

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

УДК 616.89

DOI 10.21685/2072-3059-2017-3-1

А. К. Алимуратов, А. Ю. Тычков, П. П. Чураков, А. В. Агейкин

РАЗРАБОТКА АЛГОРИТМА ОБРАБОТКИ РЕЧЕВЫХ СИГНАЛОВ ДЛЯ ОПРЕДЕЛЕНИЯ ИНФОРМАТИВНО ЗНАЧИМЫХ ПАРАМЕТРОВ ПОГРАНИЧНЫХ ПСИХИЧЕСКИХ РАССТРОЙСТВ¹

Аннотация.

Актуальность и цели. Объектом исследования явились больные с пограничными психическими расстройствами Областной психиатрической больницы им. К. Р. Евграфова, у которых имеется высокий процент ложноотрицательных результатов диагностики заболеваний. Предметом исследования являются алгоритмы обработки речевых сигналов для диагностики пограничных психических расстройств. Цель – разработка алгоритма измерения частоты основного тона для систем обнаружения паттернов пограничных психических расстройств.

Материалы и методы. В качестве материалов использовались информативные параметры речевых сигналов – паттерны. Для эффективной обработки речевых сигналов использовался метод декомпозиции на эмпирические моды и его модификация – полная множественная декомпозиция на эмпирические моды с адаптивным шумом. Результаты исследований оценивались в сравнении с известными алгоритмами измерения частоты основного тона, реализованными на основе: автокорреляционной функции и ее модификаций («YIN»), устойчивого метода отслеживания основного тона (Robust Algorithm for Pitch Tracking, RAPT) и оценки основного тона пилообразной формы (Sawtooth Waveform Inspired Pitch Estimation, SWIPE).

Результаты. Разработан алгоритм измерения частоты основного тона для систем обнаружения паттернов пограничных психических расстройств. Суть алгоритма заключается в разложении речевого сигнала на частотные составляющие с помощью адаптивного метода анализа нестационарных сигналов – улучшенной полной множественной декомпозиции на эмпирические моды с адаптивным шумом и выделении составляющей, содержащей основной тон. Представлена блок-схема разработанного алгоритма и подробное математическое описание. Проведено исследование алгоритма с использованием сформированной верифицированной базы сигналов здоровых пациентов и пациентов с психическими расстройствами мужского и женского пола в возрасте от 18 до 60 лет.

Выводы. В соответствии с результатами исследования разработанный алгоритм измерения частоты основного тона обеспечивает повышение точности

¹ Работа выполнена при финансовой поддержке Гранта Президента РФ (грант № МК-250.2017.8).

определения пограничных психических расстройств: для ошибки первого рода в среднем точнее на 10,7 % и для ошибки второго рода – на 4,7 %.

Ключевые слова: речевой сигнал, паттерн, частота основного тона, улучшенная полная множественная декомпозиция на эмпирические моды с адаптивным шумом, психогенные расстройства.

A. K. Alimuradov, A. Yu. Tychkov, P. P. Churakov, A. V. Ageykin

DEVELOPMENT OF AN ALGORITHM FOR SPEECH SIGNALS PROCESSING TO DETERMINE INFORMATIVELY SIGNIFICANT PARAMETERS OF BORDERLINE MENTAL DISORDERS

Abstract.

Background. The objects of the study are patients of the Regional Mental Hospital named after K.R. Evgrafov with borderline mental disorders, who have fairly high percentage of false-negative diagnostic results for these diseases. The subjects of the study are algorithms for speech signals processing to diagnose borderline mental disorders. The goal is to develop an algorithm to measure the pitch frequency for systems that detect patterns of borderline mental disorders.

Materials and methods. Informative parameters of speech signals – patterns – are used as research materials. To effectively process speech signals, we use the decomposition method for empirical modes and its modification-complete MDEM with adaptive noise. The results of the study are evaluated in comparison with the known algorithms pitch frequency measuring, and realized on the basis of: the auto-correlation function and its modifications ("YIN"), the stable method of main tone tracking (Robust Algorithm for Pitch Tracking, RAPT) and the sawtooth pitch estimate (Sawtooth Waveform Inspired Pitch Estimation, SWIPE).

Results. An algorithm for pitch frequency measuring for systems that detect patterns of borderline mental disorders has been developed. The essence of the algorithm is the decomposition of speech signals into frequency components using the adaptive method for analyzing non-stationary signals – improved complete multiple decomposition into empirical modes with adaptive noise and isolation of the component containing the fundamental tone. The article adduces a block diagram of the developed algorithm together with detailed mathematical description. The algorithm is investigated using the formed verified signal base of healthy patients and patients with psychogenic disorders of both genders aged from 18 to 60 years.

Conclusions. In accordance with the study results, the developed algorithm for measuring the fundamental tone frequency provides increased accuracy of borderline mental disorders detection: for an error of first kind, on the average, it is more accurate by 10.7 % and for a second type error – by 4.7 %.

Key words: speech signal, pattern, pitch frequency, improved complete multiple decomposition into empirical modes with adaptive noise, psychogenic disorders.

Введение

Оценка состояния психического здоровья на сегодня является социально значимой проблемой для каждого государства, поскольку напрямую связана с формированием здорового образа жизни населения. По данным Всемирной организации здравоохранения, современные социально значимые заболевания, являющиеся главной причиной временной нетрудоспособности, инвалидности и смертности населения, оказывая при этом негативное влия-

ние на социально-экономические факторы развития государства, напрямую связаны с психическим здоровьем населения [1].

Психические расстройства представляют обширную группу болезненных нарушений, объединенных общими признаками: преобладание невротического уровня психопатологических нарушений; взаимосвязь с вегетативными дисфункциями и соматическими проявлениями.

В настоящее время для обнаружения пограничных психических расстройств применяются различные экспериментально-статистические методики и дифференциации методов обработки сигналов по доступным каналам регистрации реакций организма человека. Особый интерес вызывают способы оценки, реализованные на основе: данных видеоряда, отражающего мимические и жестикулярные изменения [2, 3]; сигналов, отражающих параметры физиологической активности организма человека (электроэнцефалограмма, электрокардиограмма, электромиограмма и др.) [4–6]; биохимических параметров крови [7, 8]; параметров рукописного и клавиатурного написания текстов [9, 10]; параметров окулографии («*eye tracking*») [11, 12].

Существенным недостатком, ограничивающим широкое практическое применение указанных способов, является обязательное условие контактной регистрации/взятия пробы/написания, что, безусловно, влияет на психическое состояние, эффективно оценить которое уже не представляется возможным. Наиболее перспективным и адаптивным (в режиме реального времени и свободной активности) является способ на основе анализа речевых сигналов (РС) [13–15].

1. Материалы и методы

1.1. Информативные параметры речевых сигналов, отражающие пограничные психические расстройства

Важность анализа РС с целью диагностирования нарушений работы нервной системы отмечена в работе [16], в которой авторы показали, что группировка определенных информативных параметров отражает предположительно лежащую в основе патологию.

Вид и степень выраженности психических расстройств кодируются в определенные информативные параметры РС – паттерны. Обзорный анализ [17–22] в области речеобразования, психологии и психолингвистики выявил, что характеристики речи, способные служить паттернами психогенных состояний (проявляющихся на уровне голосовых сегментов, слогов, слов и целых предложений, связанных с геометрической формой и динамикой изменения речевого аппарата), можно разделить на три основные группы: спектрально-временные, кепстральные и амплитудно-частотные.

Каждая группа паттернов предназначена для описания отдельных аспектов РС и находит свое применение в обнаружении пограничных психических расстройств.

1.2. Частота основного тона

Речь представляет собой нестационарный акустический сигнал сложной формы, амплитудные и частотные характеристики которого быстро изменяются во времени. Речь состоит из вокализованных и невокализованных

участков, образующихся соответственно в результате периодических и непериодических колебаний голосовых связок. Периодические колебания голосовых связок называется основным тоном (ОТ). Частота колебаний связок является важным информативным параметром речи, называемым частотой основного тона (ЧОТ). С акустической точки зрения ЧОТ – это первая составляющая формантных частот («гармонического сита») речи. Кроме ЧОТ, вокальными характеристиками речи также являются: интенсивность ОТ, динамика изменения интенсивности ОТ, динамика изменения ЧОТ, девиация ЧОТ и отношение интенсивности гармоник к интенсивности ОТ.

Особенностью ЧОТ при расстройствах является то, что интервалы следования импульсов голосовых связок непрерывно меняются в значительных пределах. Во многих случаях длительность вокализованных участков речи невелика, значительную часть занимают переходные процессы.

