

Психогенное расстройство является наиболее распространенным аффективным расстройством, риск возникновения которого в течение жизни наблюдается у 10–20 % женщин и 5–12 % мужчин [1]. За последние 20 лет численность людей, страдающих психогенными расстройствами, стремительно возрастает. В настоящее время диагностика психогенных расстройств осуществляется посредством наблюдения врачей – профессионалами, имеющими значительный клинический опыт. Однако из-за растущего количества пациентов, возникает вероятность ошибок в клинической практике врачей (как человеческого фактора), что может привести к назначению некорректного курса лечения. По этой причине возникает необходимость разработки новых способов обнаружения маркеров психогенных расстройств [2, 3]. В настоящее время для обнаружения психических расстройств применяются различные дистанционные экспериментально-статистические методики обработки сигналов по доступным каналам регистрации реакций организма человека [4–6]. Наиболее адаптивным, работающим в режиме реального времени и свободной активности является способ на основе анализа речевых сигналов [7–9]. Вид и степень выраженности психических расстройств кодируются в определенные информативные параметры речевых сигналов – сигнальные маркеры [10]. Характеристики речи, способные служить маркерами психогенных состояний, можно разделить на три основные группы: спектрально-временные, амплитудно-частотные и кепстральные. В отличие от первых двух групп кепстральные маркеры характеризуют речевой сигнал с точки зрения частотной характеристики речевого тракта, отбрасывая при этом характеристики сигнала возбуждения. Кепстральные маркеры отображают информацию в частотной области речевого тракта более подробно и компактно, обеспечивая лучшее разделение сегментов речевого потока. Известно, что точность определения кепстральных маркеров зависит от корректной обработки речевых сигналов, которая определяется правильностью измерения его амплитудных, временных, частотных и энергетических характеристик. Основная причина низкой точности и больших погрешностей при измерениях связана с использованием неэффективных и неадаптивных методов обработки нестационарных речевых сигналов.

2. Материалы и методы

2.1. Мел-частотные кепстральные коэффициенты

В области обработки речевых сигналов кепстральный анализ получил широкую практическую популярность, объясняемую достоинством сжатия информации о сигнале при

переходе в частотную область обработки [11]. Кепстральный анализ основан на выделении кепстральных коэффициентов на мел-шкале, называемых мел-частотными кепстральными коэффициентами (МЧКК). МЧКК включают в себя два основных понятия: кепстр и мел-шкала. Кепстр – это дискретно-косинусное преобразование амплитудного спектра сигнала в логарифмическом масштабе. Кепстр сигнала определяется по формуле (2.1):

$$c(n) = DCT[\log(X(|x(n)|^2))], \quad (2.1)$$

где DCT – дискретно-косинусное преобразование, X – спектральное представление сигнала $x(n)$, n – дискретный отсчет времени ($0 < n \leq N$, N – количество дискретных отсчетов в сигнале). Метод получения МЧКК основан на модели функционирования органов слуха человека и использует частотную шкалу в мелах, которая моделирует частотную чувствительность человеческого уха [11].

2.2. Декомпозиция на эмпирические моды

Исследования методов обработки речевых сигналов выявили перспективность использования адаптивной технологии анализа нестационарных сигналов – декомпозиции на эмпирические моды (ДЭМ) [12]. ДЭМ – это адаптивный метод анализа нестационарных сигналов, возникающих в нелинейных системах. ДЭМ обеспечивает локальное разложение сигнала на быстрые и медленные колебательные функции. В результате разложения исходный сигнал может быть представлен в виде суммы амплитудных и частотных модулированных функций, называемых эмпирическими модами (ЭМ). Аналитическое выражение ДЭМ выглядит следующим образом (2.2):

$$x(n) = \sum_{i=1}^I IMF_i + r_i(n), \quad (2.2)$$

где $x(n)$ – исходный сигнал, $IMF_i(n)$ – ЭМ, $r_i(n)$ – конечный остаток, $i = 1, 2, \dots, I$ – номер ЭМ. В настоящее время для адаптивного разложения речевого сигнала перспективным является применение усовершенствованной полной множественной ДЭМ с адаптивным шумом [13]. Основная идея метода заключается в добавлении к исходному сигналу контролируемого шума для создания новых экстремумов.

В методе ПМДЭМАШ решаются все основные недостатки существующих разновидностей декомпозиции:

- явление смешивания ЭМ;
- наличие остаточного шума в ЭМ;
- наличие «паразитных» ЭМ на ранних этапах декомпозиции.

Алгоритм и математическое описание метода улучшенной ПМДЭМАШ:

Шаг 1. С помощью аппарата ДЭМ, выражая из формулы $\langle E_1(x_j(n)) \rangle = \langle x_j(n) \rangle - \langle M(x_j(n)) \rangle$ локальные средние значения шумовых копий исходного сигнала ($x_j(n) = x(n) + \beta_0 E_1(\omega_j(n))$), определяется первый остаток (2.3):

$$r_1(n) = \langle M(x_j(n)) \rangle, \quad (2.3)$$

где $E_i(\bullet)$ – аппарат извлечения ЭМ методом ДЭМ (i – номер моды), $x_j(n) = x(n) + w_j(n)$ – шумовые копии исходного сигнала ($x(n)$ – исходный речевой сигнал, $w_j(n)$ – реализации белого шума с нулевой средней единичной дисперсией), $M(\bullet)$ – аппарат, создающий локальное среднее значение применяемого сигнала, $\beta_i = \varepsilon_i \text{std}(r_i)$ – коэффициент, допускающий выбор различных значений отношения сигнал/шум.

Шаг 2. На первом этапе для $i = 1$ вычисляется первая мода: $IMF_1(n) = x(n) - r_1(n)$.

