

Научная статья

DOI: 10.15593/2224-9397/2022.2.05

УДК 004.934

**А.К. Алимуратов, А.Ю. Тычков, П.П. Чураков, Б.А. Порезанов,  
И.О. Шешкин, К.Е. Платонов, А.В. Баранова, Н.С. Бофанова**

Пензенский государственный университет, Пенза, Россия

## **СПОСОБ ВЫЯВЛЕНИЯ СКРЫТЫХ ИНФОРМАТИВНЫХ ПАРАМЕТРОВ РЕЧИ НА ОСНОВЕ ДЕКОМПОЗИЦИИ НА ЭМПИРИЧЕСКИЕ МОДЫ ДЛЯ ОЦЕНКИ ПСИХОЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА**

**Актуальность и цели:** оценка психоэмоционального состояния человека сегодня является важной социально-значимой проблемой для каждого государства, поскольку напрямую связана с формированием здорового образа жизни населения и профилактикой психических расстройств. Целью исследования является повышение эффективности оценки психоэмоциональных состояний человека по речевым сигналам. Объектом исследования является информативное пространство характеристик речевых сигналов. Предметом исследования являются скрытые информативные параметры речи, релевантные психоэмоциональным состояниям человека. **Материалы и методы:** в работе использовалась уникальная технология частотно-временного анализа для разложения нестационарных данных, не требующая априорной информации об анализируемом сигнале, – декомпозиция на эмпирические моды. Программная реализация способа была выполнена в среде математического моделирования © MatLab (MathWorks). **Результаты:** разработан способ выявления скрытых информативных параметров речи, основанный на равномерном делении исходного речевого сигнала на фрагменты, декомпозиции фрагментов на эмпирические моды, формировании набора информативных (модовых и комбинированных) речевых сигналов. Целью формирования информативных речевых сигналов является расширение информативного пространства характеристик исходного сигнала, необходимых для выявления скрытых информативных параметров. Представлены краткое описание разновидностей методов декомпозиции, их преимущества и недостатки. Подробно описан функционал предлагаемого способа и представлены результаты исследования. **Выводы:** исследование заключалось в сравнении результатов сегментации речь/пауза, полученных посредством анализа параметров новых информативных (модовых и комбинированных) сигналов и исходного речевого сигнала. Из результатов исследования следует, что разработанный способ обеспечивает выявление скрытых информативных параметров и расширение информативного пространства исходного речевого сигнала; уменьшение ошибок 1-го и 2-го рода, демонстрирующее уровень информативности выявленных скрытых параметров и их релевантность задаче сегментации речь/пауза. Применение способа для выявления скрытых информативных параметров речи позволит повысить эффективность оценки психоэмоциональных состояний человека.

**Ключевые слова:** обработка речевых сигналов, информативные параметры речи, декомпозиция на эмпирические моды, психоэмоциональное состояние человека, эмоции.

**A.K. Alimuradov, A.Yu. Tychkov, P.P. Churakov, B.A. Porezanov, I.O.  
Steshkin, K.E. Platonov, A.V. Baranova, N.S. Bofanova**

Penza State University, Penza, Russian Federation

## **EMD-BASED METHOD FOR REVEALING HIDDEN INFORMATIVE SPEECH PARAMETERS TO ASSESS HUMAN PSYCHO-EMOTIONAL STATE**

**Relevance and goals:** An assessment of human psycho-emotional state is a vital social problem worldwide nowadays, being conjoined with healthy lifestyle formation and mental disorder prevention among the population. The aim of the research is to enhance efficiency of assessing human psycho-emotional states by speech signals. The object of the research is the informative space of speech signal characteristics. The subject of the research is hidden informative speech parameters relevant to human psycho-emotional states. **Materials and methods:** A unique time-frequency analysis technology for non-stationary data decomposition, namely, empirical mode decomposition that requires no a priori information about the analyzed signal, has been used. The software implementation of the method has been performed in © MatLab (MathWorks) mathematical modeling environment. **Results:** A method for revealing hidden informative speech parameters based on the uniform splitting of the original speech signal into fragments, the empirical mode decomposition of fragments, and the formation of a set of informative (mode and composite) speech signals has been developed. Informative speech signals are formed to expand the informative space of the original speech signal characteristics, being essential to identify hidden informative parameters. A synopsis of decomposition method types, along with pros and cons thereof have been demonstrated. The functionality of the proposed method has been detailed, and the research outcomes have been reported. **Conclusions:** The results of speech/pause segmentation obtained by analyzing the parameters of new informative (mode and composite) signals and the original speech signal have been compared. It has been evidenced that the developed method provides identifying hidden informative parameters; expanding informative space of the original speech signal; reducing the 1st and 2nd kind errors to indicate the level of information content of the revealed hidden parameters and their relevance to the speech/pause segmentation task. An application of the method for revealing hidden informative speech parameters will improve the efficiency of assessing human psycho-emotional states.