1.3. Подходы измерения ЧОТ

Задача измерения ЧОТ представляет собой: выделение контура ОТ, маркировку периодов ОТ и измерение ЧОТ. В настоящее время известно большое количество способов измерения ЧОТ, которые в общем можно классифицировать как способы во временной, частотной и частотно-временной областях.

Во временной области [23, 24] измерение осуществляется за счет анализа распределения пиков, переходов через ноль, корреляции (автокорреляции, взвешенной и нормированной автокорреляции) осциллограммы сигнала. Временные методы являются самыми точными, но требуют тщательной фильтрации и настройки (работа только с вокализованными участками) исходного РС. Основным недостатком является высокая чувствительность к уровню шума в сигнале.

В частотной области [25, 26] измерение осуществляется за счет использования максимальных значений энергии спектра (пиков) и сопоставления их с частотами, кратными ЧОТ. Основным недостатком частотных методов является наличие в рассматриваемой полосе частот, помимо ЧОТ, второй или третьей гармоники с большей энергией.

В частотно-временных подходах [27, 28] измерение осуществляется за счет анализа предполагаемого контура ОТ, выделения мгновенных максимумов отдельных гармоник и разделения сигнала на вокализованные и невокализованные участки. Недостатком частотно-временных способов является большая вероятность получения мгновенного максимума энергии в невокализованных участках из-за присутствия шума в РС.

Широкое практическое применение получили подходы, реализованные на основе: автокорреляционной функции и ее модификаций («YIN») [29], устойчивого метода отслеживания основного тона (Robust Algorithm for Pitch Tracking, RAPT) [30] и оценки основного тона пилообразной формы (Sawtooth Waveform Inspired Pitch Estimation, SWIPE) [31]. Популярность данных алгоритмов обусловлена хорошей функциональностью, низким процентом грубых ошибок и наличием свободно распространяемых программных реализаций.

Тем не менее с учетом нерегулярности моторики органов речевого аппарата при пограничных психических расстройствах возможности данных

алгоритмов существенно ограничены. Ограничение обусловлено использованием неэффективных и неадаптивных методов обработки сложных нестационарных РС, приводящих к низкой точности и большим погрешностям в измерениях ЧОТ.

В работе предлагается алгоритм измерения ЧОТ для систем обнаружения паттернов пограничных психических расстройств. Исследование является развитием ранее опубликованных трудов авторов [32, 33].

1.4. Методы декомпозиции на эмпирические моды

Исследования методов обработки РС выявили перспективность использования адаптивной технологии анализа нестационарных сигналов – декомпозиции на эмпирические моды [34].

Декомпозиция на эмпирические моды (ДЭМ) [34] – 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:

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

где $x(n)$ – исходный сигнал; $IMF_i(n)$ – ЭМ; $r_i(n)$ – конечный остаток; $i = 1, 2, \dots, I$ – номер ЭМ; n – дискретный отсчет времени ($0 < n \leq N$, N – количество дискретных отсчетов в сигнале).

В результате разложения РС с помощью ДЭМ в одной ЭМ могут оказаться несоизмеримые по амплитудному и частотному масштабам колебательные функции или наоборот – соизмеримые колебательные функции могут оказаться в разных модах. Это явление называется смешиванием ЭМ. Для ее решения был предложен новый метод [35]: множественная ДЭМ (МДЭМ). Суть метода заключается в добавлении к исходному сигналу белого шума для создания новых экстремумов:

$$x_j(n) = x(n) + w_j(n),$$

где $x_j(n)$ – шумовые копии исходного сигнала; $w_j(n)$ – реализации белого шума с нулевой средней единичной дисперсией.

Аналитическое выражение МДЭМ выглядит следующим образом:

$$x_j(n) = \sum_{i=1}^I IMF_{ji}(n) + r_{ji}(n),$$

где $j = 1, 2, \dots, J$ – количество реализаций белого шума.

Таким образом, в результате разложения получаются более регулярные ЭМ с соизмеримыми масштабами колебательных функций. Основным недостатком МДЭМ – декомпозиция не является полной, т.е. каждая шумовая копия исходного сигнала $x_j(n)$ разлагается независимо от других реализаций и

для каждой из них остаток $r_{ji}(n) = r_{ji-1}(n) - IMF_{ji}(n)$ вычисляется на каждом этапе, без связи между различными реализациями. Кроме этого, в ЭМ наблюдается остаточный белый шум, а различные реализации шумовых копий могут порождать разное количество ЭМ, что затрудняет окончательное усреднение.

Другая разновидность ДЭМ – комплементарная МДЭМ [36] – качественнее решает проблему остаточного шума, используя дополнительные (добавляя и вычитая) пары шумов с прямыми и инверсными значениями амплитуды. Тем не менее проблема усреднения остается нерешенной, поскольку разные шумовые копии исходного сигнала также могут производить разное количество ЭМ:

$$\begin{bmatrix} x_j(n) \\ x_j^*(n) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} x(n) \\ w_j(n) \end{bmatrix},$$

где $x_j(n)$ – зашумленный сигнал белым шумом; $x_j^*(n)$ – зашумленный сигнал белым шумом с инверсными значениями амплитуды.

С учетом этих недостатков в работе [37] предлагается новый метод, называемый полной МДЭМ с адаптивным шумом (ПМДЭМАШ). Основная идея метода заключается в добавлении к исходному сигналу контролируемого шума для создания новых экстремумов. Первая ЭМ извлекается по методике МДЭМ, усредняя первые моды сигнала с белым шумом:

$$\overline{IMF_1}(n) = \frac{1}{J} \sum_{j=1}^J IMF_{j1}(n) = \widetilde{IMF_1}(n).$$

Затем вычисляется первый остаток, независимый от реализации шума:

$$r_1(n) = x(n) - \widetilde{IMF_1}(n).$$

Для дальнейшего извлечения мод к текущему первому остатку добавляется конкретный шум. Этот шум представляет собой ЭМ белого шума, полученную по методике ДЭМ.

Несмотря на вышеупомянутые достоинства метода ПМДЭМАД, в работе [38] авторы отмечают следующие недостатки данного метода:

- ЭМ содержат остаточный шум;
- информативные моды о сигнале при разложении извлекаются «позже», чем при МДЭМ с некоторыми «паразитными» модами на ранних этапах декомпозиции.

В работе [38] авторы решают отмеченные недостатки и предлагают усовершенствованную полную МДЭМ с адаптивным шумом.

Суть уменьшения остаточного шума заключается в использовании локальных средних значений вместо значений ЭМ.

Метод МДЭМ независимо разлагает каждую реализацию сигнала с шумом, поэтому на первом этапе каждой реализации декомпозиции имеется одно локальное среднее значение и одна мода.

Учитывая, что истинная мода может быть определена как разность между текущим остатком и усредненным значением его локальных средних величин, получим следующее выражение:

$$E_1(x(n)) = x(n) - M(x(n)),$$

где M – оператор, создающий локальное среднее значение применяемого сигнала.

Тогда для первых ЭМ, полученных методами МДЭМ и ПМДЭМАШ, выражение будет иметь следующий вид:

$$\widetilde{IMF}_1(n) = \langle E_1(x_j(n)) \rangle = \langle x_j(n) - M(x_j(n)) \rangle = \langle x_j(n) \rangle - \langle M(x_j(n)) \rangle,$$

где $\langle \rangle$ – действие усреднения.

Оценивая только локальное среднее значение и вычитая его из исходного сигнала, получим следующее выражение:

$$\widetilde{IMF}_1(n) = x(n) - \langle M(x_j(n)) \rangle.$$

Суть устранения возникающих паразитных ЭМ на ранних этапах декомпозиции заключается в уменьшении перекрытия масштабно-энергетических пространств первых двух мод.

Алгоритм и математическое описание улучшенной ПМДЭМАШ:

Шаг 1. С помощью аппарата ДЭМ и выражения из формулы $\langle E_1(x_j(n)) \rangle = \langle x_j(n) \rangle - \langle M(x_j(n)) \rangle$ локальных средних значений шумовых копий исходного сигнала ($x_j(n) = x(n) + \beta_0 E_1(w_j(n))$) определяется первый остаток $r_1(n) = \langle M(x_j(n)) \rangle$.

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 $r_1(n) + \beta_1 E_2(w_j(n))$ и определяется вторая мода:

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

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

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

Шаг 5. Вычисляется i -я мода: $\widetilde{IMF}_i(n) = r_{i-1}(n) - r_i(n)$.