Шаг 3. Вычисляется второй остаток как усредненное локальное среднее значение шумовых копий первого остатка $r_1(n) + \beta_1 E_2(\omega_j(n))$ и определяется вторая мода:

$$\widehat{IMF}_2(n) = r_1(n) - r_2(n) = r_1(n) - \langle M(r_1(n) + \beta_1 E_2(\omega_j(n))) \rangle.$$

Шаг 4. На последующих этапах для $i = 3, \dots, I$ вычисляется i -й остаток:

$$r_i(n) = \langle M(r_{i-1}(n) + \beta_{i-1} E_i(\omega_j(n))) \rangle.$$

Шаг 5. Вычисляется i -я мода:

$$\widehat{IMF}_i(n) = r_{i-1}(n) - r_i(n).$$

Шаг 6. Переход к шагу 4 для следующего значения i . Константы β_i выбираются так, чтобы получить желаемое отношение сигнал/шум между добавленным шумом и остатком, к которому добавляется шум.

3. Описание способа

На рис. 1 представлена упрощенная блок-схема автоматизированного способа определения кепстральных маркеров речевых сигналов для обнаружения пограничных психических расстройств. Суть предложенного способа заключается в разложении с помощью усовершенствованной ПМДЭМАШ речевого сигнала на частотные составляющие с формированием набора информативных компонент (концентрации информации о психогенных расстройствах) и определением МЧКК. Рассмотрим каждый этап разработанного способа подробнее.

Блок 1. Ввод осуществляется со следующими параметрами: частота дискретизации 8000 Гц, разрядность квантования 16 бит.

Блок 2. В рамках предварительной обработки осуществляется: удаление постоянной составляющей, фильтрация речевого сигнала с помощью ФВЧ Чебышева четвертого порядка частот ниже 130 Гц и коррекция естественных искажений спектра (минус 6 дБ на октаву), возникающих в речевом аппарате человека при произнесении речи.

Блок 3. Сегментация речевого сигнала на информативные участки представляет собой обнаружение границ вокализованных и невокализованных участков в общем речевом потоке. Сегментация осуществлялась на основе вычисления следующих параметров в скользящем окне [14]: скорость пересечения сигнала через нулевое значение (Zero-Crossing Rate, ZCR), автокорреляционная функция (Autocorrelation Function, ACR), энергия/мощность (PWR), линейно-частотные кепстральные коэффициенты (Linear-Frequency Cepstral Coefficients, LFCC).

Вокализованные участки (сформированные с участием голосовых связок – гласные звуки и звуки сонорных согласных) определяются на основе следующего решающего правила, с учетом физиологических аспектов формирования речевых сигналов: сегменты длительностью меньше, чем 30 мс классифицируются как участки, не содержащие вокализации (сформированные без участия голосовых связок), но сегменты, не содержащие вокализации длительностью меньше, чем 20 мс, классифицируются как вокализованные, так как такой кратковременный переход между вокализованными и невокализованными сегментами невозможен. Вокализованные участки идентифицируются, как сегменты с самыми высокими средними значениями PWR и ACR и самым низким значением ZCR.

Невокализованные участки, представленные невокализованными согласными, отделяются от участков паузы (тишины) из общего потока участков, сформированных без участия голосовых связок, на основе следующего решающего правила: сегменты длительностью больше, чем 300 мс классифицируются как участки невокализованных согласных, но сегменты невокализованных согласных длительностью меньше, чем 5 мс и сегменты, находящиеся на расстоянии от вокализованных участков более чем на 30 мс, не используются. Для идентификации невокализованных участков используются первые пять коэффициентов LFCC. Невокализованные участки идентифицировались как сегменты с самым высоким средним значением первого коэффициента LFCC, что связано с громкостью произношения. Участки паузы (тишины), включая время, необходимое для дыхания, отделяются от невокализованных участков (сформированные без участия голосовых связок) на основе следующего решающего правила: сегменты длительностью больше, чем 30 мс, обладающие минимальным средним значением первого коэффициента LFCC, классифицируются

как участки паузы (тишины). На рис. 2 представлена графическая иллюстрация работы этапа сегментации.