**Keywords:** speech signal processing, informative speech parameters, empirical mode decomposition, human psycho-emotional state, emotions.

### **Введение**

Контроль психоэмоционального состояния населения сегодня является важной социально-значимой проблемой для каждого государства, поскольку напрямую связан с формированием здорового образа жизни населения и профилактикой психоэмоциональных расстройств. Социально-значимые заболевания, являющиеся главной причиной временной нетрудоспособности, инвалидности и смертности населения, напрямую связаны с психоэмоциональным состоянием [1].

Своевременное обнаружение и классификация психоэмоционального состояния человека особенно важно в тех отраслях деятельности,

которые сопряжены с повышенным риском возникновения техногенных и биогенных аварий. В частности, постоянный мониторинг психоэмоционального состояния необходим специалистам, деятельность которых связана с повышенной ответственностью: пилоты, космонавты, военнослужащие, диспетчеры аэропортов, диспетчеры опасных производственных объектов.

В последнее время научные исследования в области обнаружения и классификации психоэмоциональных состояний человека активно поддерживаются международными фондами:

– «ParaStiChaD: Paralinguistic Speech Characteristics in Major Depressive Disorder» (#SCHU2508/8-1) («Paralinguistische Stimmmerkmale in Major Depression»), фонд/организация-грантодатель: DFG (German Research Foundation);

– «Assessing and Enhancing Emotional Competence for Well-Being (ECoWeB) in the Young: A principled, evidence-based, mobile-health approach to prevent mental disorders and promote mental wellbeing (#754657)», фонд/организация-грантодатель: EU Horizon 2020 Research & Innovation Action;

– «Promoting Early Diagnosis of Rett Syndrome through Speech-Language Pathology», (#16430), фонд/организация-грантодатель: Österreichische Nationalbank (OeNB) Jubiläums fonds.

Анализ открытых источников научной литературы в период с 2005 по 2021 г. выявил множество научных работ, посвященных: созданию речевых баз данных эмоциональной речи; исследованию информативных параметров эмоциональной речи; разработке алгоритмов, способов и средств определения параметров речи, отражающих психоэмоциональное состояние человека; проектированию систем автоматического распознавания эмоций человека по речи и др.

Однако, несмотря на большое количество проведенных исследований, ключевая проблема выделения скрытых информативных параметров речевых сигналов, релевантных эмоциональному состоянию человека, требует дальнейшей проработки [2].

В данной статье представлен новый способ выявления скрытых информативных параметров речи на основе декомпозиции на эмпирические моды (ДЭМ) [3] для применения в системах оценки психоэмоционального состояния человека. Статья является результатом научной работы коллектива авторов [4, 5], посвященной исследованию и поиску

скрытых особенностей речевых сигналов, формированию оптимального набора параметров, релевантных естественно выраженным эмоциям человека посредством применения новых адаптивных методов частотно-временного анализа.

### **Декомпозиция на эмпирические моды**

Метод ДЭМ и его модификации широко используются в различных областях, связанных с обработкой и анализом нестационарных сигналов: природные явления [6–8], биотехнологии [9–11], медицина [12–14], анализ данных и машинное обучение [15, 16] и др.

Применение метода ДЭМ и его модификаций в исследованиях эмоциональной составляющей речи началось сравнительно недавно, учитывая, что ДЭМ была разработана в 1998 г. [17–20].

В результате разложения методом ДЭМ исходный сигнал представляет собой набор внутренних функций, называемых модами или эмпирическими модами (ЭМ). Отличительной особенностью процесса разложения методом ДЭМ является то, что базисные функции, используемые при разложении, извлекаются непосредственно из исходного сигнала. Таким образом, математическая модель ЭМ не задаётся заранее, а вычисляется в процессе отсеивания. Такой способ разложения обеспечивает адаптивность базисным функциям, позволяющую учитывать локальные особенности (экстремумы и нули функции сигнала) и особенности внутренней структуры исходного сигнала (скрытые модуляции, области концентрации энергии и т.п.).

Подробный анализ известных методов ДЭМ, применяемых для анализа сигналов естественной природы, выявил, что наиболее адаптивными к нестационарной речи являются: множественная ДЭМ (МДЭМ) [21] и улучшенная полная МДЭМ с адаптивным шумом (ПМДЭМАШ) [22].