Шаг 6. Переход к шагу 4 для следующего значения i .

Константы $\beta_i = \varepsilon_i \text{std}(r_i(n))$ выбираются таким образом, чтобы получить желаемое отношение сигнал/шум между добавленным шумом и остатком, к которому добавляется шум. Обратите внимание, что при МДЭМ отношение сигнал/шум между добавленным шумом и остатком увеличивается на порядок i . Это связано с тем, что энергия шума в i -м остатке ($i > 1$) является лишь малой энергией шума, добавленного в начале алгоритма. Чтобы эмулировать это поведение, в данной работе мы установим β_0 так, чтобы ε_0 было прямо противоположно желаемому отношению сигнал/шум между первым добавленным шумом и анализируемым сигналом: если мы выражаем

отношение сигнал/шум как фактор стандартных отклонений, то имеем $\beta_0 = \varepsilon_0 \text{std}(x(n)) / \text{std}(E_1(w_f(n)))$. Для получения шумовых реализаций с меньшими амплитудами для последних стадий декомпозиции в остальных модах мы будем воспринимать шум как результат его предварительной обработки ДЭМ, т.е. без их нормализации стандартным отклонением.

2. Описание алгоритма

Авторами разработан алгоритм измерения ЧОТ для обнаружения паттернов пограничных психических расстройств. Блок-схема алгоритма (блоки 1–8) представлена на рис. 1.

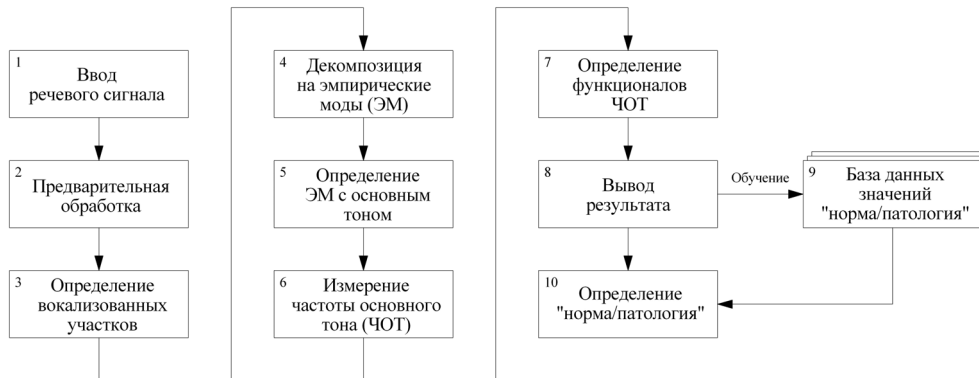


Рис. 1. Блок-схема алгоритма измерения ЧОТ для обнаружения паттернов пограничных психических расстройств

Суть алгоритма заключается в разложении РС на частотные составляющие, выделении составляющей, содержащей ОТ, и измерении ЧОТ.

Работа алгоритма осуществляется следующим образом (рис. 2):

- обработка в блоке 2 выполняется для РС в полном временном интервале;
- обработка в блоке 3 выполняется в режиме скользящего окна длительностью 15 мс с перекрытием 7,5 мс;
- обработка в блоках 4–8 выполняется в режиме скользящего окна длительностью 20 мс (возможны вариации) с перекрытием 10 мс.

Рассмотрим каждый этап обработки подробнее.

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

Предварительная обработка. В рамках предварительной обработки в первую очередь осуществляется удаление постоянной составляющей (смещение сигнала относительно нуля на некую постоянную величину), которая обычно возникает в аналого-цифровом преобразователе. Для того чтобы удалить постоянную составляющую, или, другими словами, выровнять сигнал относительно нуля, определяется среднее арифметическое значение всех отсчетов сигнала и вычитается из исходного сигнала.

Следующим этапом предобработки является фильтрация РС с помощью фильтра высоких частот Чебышева четвертого порядка для удаления частот ниже 130 Гц, которые включают в себя основной гул, треск и другие шумы в этом диапазоне.

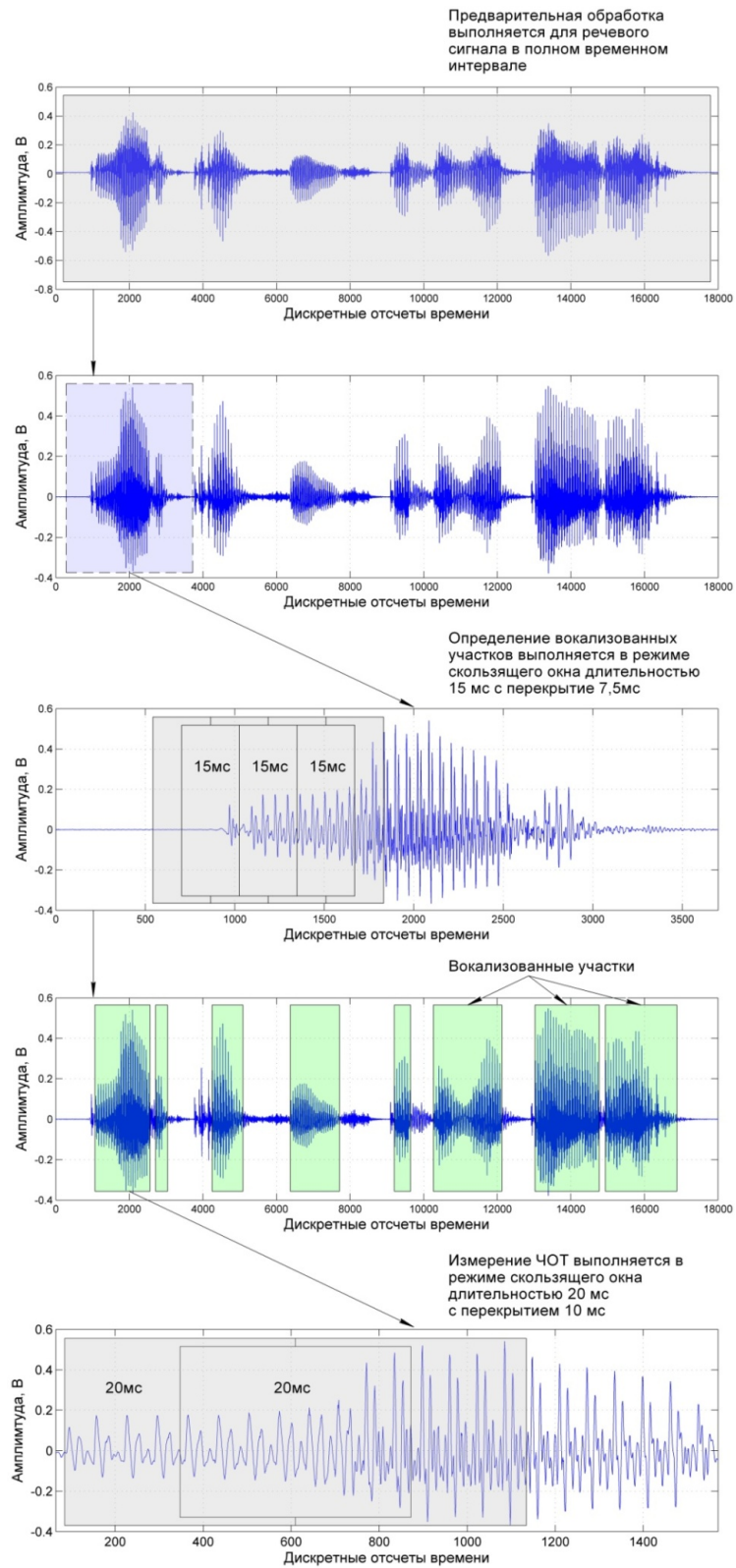


Рис. 2. Работа алгоритма

Фильтрация с частотным срезом на 130 Гц не влияет на полезную информацию в сигнале [39]. В завершение осуществляется коррекция естественных искажений спектра (–6 дБ на октаву), возникающих в речевом аппарате человека при произнесении речи [40]. РС пропускают через корректирующий фильтр с передаточной функцией:

$$W(z) = \sum_{k=0}^m a_k z^{-k},$$

где a_k – постоянные коэффициенты; m – целое число ($m > 0$); k – номер коэффициента. Чаще всего $m = 1$, а передаточная функция имеет вид

$$W(z) = a_0 - a_1 z^{-1}.$$

Определение вокализованных участков. Вокализованные участки выделяются из РС с использованием кластерного анализа в пространстве вычисленных упомянутых выше значений ZCR, ACR и PWR [41]. Сегменты вокализации были идентифицированы как сегменты с самым высоким значением PWR и самыми низкими значениями ZCR и ACR. Решения сглаживались с использованием медианного фильтра 5-го порядка на основе решающего правила принятия решения: вокализованные сегменты короче 30 мс классифицировались как сегменты, не содержащие вокализации (сформированные без участия голосовых связок); но сегменты, не содержащие вокализации, короче 20 мс были классифицированы как вокализованные в силу физиологического аспекта формирования речи [41].