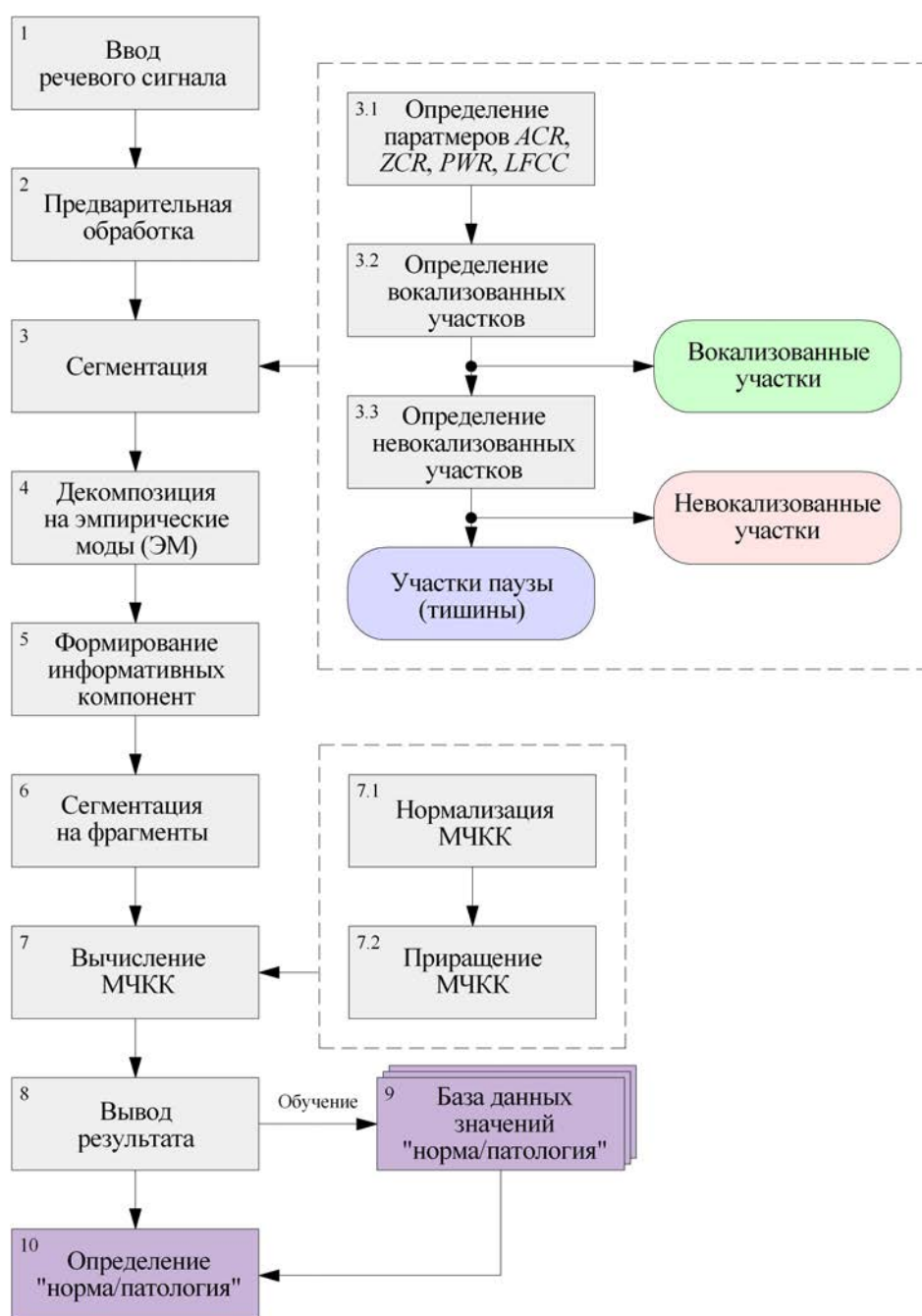


Рис. 1. Блок-схема способа

Блок 4. В представленном способе разложение речевых сигналов на частотные составляющие для последующего формирования информативных компонент осуществляется на основе улучшенной ПМДЭМАШ. Использование улучшенной ПМДЭМАШ обеспечит:

- адаптивное разложение, так как базисные функции, используемые при декомпозиции, извлекаются непосредственно из исходного речевого сигнала и позволяют учитывать только ему свойственные особенности (скрытые модуляции, области концентрации энергии и т.п.);
- минимальный уровень остаточного шума;
- отсутствие паразитных ЭМ, возникающих на ранних этапах декомпозиции вследствие перекрытия масштабно-энергетических пространств мод.

На рис. 3 и 4 представлены результаты разложения вокализованного и невокализованного участков речевого сигнала.

Блок 5. Основным понятием при формировании информативных компонент является концентрация информации о психогенных расстройствах в отдельные компоненты.

Для абсолютно произвольного сигнала все ЭМ можно разбить на две категории [10]: информативные ЭМ с шумовыми и сигнальными составляющими; неинформативные ЭМ с трендовыми составляющими.

В основе формирования информативных компонент заложено предположение, что информативные ЭМ имеют большую энергию, чем неинформативные. Амплитудное распределение ЭМ хорошо описывается с помощью функции кратковременной энергии. В разработанном способе для сжатия амплитуды сигнала в большом динамическом диапазоне применяется логарифмирование энергии, максимально приближая работу способа к работе слухового аппарата человека. На рис. 5 представлены кривые зависимости логарифмов энергии от номера ЭМ для вокализованных и невокализованных участков речевого сигнала.

Формирование информативных компонент заключается в вычитании из исходного сигнала информативных шумовых и неинформативных ЭМ. Информативными шумовыми обычно являются первые две или три ЭМ, в зависимости от интенсивности присутствующего в сигнале шума. Неинформативными являются последние три или четыре ЭМ, в зависимости от общего количества мод (число ЭМ примерно равно двоичному логарифму от числа дискретных отсчетов в сигнале). Как видно из рисунков 3–5, информативными модами (выделенными областью зеленого цвета) для вокализованного участка являются ЭМ2–ЭМ5, а для невокализованного участка ЭМ1–ЭМ5. Во втором случае учитывается, что невокализованные согласные звуки состоят из шумовых компонент, например шумовые согласные: п, б, ф, в, т, д, с, з, ц, ч, ш, ж, к, г, х.

Целью формирования информативных компонент является сбор информации, отражающей нарушения работы органов речевого аппарата вследствие психогенного расстройства. Формирование набора информативных компонент осуществляется по формуле (3.1):

$$x_{ab,i}(n) = x(n) - (a * \sum_{i=0}^2 IMF_{i+1}(n) + b * \sum_{i=0}^2 IMF_{i-1}(n)), \quad (3.1)$$

где $x_{ab,i}(n)$ – информативная компонента, a, b – коэффициенты, определяющие участие ЭМ в формировании информативных компонент.

Блок 6. Сегментация на фрагменты ЭМ – это линейное деление на составляющие отрезки. Разработанный способ основан на предположении о том, что свойства речевого сигнала с течением времени изменяются медленно. Это предположение приводит к кратковременному анализу, в котором фрагменты ЭМ выделяются и обрабатываются так, как если бы они были короткими участками с отличающимися свойствами.