С точки зрения отсеивания ЭМ методы МДЭМ и улучшенной ПМДЭМАШ аналогичны. Добавление контролируемого шума малой амплитуды на каждом этапе отсеивания (для создания новых экстремумов) позволяет избежать известных недостатков декомпозиции (смешивание мод, неполнота декомпозиции, остаточный шум, неинформативные «паразитные» моды).

$$x_j(n) = x(n) + w_j(n), \quad (1)$$

где  $x_j(n)$  – зашумленные сигналы;  $n$  – дискретные отсчеты времени;  $x(n)$  – исходный речевой сигнал;  $w_j(n)$  – белый шум малой амплитуды;  $j$  – количество реализаций белого шума,  $j = 1, 2, \dots, J$ ;

$$x_j(n) = \sum_{i=1}^I IMF_{ji}(n) + r_{jI}(n), \quad (2)$$

$$IMF_i(n) = \sum_{j=1}^J \frac{IMF_{ji}(n)}{J}, \quad (3)$$

$$r_i(n) = \sum_{j=1}^J \frac{r_{ji}(n)}{J}, \quad (4)$$

где  $IMF(n)$  – ЭМ;  $r(n)$  – конечный неделимый остаток;  $i = 1, 2, \dots, I$  – количество ЭМ.

Важными параметрами настройки методов МДЭМ и улучшенной ПМДЭМАШ, влияющими на результат разложения, являются:  $Nstd$  – стандартное отклонение амплитуды добавляемого белого шума (в процентном отношении от исходного сигнала),  $NR$  – количество реализаций (для дальнейшего усреднения),  $MaxIter$  – количество итераций отсеивания ЭМ,  $SNRFlag$  – отношение сигнал/шум для каждого этапа разложения (только для метода улучшенной ПМДЭМАШ).

### Описание способа

На рис. 1 представлен пример, иллюстрирующий функционал предлагаемого способа (блоки 1–4). В основе способа заложены следующие этапы обработки: сегментация исходного сигнала на фрагменты одинаковой длительности; разложение фрагментов на ЭМ; формирование набора информативных (модовых и комбинированных) речевых сигналов. Рассмотрим подробнее этапы предлагаемого способа обработки речевых сигналов.

**Блок 1.** Процесс линейного разделения исходного речевого сигнала  $x(n)$  на равномерные отрезки, записываемые в отдельные переменные  $x_s(n)$ , называется фрагментированием. В предлагаемом способе каждый фрагмент исходного сигнала обрабатывается как отдельный кратковременный речевой сигнал. Линейное разделение речевого сигнала на фрагменты осуществляется по следующим формулам:

$$S = \frac{x(n)}{L}, \quad (5)$$

$$x_{s+1}(n) = x[(s \cdot L) + 1 : (s + 1)L], \quad (6)$$

где  $s$  – номер фрагмента,  $s = 0, 1, 2, \dots, S$ ;  $S$  – количество фрагментов в исходном речевом сигнале;  $L$  – количество дискретных отсчетов времени в одном фрагменте.

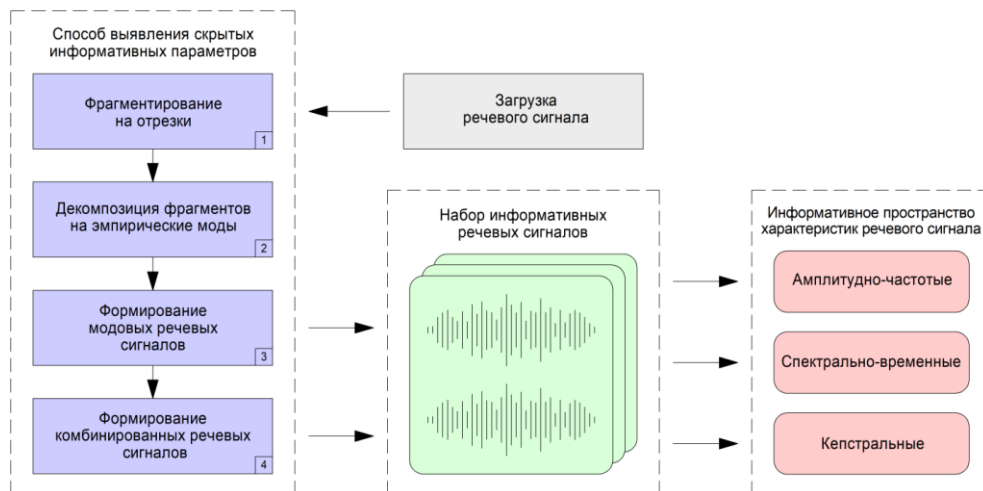


Рис. 1. Функционал предлагаемого способа для выявления скрытых информативных параметров речи

Длительность фрагментов влияет на последующие этапы обработки речевого сигнала в предлагаемом способе. В работе [23] авторами представлены результаты исследования, посвященного влиянию длительности анализируемых сигналов на частотно-избирательные свойства методов декомпозиции. Оптимальной длительностью, при которой частотно-избирательные свойства декомпозиции обеспечивают необходимую и достаточную полноту разложения, является 50 мс и выше.