Декомпозиция на эмпирические моды. На основе детализированного анализа преимуществ и недостатков различных видов декомпозиции и с учетом специфики РС при пограничных психических расстройствах авторы приняли решение использовать для разложения сигнала на частотные составляющие улучшенную ПМДЭМАШ [39].

Результат разложения вокализованного участка РС с использованием улучшенной ПМДЭМАШ приведен на рис. 3. Параметры декомпозиции: стандартное отклонение шума – 0,2 мВ, количество реализаций – 500, максимально допустимое количество просеивающих итераций – 5000.

Как видно из рис. 3, вокализованный участок РС разложен на 9 ЭМ. Две первые моды (черный цвет) содержат основной шум, присутствующий в исходном сигнале. Шестая мода и последующие являются низкочастотными и соответствуют присутствующему в сигнале тренду. Ценная высокочастотная информация, связанная со смыканием голосовых складок, появляется с третьей по пятую ЭМ.

Определение ЭМ с основным тоном. Суть определения ЭМ с основным тоном заключается в предположении, что информативные моды (ЭМ3 – ЭМ5) имеют большую энергию, чем трендовые моды. Амплитудное распределение ЭМ хорошо описывается с помощью функции кратковременной энергии. В разработанном алгоритме для сжатия амплитуды сигнала в большом динамическом диапазоне применяется логарифмирование энергии по аналогии с функционированием слухового аппарата человека:

$$LE_i = \log_2 \sum_{n=1}^N (IMF_i(n))^2,$$

где LE_i – логарифм энергии ЭМ.

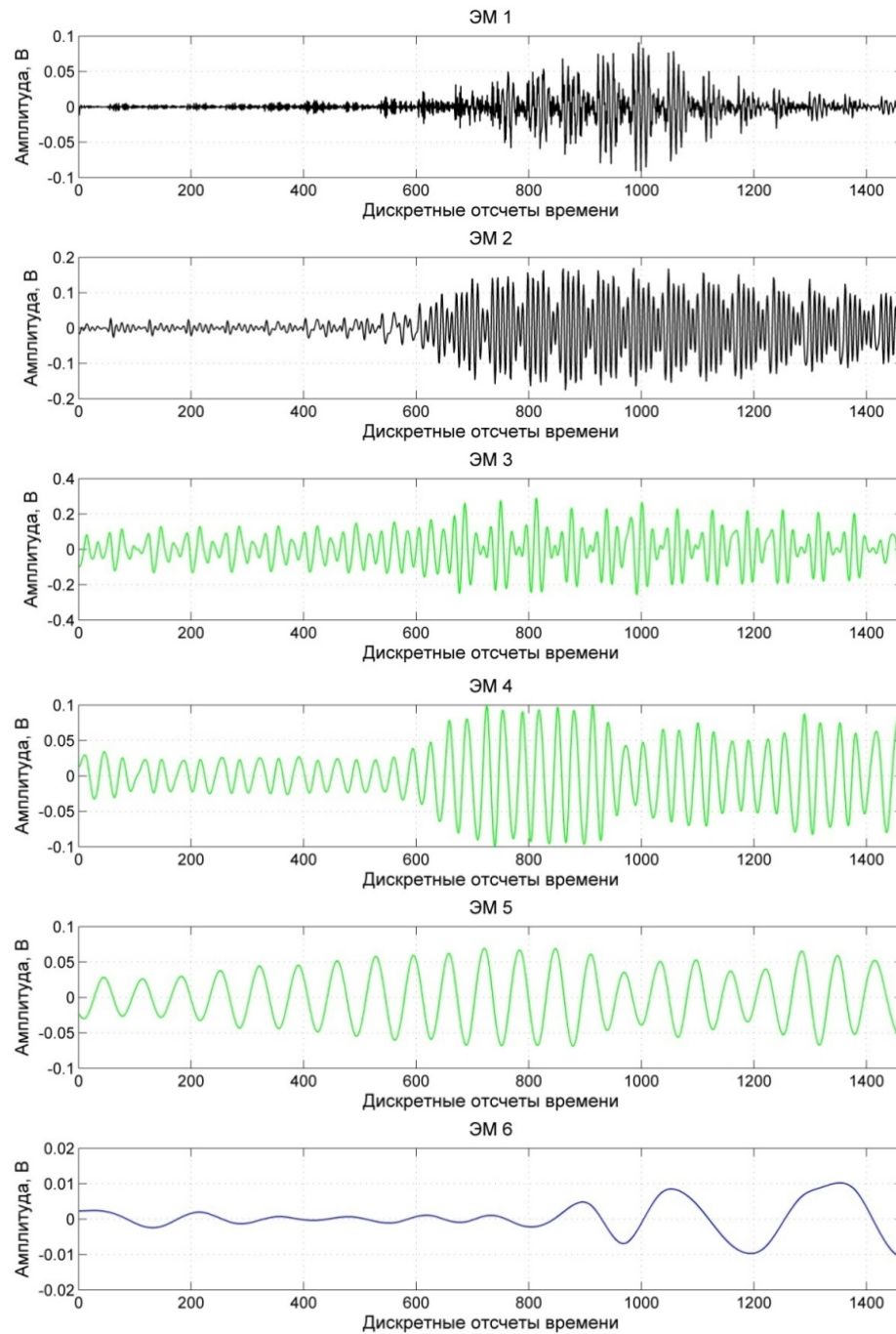


Рис. 3. Результат разложения вокализованного участка РС методом улучшенной ПМДЭМАШ

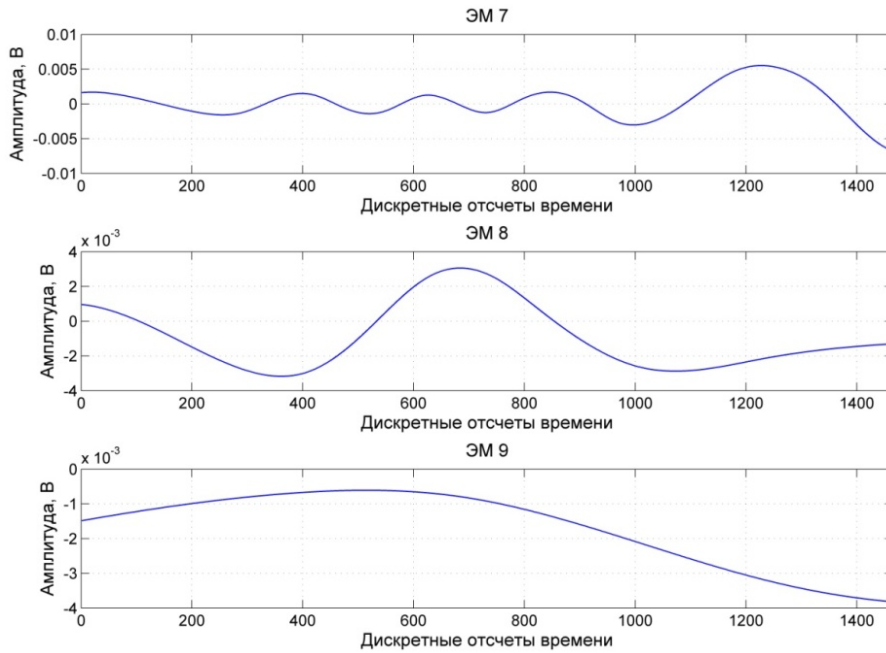


Рис. 3. Окончание

Процесс определения ЭМ с ОТ заключается в последовательном вычислении разницы значений логарифмов энергии по модулю между текущей и последующей модами:

$$d = |LE_i - LE_{i+1}|,$$

где d – разница между значениями логарифмов энергии между текущей и последующей ЭМ.