Блок 7. Как отмечалось ранее, в качестве кепстальных характеристик в способе вычисляются МЧКК. Основными этапами вычисления МЧКК являются [15]:

- перевод сигнала из шкалы герц в шкалу мелов;
- спектральное преобразование и определение периодограммы сигнала;
- фильтрация периодограммы сигнала набором мел-фильтров;
- логарифмирование энергии сигнала в каждом мел-фильтре;
- дискретно-косинусного преобразования логарифма энергии.

Итогом вычисления являются МЧКК ($MFCC(c)$, где $c = 1, 2, \dots, C$ – номер МЧКК, C – желаемое количество коэффициентов).

В представленном способе используются 24 МЧКК, так как чем выше индекс коэффициента, тем быстрее изменяется энергия в наборе фильтров [15]. Также в результате экспериментальных исследований выяснилось, что первый коэффициент в основном несет информацию об интенсивности речевых сигналов.

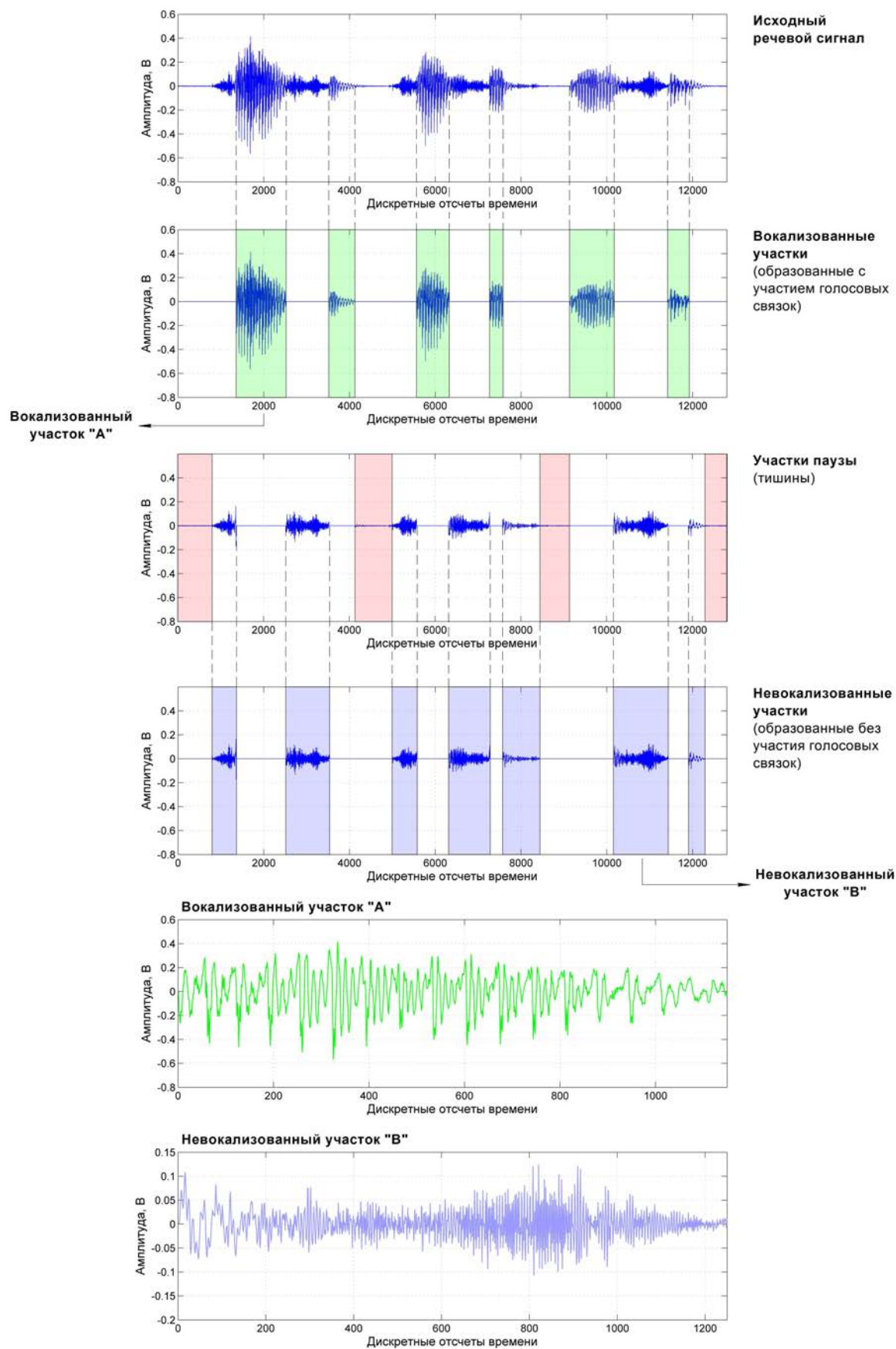


Рис. 2. Работа способа на этапе сегментации

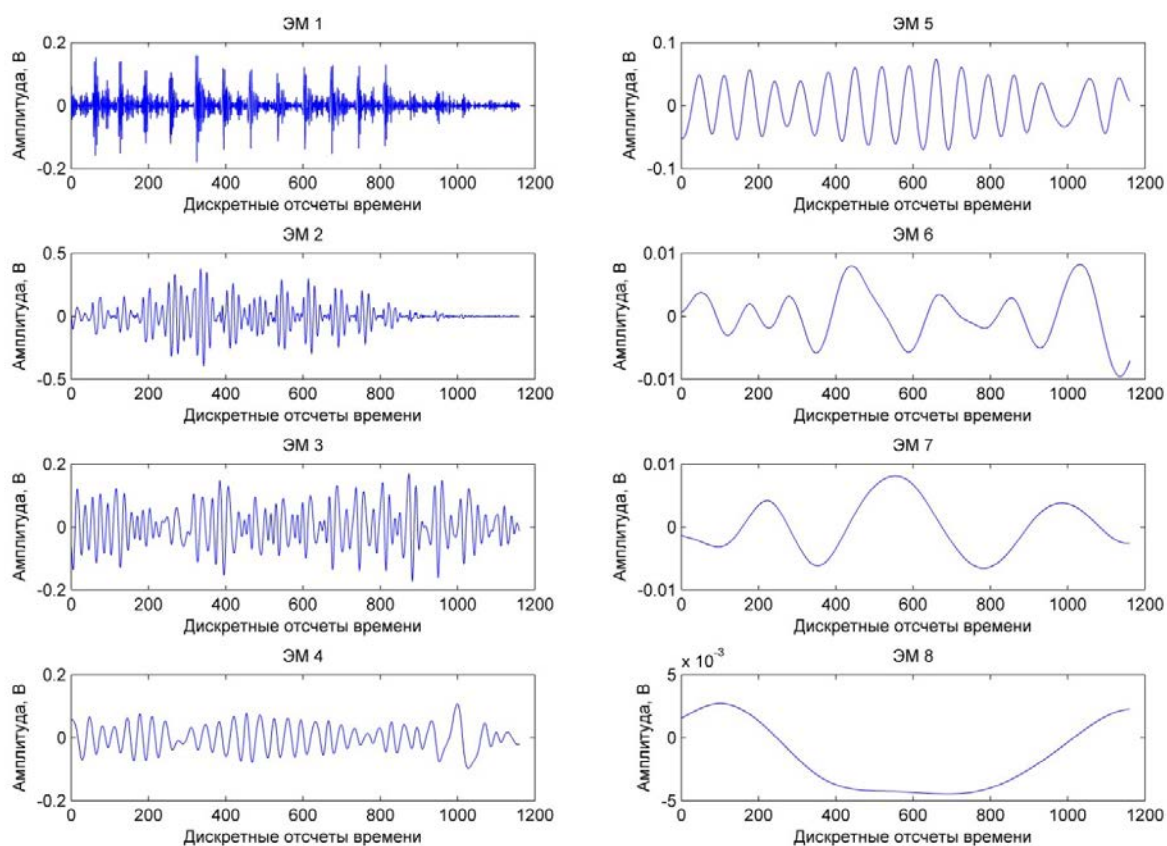


Рис. 3. Результат разложения вокализованного участка

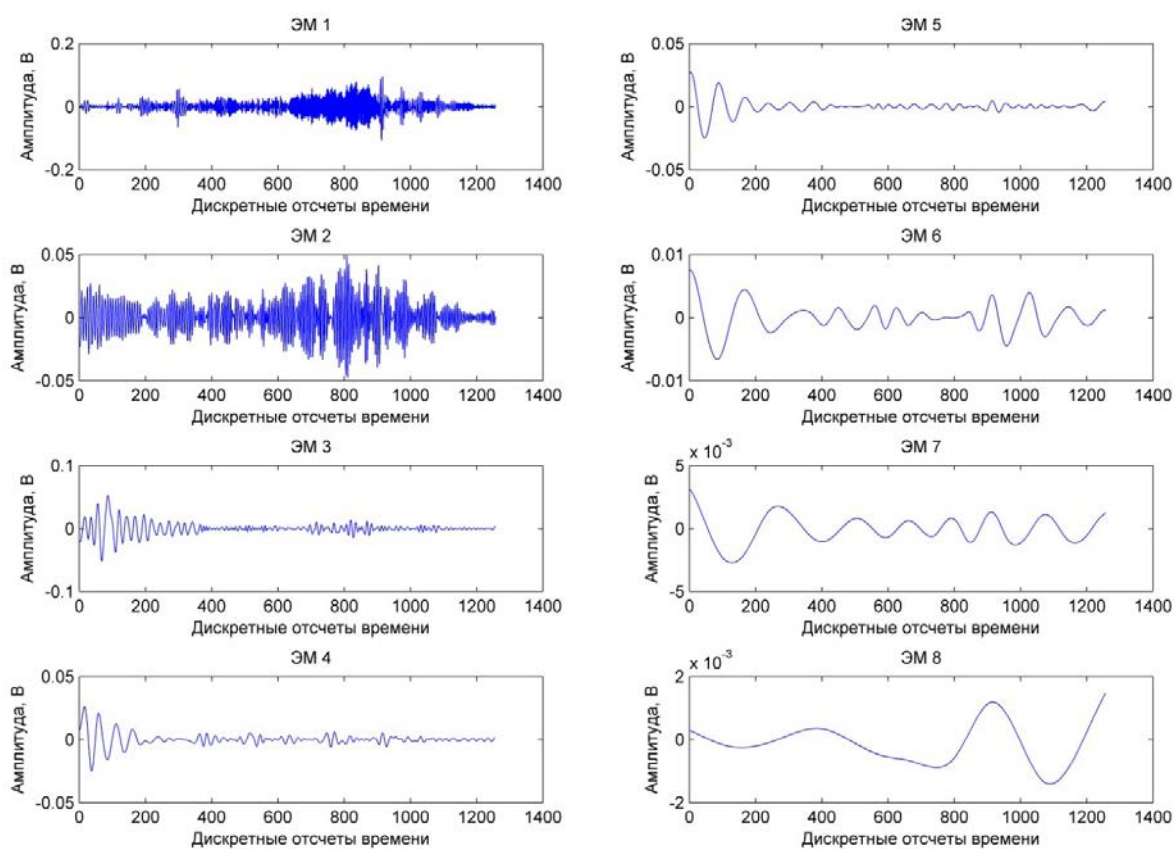


Рис. 4. Результат разложения невокализованного участка

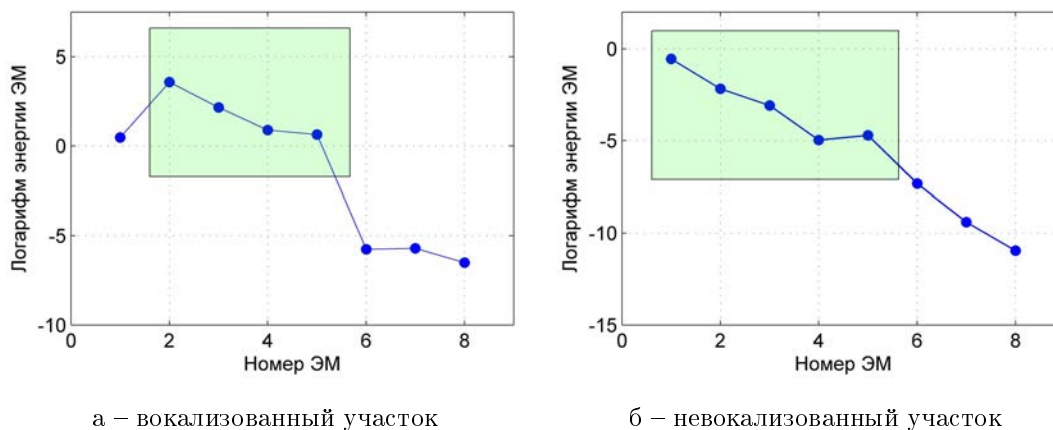


Рис. 5. Логарифм энергии ЭМ

Операция нормализации используется для придания равнозначности каждому МЧКК во фрагменте. Как известно, высокие частоты менее восприимчивы, и МЧКК на этих частотах менее информативны по сравнению с МЧКК на низких частотах. Нормализация МЧКК – это умножение каждого коэффициента на число, которое увеличивается с номером коэффициента. Таким образом, первые коэффициенты по уровню уменьшаются, а последние коэффициенты увеличиваются. Для этой операции используется следующая формула (3.2):

$$MFCC_N(c) = MFCC(c) * (1 + \frac{L}{2} \sin(\frac{\pi c}{2})), \quad (3.2)$$