**Блок 2.** Наиболее адаптивными к нестационарной речи являются методы МДЭМ и улучшенной ПМДЭМАШ. Особенностью метода улучшенной ПМДЭМАШ является локальное разложение белого шума на ЭМ с целью добавления полученных шумовых мод в качестве контролируемого шума на каждом этапе декомпозиции. Данная особенность обеспечивает методу улучшенной ПМДЭМАШ полноту разложения в сравнении с методом МДЭМ.

**Блоки 3, 4.** Целью формирования набора информативных (модовых и комбинированных) речевых сигналов является расширение информативного пространства амплитудно-частотных, спектрально-временных и кепстральных характеристик исходного сигнала, необходимых для выявления скрытых информативных параметров. Каждый

информативный (модовый и комбинированный) сигнал содержит в себе скрытые особенности внутренней структуры исходного речевого сигнала (скрытые модуляции, области концентрации энергии и т.п.).

В соответствии с результатом декомпозиции каждый фрагмент исходного речевого сигнала представлен набором ЭМ. Формирование модовых сигналов представляет собой процесс объединения ЭМ фрагментов исходного речевого сигнала:

$$xmode_i(n) = \sum_{s=1}^S IMF_{s,i}[(s \cdot L) + 1 : (s + 1) \cdot L], \quad (7)$$

где  $xmode_i(n)$  – модовый речевой сигнал.

Формирование комбинированных сигналов осуществляется посредством суммирования модовых речевых сигналов по следующей формуле:

$$xcomb_k(n) = \sum_i^I a_i \cdot xmode_i(n), \quad (8)$$

где  $a_i$  – коэффициенты, принимающие два значения (0 или 1) и определяющие, какие модовые речевые сигналы участвуют в формировании комбинированных сигналов;  $I$  – количество модовых сигналов;  $k$  – номер комбинированного сигнала (количество комбинированных сигналов ограничено выражением  $2^I$ ).

Формирование набора информативных сигналов осуществляется посредством объединения модовых и комбинированных речевых сигналов:  $xinfo_m(n) = xcomb_k(n) + xmode_i(n)$ , где  $m$  – количество информативных сигналов,  $m = k + i$ .

### **Исследование способа**

Суть исследования заключается в сравнении результатов сегментации речь/пауза, полученных посредством анализа параметров набора информативных сигналов и исходного речевого сигнала. Сегментация осуществлялась способами, основанными на анализе следующих информативных параметров: количество пересечений функции сигнала через нулевую ось (Zero-Crossing Rate, ZCR); кратковременная энергия (Short Time Energy, STE); одномерное расстояние Махаланобиса (One Dimensional Mahalanobis Distance, ODMD). Подробно математическое описание и результаты детализированных исследований данных способов представлены в работе [24].

Эффективность сегментации оценивалась в соответствии с полученными значениями ошибок 1-го и 2-го рода. Ошибкой 1-го рода считалась ситуация, когда фрагменту речи присваивался статус «пауза». Ошибкой 2-го рода считалась ситуация, когда фрагменту паузы присваи-

вался статус «речь». Основной задачей сегментации считалось обнаружение фрагментов речи. Ошибки определялись по результатам сопоставления полученных данных сегментации с данными сегментации, осуществленной вручную. Для разложения исходного сигнала методами МДЭМ и улучшенной ПМДЭМАШ в рамках исследования изменялись следующие параметры:  $NR = 50$ ,  $MaxIter = 100$ . Параметры  $Nstd = 20\%$  и  $SNRFlag = 2$  оставались неизменными. Реализация методов декомпозиции в программе MatLab была заимствована из работы [25].

Для обеспечения полноты разложения продолжительность анализируемых фрагментов речи была установлена равной 300 мс. Для формирования модовых речевых сигналов использовалось восемь информативных ЭМ, полученных для каждого фрагмента.