В результате из последовательности полученных значений d большему из них соответствует резкий спад энергии между информативной ЭМ, содержащей ОТ, и трендовой [42, 43]. На рис. 4 представлена графическая интерпретация процесса определения моды, содержащей ОТ. В соответствии с правилом определения из рис. 4,а следует, что 5-я ЭМ может содержать ОТ. Анализ спектрального распределения 5-й ЭМ и исходного РС подтверждает правильность определения моды, содержащей ОТ: единственная гармоническая составляющая ЧОТ 5-й ЭМ соответствует первой составляющей «гармонического сита» исходного РС (рис. 4,б). На рис. 5 представлены осциллограммы исходного сигнала 5-й моды с ОТ.

В качестве единицы измерения спектральной плотности мощности выбрана составная величина дБВт/Гц с опорным уровнем в 1 Вт/Гц (мощность, выделяющаяся в полосе частот шириной в 1 Гц).

Измерение ЧОТ осуществляется с использованием функции измерения мгновенной энергии сигнала – оператора Тигра, обладающего простотой, эффективностью и хорошей восприимчивостью к изменению РС:

$$T(n) = \left(IMF_{i,PF}(n) \right)^2 - IMF_{i,PF}(n-1) \times IMF_{i,PF}(n+1),$$

где $T(n)$ – функция оператора Тигра; $IMF_{i,PF}(n)$ – ЭМ, содержащая ОТ.

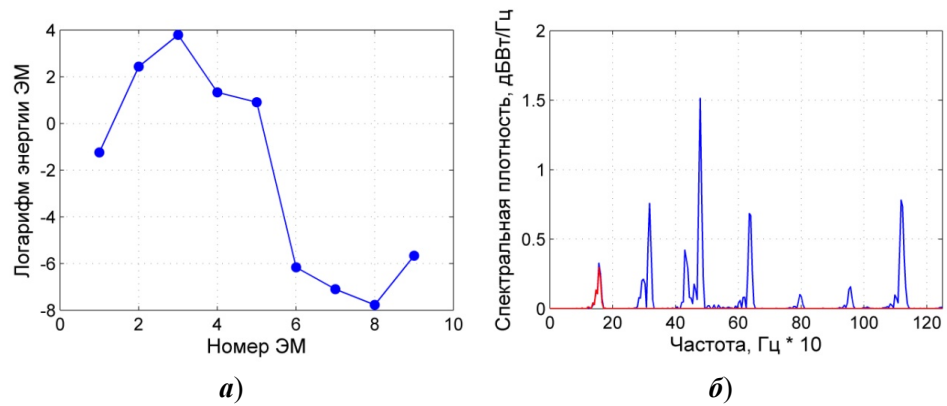


Рис. 4. Определения ЭМ, содержащей ОТ: *а* – логарифмы энергии ЭМ; *б* – спектральное распределение исходного РС и 5-й ЭМ

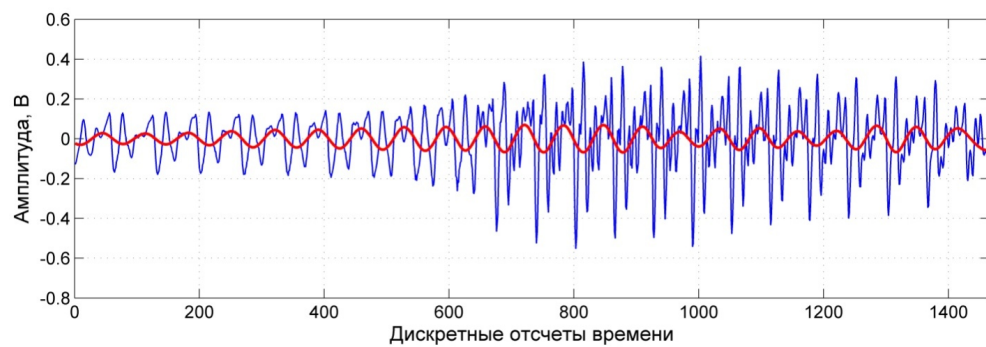


Рис. 5. Осциллограммы исходного сигнала пятой моды с ОТ

На рис. 6 представлены осциллограмма и функция оператора Тигра 5-й ЭМ.

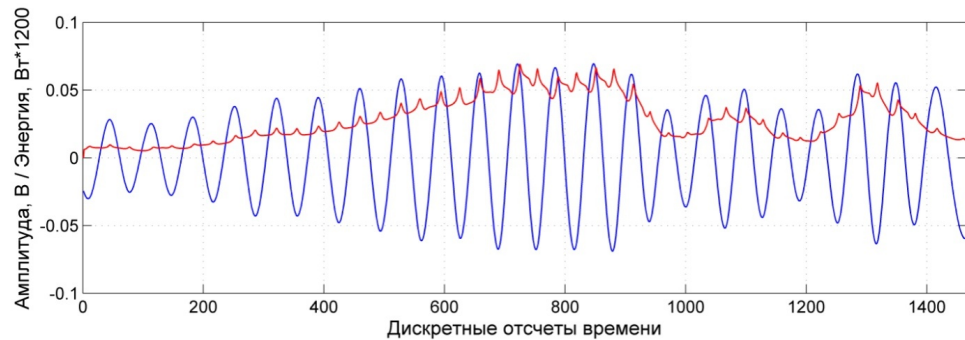


Рис. 6. Осциллограмма и функция оператора Тигра 5-й ЭМ

Для измерения частоты используются близкорасположенные максимумы, функции оператора Тигра, между которыми определяется разница в дискретных отсчетах времени, вычисляется период ОТ (в секундах) и ЧОТ (в герцах):

$$P_0 = \frac{T_{\max}(n+2) - T_{\max}(n)}{f_d}, \quad f_0 = \frac{1}{P_0},$$

где P_0 – ОТ, f_0 – ЧОТ; $T_{\max}(n)$, $T_{\max}(n+1)$ – максимумы функции оператора Тигра; f_d – частота дискретизации.

Определение ЧОТ. Для расширения информационного пространства о частоте ОТ определяются следующие ее значения:

– среднее значение ЧОТ, Гц:

$$f_{0,\text{mean}} = \frac{1}{P} \sum_{p=1}^P f_{0,p},$$

где $p = 1, 2, \dots, P$ – номер периода ОТ;

– максимальное $\max(f_0)$ и минимальное $\min(f_0)$ значения ЧОТ, Гц;

– стандартное отклонение контура ЧОТ:

$$SD_{f_0} = \frac{1}{P-1} \sum_{p=1}^P (f_{0,p} - f_{0,\text{mean}})^2;$$

– диапазон фоновых частот:

$$PFR = 12 \times \frac{\log\left(\frac{\max(f_0)}{\min(f_0)}\right)}{\log_2};$$

– среднее абсолютное значение джиттера:

$$MAJ = \frac{1}{P-2} \sum_{p=P-1}^1 |f_{0,p-1} - f_{0,p}|;$$

– джиттер:

$$J = \frac{MAJ}{f_{0,\text{mean}}};$$

– среднее относительное возмущение ЧОТ, сглаженное за три периода ОТ:

$$RAP = \frac{\frac{1}{P-2} \sum_{p=2}^{P-1} \left| \frac{f_{0,p+1} + f_{0,p} + f_{0,p-1}}{3} - f_{0,p} \right|}{f_{0,\text{mean}}} \cdot 100;$$

– коэффициент возмущения ЧОТ, сглаженный за пять периодов ОТ:

$$PPQ = \frac{\frac{1}{P-4} \sum_{p=3}^{P-2} \left| \frac{\sum_{k=p-2}^{p+2} f_{0,k}}{5} - f_{0,i} \right|}{f_{0,\text{mean}}} \cdot 100.$$

Вывод результатов. На данном этапе работы алгоритма осуществляется формирование векторов полученных паттернов ЧОТ и ее значений в удобный для дальнейшего определения «норма/патология» вид (рис. 1, блоки 9, 10).

3. Исследование алгоритма

3.1. Описание базы данных речевых сигналов

Для проведения исследований разработанного алгоритма сформирована 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 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, которые поступили 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 неработающих. Большинство пациентов имели высшее или неоконченное высшее образование (69,2 %).

3.2. Результаты исследования

Для оценки эффективности алгоритма использовался параметр – ошибки первого и второго рода. В рамках данного исследования задачей является определение паттернов пограничных психических расстройства, поэтому ошибкой первого рода будет ложное присваивание статуса «норма» РС, произнесенному человеком с психогенным расстройством, а ошибкой второго рода – ложное присваивание статуса «патология» РС, произнесенному здоровым человеком.