где L – величина, подбираемая эмпирически и равна 22 [15]. Вычисление первого и второго приращений значений МЧКК позволяет получить динамическую информацию о коэффициентах. Вектор коэффициентов описывает фиксированную спектральную огибающую одного фрагмента, но, очевидно, что речевые сигналы несут информацию и о динамике в виде незначительного изменения коэффициентов с течением времени [15]:

$$MFCC_D(c) = \frac{\sum_{d=1}^D d(MFCC(c+d)) - MFCC(c-d)}{2 \sum_{d=1}^D d^2}, \quad (3.3)$$

$$MFCC_DD(c) = \frac{\sum_{d=1}^D d(MFCC_D(c+d)) - MFCC_D(c-d)}{2 \sum_{d=1}^D d^2}, \quad (3.4)$$

где $MFCC_D(c)$, $MFCC_{DD}(c)$ – первое и второе приращения МЧКК, $MFCC(c)$ – статические МЧКК, D – типовое значение приращения, равное 2.

Блок 8. На данном этапе работы способа осуществляется формирование вектора полученных МЧКК (первичных, нормализованных и после приращения), удобного для дальнейшего определения «норма/патология».

4. Исследование способа

4.1. Описание базы данных речевых сигналов

Для проведения исследований разработанного способа сформированы группа пациентов и верифицированная база сигналов при поддержке Областной клинической больницы им. К.Р. Евграфова (г. Пенза, Российская Федерация) и Пензенского государственного университета. Группа пациентов сформирована в соответствии с клинической картиной расстройства следующих диагностических рубрик международной классификации болезней (МКБ-10): F48.0, F45.3, F43.2, F41.2.

В группу пациентов с психогенными расстройствами отобрано 100 человек мужского и женского пола в возрасте от 18 до 60 лет, поступившие с явно выраженной симптоматикой. В том же количестве, 100 человек, сформирована база контрольной группы пациентов без признаков пограничных психических расстройств (условно здоровые). Средний возраст в экспериментальной группе пациентов с пограничными психическими расстройствами составил 40,2 года, в контрольной группе сравнения – 35,4 года.