Результаты исследования разработанного алгоритма оценивались в сравнении с упомянутыми в начале статьи алгоритмами измерения ЧОТ, программная реализация которых имеется в открытом доступе: «YIN», RAPT SWIPE.

В табл. 1 представлены результаты определения пограничных психических расстройств для трех сравниваемых алгоритмов.

4. Обсуждение и выводы

Из полученных результатов следует, что разработанный авторами алгоритм обеспечивает повышение точности измерения ЧОТ:

- для ошибки первого рода точнее на 8 %, чем у алгоритма RAPT, на 13 % – чем у алгоритма «YIN», и на 11 % – чем у алгоритма SWIPE;
- для ошибки второго рода точнее на 4, 7 и 3 % соответственно.

Таблица 1

Результаты определения пограничных психических расстройств

Прогнозируемый результат	Результат определения		Ошибки первого и второго рода, %	
	Патология	Норма		
Алгоритм на основе устойчивого отслеживания основного тона (RAPT)				
Патология	84 чел.	16 чел.	1-го	16
Норма	8 чел.	92 чел.	2-го	8
Алгоритм на основе автокорреляционной функции («YIN»)				
Патология	79 чел.	21 чел.	1-го	21
Норма	11 чел.	89 чел.	2-го	11
Алгоритм на основе оценки основного тона пилообразной формы (SWIPE)				
Патология	81 чел.	19 чел.	1-го	19
Норма	7 чел.	93 чел.	2-го	7
Разработанный алгоритм				
Патология	92 чел.	8 чел.	1-го	8
Норма	4 чел.	96 чел.	2-го	4

Данные результаты позволяют сделать вывод, что разработанный алгоритм на основе метода улучшенной ПМДЭМАШ может быть успешно использован в системах обнаружения паттернов пограничных психических расстройств и внедрен в клиническую практику врача-психиатра.

Библиографический список

1. Современные представления о специфике и проблемах диагностики пограничных психических расстройств / А. Ю. Тычков, А. В. Агейкин, А. К. Алимуратов, В. Б. Калистратов, С. Ю. Митрошина // Психическое здоровье. – 2017. – № 5. – С. 69–75.
2. **Pantic, K. M.** A dynamic texture-based approach to recognition of facial actions and their temporal models / Koelstra M. Pantic, Ioannis Patras // IEEE transactions on pattern analysis and machine intelligence. – 2010. – Vol. 32. – P. 1940–1954.
3. **Бобков, А. С.** Модели и методы определения эмоционального состояния человека на основе автоматизированной обработки и анализа видеoinформации : дис. ... канд. техн. наук : 05.13.01 / Бобков Артем Сергеевич. – Москва, 2013. – 170 с.
4. **Кислова, О. О.** Психофизиологические предпосылки успешности распознавания эмоциональной речевой экспрессии : дис. ... канд. биол. наук : 03.00.13 / Кислова Ольга Олеговна. – Москва, 2009. – 151 с.
5. **Agrafioti, F.** ECG in Biometric Recognition: Time Dependency and Application Challenges / Foteini Agrafioti, Doctor of Philosophy Graduate Department of Electrical and Computer Engineering University of Toronto. – Toronto, 2011.
6. **Kappeler-Setz, C.** Towards long term monitoring of electrodermal activity in daily life / Cornelia Kappeler-Setz, Franz Gravenhorst, Johannes Schumm, Bert Arnrich, Gerhard Tröster // Journal Personal and Ubiquitous Computing. – 2013. – Vol. 17, iss. 2. – P. 261–271.
7. **Farwell, L. A.** Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials / Lawrence A. Farwell // Cogn Neurodyn. – 2012. – Vol. 6. – P. 115–154. DOI 10.1007/s11571-012-9192-2.

8. **Quan, C.** Sentence Emotion Analysis and Recognition Based on Emotion Words Using Ren-CECps / Changqin Quan, Fuji Ren // *International Journal of Advanced Intelligence*. – 2010. – Vol. 2, № 1. – P.105–117.
9. **Заболеева-Зотова, А. В.** Задача создания системы автоматизированного распознавания эмоций / А. В. Заболеева-Зотова, Ю. А. Орлова, В. Л. Розалиев, А. С. Бобков // *Материалы международной научно-технической конференции OSTIS*. – Минск : БГУИР, 2012. – С. 347–350.
10. **Likforman-Sulem, L.** EMOTHAW: A Novel Database for Emotional State Recognition From Handwriting and Drawing / Laurence Likforman-Sulem, Anna Esposito, Marcos Faundez-Zanuy, Stéphan Cléménçon, Gennaro Cordasco // *IEEE Transactions on Human-Machine Systems*. – 2017. – Vol. 47, iss. 2. – P. 273–284. DOI: 10.1109/THMS.2016.2635441.
11. **Барабанщиков, В. А.** Методы регистрации движений глаз в психологии: основы учебно-методического комплекса / В. А. Барабанщиков, А. В. Жегалло // *Экспериментальная психология*. – 2014. – Т. 7, № 1. – С. 132–137.
12. **Loughland, C. M.** Visual scanpaths to positive and negative facial emotions in an outpatient schizophrenia sample / C. M. Loughland, L. M. Williams, E. Gordon // *Schizophrenia Research*. – 2002. – Vol. 55, iss. 1-2. – P. 159–170.
13. **Филатова, Н. Н.** Модель интерпретации знака эмоций по естественной речи / Н. Н. Филатова, К. В. Сидоров // *Известия Южного федерального университета. Технические науки*. – 2012. – № 9 (134). – С. 39–45.
14. **Давыдов, А. Г.** Выбор оптимального набора информативных признаков для классификации эмоционального состояния диктора по голосу / А. Г. Давыдов, В. В. Киселев, Д. С. Кочетков, А. В. Ткачя // *Компьютерная лингвистика и интеллектуальные технологии : по материалам ежегод. Междунар. конф. «Диалог»*. – Вып. 11 : в 2 т. Т. 1 : Основная программа конференции. – М. : Изд-во РГГУ, 2012. – С. 122–128.
15. **Williamson, J. R.** Vocal and facial biomarkers of depression based on motor incoordination and timing / James R Williamson, Thomas F. Quatieri, Brian S. Helfer, Gregory Ciccarelli, Daryush D. Mehta // *Proceedings of the 4th International Workshop on Audio-Visual Emotion – Challenge*, 2014. – С. 65–72.
16. **Darley, F. L.** Differential diagnostic patterns of dysarthria / F. L. Darley, A. E. Aronson, J. R. Brown // *J. Speech Lang. Hear. Res.* – 1969. – Vol. 12. – P. 47–57.
17. **Dorry, M. K.** Emotion Identification from Spontaneous Communication: A Thesis Submitted to the Department of Computer Science in Partial Fulfilment for the Degree of Master of Science in Computer Science / Mikiyas Kebede Dorry ; Addis Ababa University College of Natural Sciences. – Addis Ababa, Ethiopia, 2016. – P. 93.
18. **Schwenzer, M.** Impaired pitch identification as a potential marker for depression / Michael Schwenzer, Eva Zattarin, Michael Grozinger and Klaus Mathiak // *BMC Psychiatry*. – 2012. – Vol. 12, iss. 1. – P. 32–37.
19. **Morist, M. U.** Emotional speech synthesis for a radio dj: corpus design and expression modeling: master thesis MTG-UPF dissertation / M. U. Morist. – Barcelona, 2010.
20. **Yingthawornsuk, T.** Analysis of Speech Based on Spectral Entropy in Detecting Depressed among Control Subjects / Thaweesak Yingthawornsuk // *International Conference on Computer and Communication Technologies (ICCCCT'2012)*, May 26–27, 2012. – Phuket, 2012.
21. **Yu, B.** Cognitive impairment prediction in the elderly based on vocal biomarkers / Be Yu, Thomas F. Quatieri, James R. Williamson, James C. Mundt // *Sixteenth Annual Conference of the International Speech Communication Association*. – Dresden Germany, 2015. – P. 3734–3738.
22. **Grossberg, S.** The Adaptive Self-organization of Serial Order in Behavior: Speech, Language / Stephen Grossberg // *Pattern recognition by humans and machines: Speech perception*. – 2013. – P. 187.

23. **Staudacher, M.** Fast fundamental frequency determination via adaptive autocorrelation / Michael Staudacher, Viktor Steixner, Andreas Griessner and Clemens Zierhofer // EURASIP Journal on Audio, Speech, and Music Processing. – 2016. – Issue 1. Dec. – P. 8.
24. **Kumar, S.** Efficient Time Domain Fundamental Frequency Estimation (Pitch Estimation) / Sunil Kumar, Manisha Jangra, // International Journal of Electronics & Communication Technology (IJECT). – 2015. – Vol. 6, iss. 3.
25. **Eun-Joo, S.** Feature Extraction Method for Predicting Depression by Frequency Domain Analysis / Eun-Joo Seo, Kwang-Seok Hong // International Conference on Civil, Materials and Environmental Sciences (CMES 2015). – London, UK, 2015. – P. 600–603.
26. **Haraa, Y.** Method for estimating pitch independently from power spectrum envelope for speech and musical signal / Yoshifumi Haraa, Mitsuo Matsumoto, Kazunori Miyoshi // Journal of Temporal Design in Architecture and the Environment, 2009. – Vol. 9, iss. 9. – P. 121–124.
27. **Shahnaz, C.** A Spectro-Temporal algorithm for pitch frequency estimation from noisy observations / C. Shahnaz, W.-P. Zhu, and M. O. Ahmad // IEEE International Symposium on Circuits and Systems (ISCAS2008), May 18–21, 2008. – Seattle USA, 2008. – P. 1704–1707.
28. **Klapuri, A. P.** Multiple Fundamental Frequency Estimation Based on Harmonicity and Spectral Smoothness / Anssi P. Klapuri // IEEE Transactions on speech and audio processing. – 2003. – Vol. 11, № 6. – P. 804–816.
29. **Cheveigne, A.** «YIN» a fundamental frequency estimator for speech and music / A. Cheveigne, H. Kawahara // The Journal of the Acoustical Society of America. – 2002. – Vol. 111, № 4. – P. 1917–1930.
30. **Talkin, D.** A Robust Algorithm for Pitch Tracking (RAPT) // Chapter 14 in Speech Coding & Synthesis / D. Talkin ; ed. by W. B. Kleijn and K. K. Paliwan. – New York, USA, Elsevier Science, 1995. – P. 495–518.
31. **Camacho, A.** A sawtooth waveform inspired pitch estimator for speech and music / A. Camacho, J. G. Harris // The Journal of the Acoustical Society of America. – 2008. – Vol. 123, № 4. – P. 1638–1652.
32. **Alimuradov, A. K.** Research of Frequency-Selective Properties of Empirical Mode Decomposition Methods for Speech Signals' Pitch Frequency Estimation / A. K. Alimuradov // International Conference on Engineering and Telecommunication (EnT), Nov. 18–19, 2015, Russia, Moscow, – Moscow, 2015. – P. 77–79. DOI: 10.1109/EnT.2015.9.
33. Speech/pause detection algorithm based on the adaptive method of complementary decomposition and energy assessment of intrinsic mode functions / A. K. Alimuradov, A. Yu. Tychkov, A. V. Ageykin, P. P. Churakov, Y. S. Kvitka and A. P. Zaretskiy // XX IEEE International Conference on Soft Computing and Measurements (SCM), May 24–26, 2017, Russia, St. Petersburg. – St. Petersburg, 2017. – P. 610–613. DOI: 10.1109/SCM.2017.7970665.
34. **Huang, N. E.** The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis / N. E. Huang, Sh. Zheng, and R. L. Steven // Proc. R. Soc. Lond. – 1998. – Vol. A454. – P. 903–995.
35. **Zhaohua, Wu.** Ensemble empirical mode decomposition: A noise - assisted data analysis method / Wu Zhaohua, N. E. Huang // Advances in Adaptive Data Analysis. – 2009. – Vol. 1, № 1. – P. 1–41.
36. **Yeh, J.-R.** Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method / J.-R. Yeh, J.-S. Shieh, and N. E. Huang // Adv. Adapt. Data Anal. – 2010. – Vol. 2, № 2. – P. 135–156.
37. **Torres, M. E.** A complete Ensemble Empirical Mode decomposition with adaptive noise / M. E. Torres, M. A. Colominas, G. Schlotthauer, P. Flandrin // International

- Conference on Acoustics, Speech and Signal Processing (ICASSP-11). – Prague (CZ), 2011. – P. 4144–4147.
38. **Colominasa, M. A.** Improved complete ensemble EMD: A suitable tool for biomedical signal processing / Marcelo A. Colominasa, Gaston Schlotthauer, Maria E. Torres // *Biomedical Signal Processing and Control*. – 2014. – Vol. 14. – P. 19–29.
39. **Huang, X.** Spoken Language Processing. Guide to Algorithms and System Development / X. Huang, A. Acero, H.-W. Hon // Prentice Hall. – New Jersey, 2001. – 980 p.
40. **Фант, Г. К.** Акустическая теория речеобразования / Г. К. Фант ; пер. с англ. Л. А. Варшавского и В. И. Медведева ; науч. ред. В. С. Григорьева. – М. : Наука, 1964. – 284 с.
41. **Попов, Д. И.** Оптимизация цифровых систем обработки сигналов / Д. И. Попов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 1 (41). – С. 96–105.
42. **Alimuradov, A. K.** An Algorithm for Measurement of the Pitch Frequency of Speech Signals Based on Complementary Ensemble Decomposition Into Empirical Modes / A. K. Alimuradov // *Measurement techniques*. – 2017. – Vol. 59, № 12. – P. 1316–1323. DOI 10.1007/s11018-016-0994-1.
43. **Alimuradov, A. K.** Program Implementation of an Algorithm for Recognition of Speech Signals in the Labview Graphics Programming Environment / A. K. Alimuradov, A. Yu. Tychkov, M. V. Frantsuzov, P. P. Churakov // *Measurement techniques*. – 2015. – Vol. 58, iss. 9. – P. 965–969.

References

1. Tychkov A. Yu., Ageykin A. V., Alimuradov A. K., Kalistratov V. B., Mitroshina S. Yu. *Psikhicheskoe zdorov'e* [Mental health]. 2017, no. 5, pp. 69–75.
2. Pantic K. M., Patras Ioannis *IEEE transactions on pattern analysis and machine intelligence*. 2010, vol. 32, pp. 1940–1954.
3. Bobkov A. S. *Modeli i metody opredeleniya emotsional'nogo sostoyaniya cheloveka na osnove avtomatizirovannoy obrabotki i analiza videoinformatsii: dis. kand. tekhn. nauk: 05.13.01* [Human emotional condition determination methods and models on the basis of automatic videoinformation processing and analysis: dissertation to apply for the degree of the candidate of engineering sciences]. Moscow, 2013, 170 p.
4. Kislova O. O. *Psikhofiziologicheskie predposylki uspehnosti raspoznavaniya emotsional'noy rechevoy ekspresii: dis. kand. biol. nauk: 03.00.13* [Psychophysical prerequisites of successful detection of emotion speech expression: dissertation to apply for the degree of the candidate of biological sciences]. Moscow, 2009, 151 p.
5. Agraftioti F. *ECG in Biometric Recognition: Time Dependency and Application Challenges*. Toronto, 2011.
6. Kappeler-Setz C., Gravenhorst Franz, Schumm Johannes, Arnrich Bert, Tröster Gerhard *Journal Personal and Ubiquitous Computing*. 2013, vol. 17, iss. 2, pp. 261–271.
7. Farwell L. A. *Cogn Neurodyn*. 2012, vol. 6, pp. 115–154. DOI 10.1007/s11571-012-9192-2.
8. Quan C., Ren F. *International Journal of Advanced Intelligence*. 2010, vol. 2, no. 1, pp. 105–117.
9. Zabooleeva-Zotova A. V., Orlova Yu. A., Rozaliev V. L., Bobkov A. S. *Materialy mezhdunarodnoy nauchno-tekhnicheskoy konferentsii OSTIS* [Proceedings of an International Scientific and Technical Conference OSTIS]. Minsk: BGUIR, 2012, pp. 347–350.
10. Likforman-Sulem L., Esposito A., Faundez-Zanuy M., Cléménçon S., Cordasco G. *IEEE Transactions on Human-Machine Systems*. 2017, vol. 47, iss. 2, pp. 273–284. DOI: 10.1109/THMS.2016.2635441.
11. Barabanshchikov V. A., Zhegallo A. V. *Ekspperimental'naya psikhologiya* [Experimental psychology]. 2014, vol. 7, no. 1, pp. 132–137.