В обеих группах преобладали женщины (75%), возраст преимущественно для женщин от 40 до 59 лет, для мужчин от 50 до 59 лет. Большинство пациентов были работающими (90,8%), среди которых выделялись по численности служащие предприятий и организаций (65,0%). Меньшие доли приходились на рабочих (14,2%), творческих работников (12,5%) и единицы – на учащихся, студентов и не работающих. Большинство пациентов имели высшее или неоконченное высшее образование (69,2%).

4.2. Результаты исследования

Для оценки эффективности разработанного способа использовался параметр – ошибки первого и второго рода. В рамках исследования главной задачей являлось определение маркеров психогенных расстройств, поэтому ошибкой первого рода будет ложное присваивание статуса «норма» речевому сигналу, произнесенному человеком с психогенным расстройством, а ошибкой второго рода – ложное присваивание статуса «патология» речевому сигналу, произнесенному здоровым человеком.

Исследование кепстральных маркеров речевых сигналов, отражающих нарушения работы органов речевого аппарата вследствие психогенного расстройства, проводилось для:

- сигналов, состоящих только из вокализованных участков (сформированных с участием голосовых связок);
- сигналов, состоящих из невокализованных участков (сформированных без участия голосовых связок);
- сигналов, состоящих из вокализованных и невокализованных участков.

Для каждой группы сигналов сформированы наборы из трех информативных компонент, полученных вследствие вычитания из исходного сигнала информативных шумовых и неинформативных ЭМ.

В таблице 1 представлены результаты определения психических расстройств на основе вычисления кепстральных маркеров для трех сигналов.