12. Loughland C. M., Williams L. M., Gordon E. *Schizophrenia Research*. 2002, vol. 55, iss. 1-2, pp. 159–170.
13. Filatova N. N., Sidorov K. V. *Izvestiya Yuzhnogo federal'nogo universiteta. Tekhnicheskies nauki* [Proceedings of South Federal University. Engineering sciences]. 2012, no. 9 (134), pp. 39–45.
14. Davydov A. G., Kiselev V. V., Kochetkov D. S., Tkachenya A. V. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po materialam ezhegod. Mezhdunar. konf. «Dialog»*. Vyp. 11: v 2 t. T. 1: *Osnovnaya programma konferentsii* [Computer linguistics and intelligent technologies: based on proceedings of the annual International Conference “Dialogue”. Issue 11: in 2 volumes. Vol.1: Main conference program]. Moscow: Izd-vo RGGU, 2012, pp. 122–128.
15. Williamson J. R., Quatieri T. F., Helfer B. S., Ciccarelli G., Mehta D. D. *Proceedings of the 4th International Workshop on Audio-Visual Emotion*. Challenge, 2014, pp. 65–72.
16. Darley F. L., Aronso A. E., Brown J. R. *J. Speech Lang. Hear. Res.* 1969, vol. 12, pp. 47–57.
17. Dorry M. K. *Emotion Identification from Spontaneous Communication: A Thesis Submitted to the Department of Computer Science in Partial Fulfilment for the Degree of Master of Science in Computer Science*. Addis Ababa University College of Natural Sciences. Addis Ababa, Ethiopia, 2016, p. 93.
18. Schwenzer M., Zattarin E., Grozinger M., Mathiak K. *BMC Psychiatry*. 2012, vol. 12, iss. 1, pp. 32–37.
19. Morist M. U. *Emotional speech synthesis for a radio dj: corpus design and expression modeling: master thesis MTG-UPF dissertation*. Barcelona, 2010.
20. Yingthawornsuk T. *International Conference on Computer and Communication Technologies (ICCCT'2012)*, May 26–27, 2012. Phuket, 2012.
21. Yu B., Quatieri T. F., Williamson J. R., Mundt J. C. *Sixteenth Annual Conference of the International Speech Communication Association*. Dresden Germany, 2015, pp. 3734–3738.
22. Grossberg S. *Pattern recognition by humans and machines: Speech perception*. 2013, p. 187.
23. Staudacher M., Steixner V., Griessner A., Zierhofer C. *EURASIP Journal on Audio, Speech, and Music Processing*. 2016, iss. 1, Dec., p. 8.
24. Kumar S., Jangra M. *International Journal of Electronics & Communication Technology (IJECT)*. 2015, vol. 6, iss. 3.
25. Eun-Joo S., Kwang-Seok Hong *International Conference on Civil, Materials and Environmental Sciences (CMES 2015)*. London, UK, 2015, pp. 600–603.
26. Haraa Y., Matsumotob M., Miyoshi K. *Journal of Temporal Design in Architecture and the Environment*. 2009, vol. 9, iss. 9, pp. 121–124.
27. Shahnaz C., Zhu W.-P., Ahmad M. O. *IEEE International Symposium on Circuits and Systems (ISCAS2008)*, May 18–21, 2008. Seattle USA, 2008, pp. 1704–1707.
28. Klapuri A. P. *IEEE Transactions on speech and audio processing*. 2003, vol. 11, no. 6, Nov., pp. 804–816.
29. Cheveigne A., Kawahara H. *The Journal of the Acoustical Society of America*. 2002, vol. 111, no. 4, pp. 1917–1930.
30. Talkin D. *Chapter 14 in Speech Coding & Synthesis*. Ed. by W. B. Kleijn and K. K. Paliwan. Elsevier Science. New York USA, 1995, pp. 495–518.
31. Camacho A., Harris J. G. *The Journal of the Acoustical Society of America*. 2008, vol. 123, no. 4, pp. 1638–1652.
32. Alimuradov A. K. *International Conference on Engineering and Telecommunication (EnT)*, Nov. 18–19, 2015, Russia, Moscow. Moscow, 2015, pp. 77–79. DOI: 10.1109/EnT.2015.9.
33. Alimuradov A. K., Tychkov A. Yu., Ageykin A. V., Churakov P. P., Kvitka Y. S., Zaretskiy A. P. *XX IEEE International Conference on Soft Computing and Measurements*

- (SCM), May 24–26, 2017, Russia, St. Petersburg. Saint-Petersburg, 2017, pp. 610–613. DOI: 10.1109/SCM.2017.7970665.
34. Huang N. E., Zheng Sh., Steven R. L. *Proc. R. Soc. Lond.* 1998, vol. A454, pp. 903–995.
35. Zhaohua Wu., Huang N. E. *Advances in Adaptive Data Analysis*. 2009, vol. 1, no. 1, pp. 1–41.
36. Yeh J.-R., Shieh J.-S., Huang N. E. *Adv. Adapt. Data Anal.* 2010, vol. 2, no. 2, pp. 135–156.
37. Torres M. E., Colominas M. A., Schlotthauer G., Flandrin P. *International Conference on Acoustics, Speech and Signal Processing (ICASSP-11)*. Prague (CZ), 2011, pp. 4144–4147.
38. Colominas M. A., Schlotthauer G., Torres M. E. *Biomedical Signal Processing and Control*. 2014, vol. 14, pp. 19–29.
39. Huang X., Acero A., Hon H.-W. *Prentice Hall*. New Jersey, 2001, 980 p.
40. Fant G. K. *Akusticheskaya teoriya recheobrazovaniya: per. s ang. L. A. Varshavskogo i V. I. Medvedeva; nauch. red. V. S. Grigor'eva* [The acoustic theory of speech production: translation from English language by L.A. Varshavskiy and V.I. Medvedev: scientific editing by V.S. Grigoryev]. Moscow: Nauka, 1964, 284 p.
41. Popov D. I. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* [University proceedings. Volga region. Engineering sciences]. 2017, no. 1 (41), pp. 96–105.
42. Alimuradov A. K. *Measurement techniques*. 2017, vol. 59, no. 12, pp. 1316–1323. DOI 10.1007/s11018-016-0994-1.
43. Alimuradov A. K., Tychkov A. Yu., Frantsuzov M. V., Churakov P. P. *Measurement techniques*. 2015, vol. 58, iss. 9, pp. 965–969.
-

Алимурадов Алан Казанферович

кандидат технических наук,
директор студенческого научно-
производственного бизнес-инкубатора,
Пензенский государственный
университет (Россия, г. Пенза,
ул. Красная, 40)

E-mail: alansapfir@yandex.ru

Alimuradov Alan Kazanferovich

Candidate of engineering sciences,
director of the student research
and production business- incubator,
Penza State University (40 Krasnaya
street, Penza, Russia)

Тычков Александр Юрьевич

кандидат технических наук, заместитель
директора, Научно-исследовательский
институт фундаментальных
и прикладных исследований, Пензенский
государственный университет (Россия,
г. Пенза, ул. Красная, 40)

E-mail: tychkov-a@mail.ru

Tychkov Aleksandr Yur'evich

Candidate of engineering sciences,
deputy director, Research Institute
of Fundamental and Applied Research,
Penza State University (40 Krasnaya
street, Penza, Russia)

Чураков Петр Павлович

доктор технических наук, профессор,
кафедра информационно-измерительной
техники и метрологии, Пензенский
государственный университет (Россия,
г. Пенза, ул. Красная, 40)

E-mail: ivan@pniei.penza.ru

Churakov Petr Pavlovich

Doctor of engineering sciences, professor,
sub-department of information measuring
technologies and metrology, Penza
State University (40 Krasnaya street,
Penza, Russia)

Агейкин Алексей Викторович

младший научный сотрудник,
Научно-исследовательский институт
фундаментальных и прикладных
исследований, Пензенский
государственный университет (Россия,
г. Пенза, ул. Красная, 40)

E-mail: keokushinka@yandex.ru

Ageykin Aleksey Viktorovich

Junior researcher, Research Institute
of Fundamental and Applied Research,
Penza State University (40 Krasnaya
street, Penza, Russia)

УДК 616.89

Алимурадов, А. К.

**Разработка алгоритма обработки речевых сигналов для определе-
ния информативно значимых параметров пограничных психических
расстройств / А. К. Алимурадов, А. Ю. Тычков, П. П. Чураков, А. В. Агей-
кин // Известия высших учебных заведений. Поволжский регион. Техниче-
ские науки. – 2017. – № 3 (43). – С. 4–25. DOI 10.21685/2072-3059-2017-3-1**