Т а б л и ц а 1

Результаты определения пограничных психических расстройств

Прогнозируемый результат	Результат определения		Ошибки первого рода и второго рода, %	
	Патология	Норма		
Информативная компонента № 1 (вычитание ЭМ6, ЭМ7, ЭМ8)				
Патология	78 чел.	22 чел.	1-го	22
Норма	16 чел.	84 чел.	2-го	16
Информативная компонента № 2 (вычитание ЭМ1, ЭМ6, ЭМ7, ЭМ8)				
Патология	19 чел.	81 чел.	1-го	81
Норма	51 чел.	49 чел.	2-го	51
Информативная компонента № 3 (вычитание ЭМ1, ЭМ2, ЭМ6, ЭМ7, ЭМ8)				
Патология	83 чел.	17 чел.	1-го	17
Норма	15 чел.	85 чел.	2-го	15

В табл. 2 представлены результаты определения пограничных психических расстройств на основе вычисления кепстральных маркеров для информативных компонент сигнала, состоящего из вокализованных участков.

Т а б л и ц а 2

**Результаты определения пограничных психических расстройств
информативных компонент сигнала, состоящего из вокализованных участков**

Прогнозируемый результат	Результат определения		Ошибки первого рода и второго рода, %	
	Патология	Норма		
Информативная компонента № 1 (вычитание ЭМ6, ЭМ7, ЭМ8)				
Патология	81 чел.	19 чел.	1-го	19
Норма	11 чел.	89 чел.	2-го	11
Информативная компонента № 2 (вычитание ЭМ1, ЭМ6, ЭМ7, ЭМ8)				
Патология	86 чел.	14 чел.	1-го	14
Норма	10 чел.	90 чел.	2-го	10
Информативная компонента № 3 (вычитание ЭМ1, ЭМ2, ЭМ6, ЭМ7, ЭМ8)				
Патология	44 чел.	56 чел.	1-го	56
Норма	39 чел.	61 чел.	2-го	39

Т а б л и ц а 3

**Результаты определения пограничных психических расстройств для
информативных компонент сигнала, состоящего из невокализованных
участков**

Прогнозируемый результат	Результат определения		Ошибки первого рода и второго рода, %	
	Патология	Норма		
Информативная компонента № 1 (вычитание ЭМ6, ЭМ7, ЭМ8)				
Патология	21 чел.	79 чел.	1-го	79
Норма	56 чел.	44 чел.	2-го	56
Информативная компонента № 2 (вычитание ЭМ1, ЭМ6, ЭМ7, ЭМ8)				
Патология	16 чел.	84 чел.	1-го	84
Норма	71 чел.	29 чел.	2-го	71
Информативная компонента № 3 (вычитание ЭМ1, ЭМ2, ЭМ6, ЭМ7, ЭМ8)				
Патология	9 чел.	81 чел.	1-го	81
Норма	79 чел.	21 чел.	2-го	79

Т а б л и ц а 4

**Результаты определения пограничных психических расстройств для
информативных компонент сигнала, состоящего из вокализованных и
невокализованных участков**

Прогнозируемый результат	Результат определения		Ошибки первого рода и второго рода, %	
	Патология	Норма		
Информативная компонента № 1 (вычитание ЭМ6, ЭМ7, ЭМ8)				
Патология	86 чел.	14 чел.	1-го	14
Норма	13 чел.	87 чел.	2-го	13
Информативная компонента № 2 (вычитание ЭМ1, ЭМ6, ЭМ7, ЭМ8)				
Патология	89 чел.	11 чел.	1-го	11
Норма	12 чел.	88 чел.	2-го	12
Информативная компонента № 3 (вычитание ЭМ1, ЭМ2, ЭМ6, ЭМ7, ЭМ8)				
Патология	48 чел.	52 чел.	1-го	52
Норма	37 чел.	63 чел.	2-го	37

В табл. 3 представлены результаты определения пограничных психических расстройств на основе вычисления кепстральных маркеров для информативных компонент сигнала, состоящего из невокализованных участков.

В табл. 4 представлены результаты определения пограничных психических расстройств на основе вычисления кепстральных маркеров для информативных компонент сигнала, состоящего из вокализованных и невокализованных участков.

5. Обсуждение и выводы

В соответствии с полученными данными можно сделать вывод, что в сравнении между двумя сигналами, состоящими из вокализованных и невокализованных участков соответственно, первые отражают больше информации о нарушениях работы органов речевого аппарата и значит, психических расстройствах человека. Это объясняется тем, что психогенные расстройства в большей степени влияют на вокализованные характеристики речевого тракта и достаточно полно отображаются в кепстральных маркерах.

Данные результаты позволяют сделать вывод, что разработанный автоматизированный способ определения кепстральных маркеров речевых сигналов на основе метода улучшенной ПМДЭМАШ может быть успешно использован в системах дистанционного мониторинга психогенных расстройств с целью упрощения задачи диагностики и внедрения в клиническую практику врача-психиатра для ускорения процесса лечения. Авторы благодарят Российский научный фонд за финансовую поддержку проекта «Поиск скрытых паттернов пограничных психических расстройств и разработка системы экспресс оценки состояния психического здоровья человека», № 17-71-20029.

Литература

1. *Schwartz T.L., Petersen, T.J.* Depression: Treatment Strategies and Management (Medical Psychiatry Series). CRC Press. Boca Raton, 2009.
2. *Тычков А.Ю., Агейкин А.В., Алимуратов А.К., Чураков П.П., Тычкова А.Н.* Анализ и оценка сигнальных систем диагностики пограничных психических расстройств // Биотехносфера. 2017. № 1(49). С. 35–39.
3. *Tychkov A.Yu., Alimuradov A.K., Churakov P.P.* Adaptive signal processing method for speech organ diagnostics // Meas. Tech. 2016. 59(5), 485–490.
4. *Koelstra S., Pantic M., Patras I.* A dynamic texture-based approach to recognition of facial actions and their temporal models // IEEE TPAMI. 2010. V. 32. P. 1940–1954.
5. *Zaboleeva-Zotova A.V., Orlova Yu.A., Rozaliev V.L., Bobkov A.S.* A task of creating a system of automated recognition of emotions // International Conference on Open Semantic Technology for intelligent Systems (OSTIS). 2012. P. 347–350.
6. *Barabanshikov V.A., Zhegallo A.V.* Methods of eye tracking in psychology: educational program // Exp. Psych. 2014. V. 7(1). P. 132–137.
7. *Filatova N.N., Sidorov K.V.* The model for the interpretation of sign of emotions on natural speech // Izvestiya SFedU. Eng. Sci. 2012. V. 9(134). P. 39–45.
8. *Davydov A.G., Kiselev V.V., Kochetkov D.S., Tkachenya A.V.* A choice for the optimal set of informative signs for classification of the speaker's emotional state by voice // International Conference on Computational Linguistics and Intellectual Technologies (Dialogue-2012). 2011. V. 1(11). 122–128.
9. *Williamson J.R., Quatieri T.F., Helfer B.S., Ciccarelli G., Mehta D.D.* Vocal and facial biomarkers of depression based on motor incoordination and timing // 4th International Workshop on Audio/Visual Emotional Challenge (AVEC). 2014. P. 65–72.
10. *Darley F.L., Aronson A.E., Brown J.R.* Differential diagnostic patterns of dysarthria // J. Speech Lang. Hear. R. 1969. V. 12. P. 47–57.
11. *Huang X., Acero A., Hon H.-W.* Spoken Language Processing. Guide to Algorithms and System Development. Prentice Hall, Upper Saddle River. 2001.

12. *Huang N.E., Zheng Sh., Steven R.L.* The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis // Proc. R. Soc. Lond. 1998. A 454. P. 903–995.
13. *Colominasa M.A., Schlotthauera G., Torres M.E.* Improved complete ensemble EMD: a suitable tool for biomedical signal processing // Biomed. Signal Proces. 2014. V. 14. P. 19–29.
14. *Hlavnicka J., Cmejla R., Tykalova T., Sonka K., Ruzicka E., Rusz J.* Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder // Sci. Rep. 2017. V. 7(12). P. 13.
15. *Алимуратов А.К., Муртазов Ф.Ш.* Методы повышения эффективности распознавания речевых сигналов в системах голосового управления // Измерительная техника. 2015. № 10. С. 20–24.

References

1. *Schwartz T.L., Petersen, T.J.* Depression: Treatment Strategies and Management (Medical Psychiatry Series). CRC Press. Boca Raton, 2009.
2. *Tychkov A.Yu., Ageykin A.V., Alimuradov A.K., Churakov P.P., Tychkova A.N.* Analysis and evaluation of signaling systems for the diagnosis of borderline mental disorders. Biotechnosphere. 2017. N 1(49). P. 35–39.
3. *Tychkov A.Yu., Alimuradov A.K., Churakov P.P.* Adaptive signal processing method for speech organ diagnostics. Meas. Tech. 2016. 59(5), 485–490.
4. *Koelstra S., Pantic M., Patras I.* A dynamic texture-based approach to recognition of facial actions and their temporal models. IEEE TPAMI. 2010. V. 32. P. 1940–1954.
5. *Zaboleeva-Zotova A.V., Orlova Yu.A., Rozaliev V.L., Bobkov A.S.* A task of creating a system of automated recognition of emotions. International Conference on Open Semantic Technology for intelligent Systems (OSTIS). 2012. P. 347–350.
6. *Barabanschikov V.A., Zhegallo A.V.* Methods of eye tracking in psychology: educational program. Exp. Psych. 2014. V. 7(1). P. 132–137.
7. *Filatova N.N., Sidorov K.V.* The model for the interpretation of sign of emotions on natural speech. Izvestiya SFedU. Eng. Sci. 2012. V. 9(134). P. 39–45.
8. *Davydov A.G., Kiselev V.V., Kochetkov D.S., Tkachenya A.V.* A choice for the optimal set of informative signs for classification of the speaker's emotional state by voice. International Conference on Computational Linguistics and Intellectual Technologies (Dialogue-2012). 2011. V. 1(11). 122–128.
9. *Williamson J.R., Quatieri T.F., Helfer B.S., Ciccarelli G., Mehta D.D.* Vocal and facial biomarkers of depression based on motor incoordination and timing. 4th International Workshop on Audio/Visual Emotional Challenge (AVEC). 2014. P. 65–72.
10. *Darley F.L., Aronson A.E., Brown J.R.* Differential diagnostic patterns of dysarthria. J. Speech Lang. Hear. R. 1969. V. 12. P. 47–57.
11. *Huang X., Acero A., Hon H.-W.* Spoken Language Processing. Guide to Algorithms and System Development. Prentice Hall, Upper Saddle River. 2001.
12. *Huang N.E., Zheng Sh., Steven R.L.* The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. 1998. A 454. P. 903–995.
13. *Colominasa M.A., Schlotthauera G., Torres M.E.* Improved complete ensemble EMD: a suitable tool for biomedical signal processing. Biomed. Signal Proces. 2014. V. 14. P. 19–29.

14. *Hlavnicka J., Stejla R., Tykalova T., Sonka K., Ruzicka E., Ruz J.* Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Sci. Rep.* 2017. V. 7(12). P. 13.
15. *Alimuradov A.K., Murtazov F.Sh.* Methods to improve the efficiency of recognition of speech signals in voice control systems. *Measurement Techniques.* 2015. N 10. P. 20–24.

Поступила в редакцию 13.11.2017