Для исследования предлагаемого способа сформирована речевая база данных. Речевые сигналы длительностью не более 10 с были зарегистрированы посредством специализированных методики и технических средств. Подготовленные дикторы в количестве 20 человек воспроизводили речь на русском языке, содержащую публицистический текст (30 записей), текст из литературного произведения (30 записей) и счёт чисел от 0 до 100 (10 записей). Общее количество зарегистрированных речевых сигналов – 1400.

В табл. 1 представлены результаты исследования – усредненные значения ошибок 1-го и 2-го рода, полученные по результатам сегментации 1400 речевых сигналов способами ZCR, STE и ODMD.

Таблица 1

Усредненные значения ошибок 1-го и 2-го рода, полученные по результатам сегментации с помощью способов на основе анализа ZCR, STE и ODMD

Способ сегментации речь/пауза	Анализ исходного речевого сигнала		Анализ информативных речевых сигналов			
	Ошибка 1-го рода, %	Ошибка 2-го рода, %	МДЭМ		Улучшенная ПМДЭМАШ	
			Ошибка 1-го рода, %		Ошибка 2-го рода, %	
Способ на основе анализа ODMD	22,197	0,888	2,517	1,065	3,661	1,065
Способ на основе ана- лиза ZCR	24,256	8,881	7,322	0,888	5,263	8,703
Способ на основе ана- лиза STE	8,466	3,907	2,059	1,598	2,059	1,776



## **Анализ результатов исследования**

В соответствии с данными в табл. 1 можно отметить следующие закономерности:

1) повышение эффективности сегментации вокализованных и невокализованных фрагментов речи (уменьшение ошибки 1-го рода) для всех способов:

– на основе анализа ODMD на 19,68 % для МДЭМ и 18,536 % для улучшенной ПМДЭМАШ;

– на основе анализа ZCR на 16,934 % для МДЭМ и 18,993 % для улучшенной ПМДЭМАШ;

– на основе анализа STE на 6,407 % для МДЭМ и улучшенной ПМДЭМАШ;

2) повышение эффективности сегментации фрагментов пауз (уменьшение ошибки 2-ого) для способов:

– на основе анализа ZCR на 7,993 % для МДЭМ и 0,178 % для улучшенной ПМДЭМАШ;

– на основе анализа STE на 2,309 % для МДЭМ и 2,131 % для улучшенной ПМДЭМАШ;

3) для способа на основе анализа ODMD отмечается незначительное увеличение ошибки 2-го рода на 0,177 %. Однако необходимо отметить, что эффективность сегментации фрагментов пауз для всех способов изначально была приемлемой.

Также необходимо подчеркнуть влияние методов декомпозиции, используемых для формирования информативных речевых сигналов, на результат сегментации речь/пауза:

– при анализе ODMD метод МДЭМ обеспечивает улучшение на 1,144 % для ошибки 1-го рода;

– при анализе ZCR метод улучшенной ПМДЭМАШ обеспечивает улучшение на 2,059 % для ошибки 1-го рода и ухудшение на 7,815 % для ошибки 2-го рода;

– при анализе STE метод МДЭМ обеспечивает улучшение на 0,178 % для ошибки 2-го рода.

В целом оба метода декомпозиции обеспечивают достаточно приемлемые результаты сегментации.

На рис. 2–7 представлен пример сегментации одного речевого сигнала из сформированной речевой базы данных. Линией зеленого цвета отмечен результат сегментации, выполненной вручную. Линией

черного цвета отмечен результат сегментации, выполненной посредством анализа параметров исходного речевого сигнала. Линией красного цвета отмечен результат сегментации, выполненной посредством анализа параметров информативного речевого сигнала.

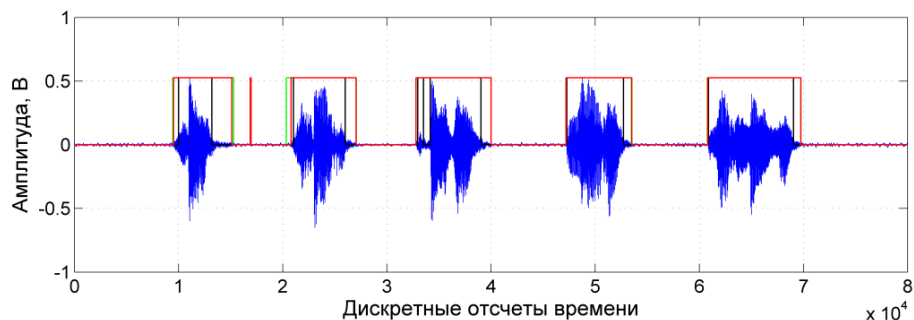


Рис. 2. Результат сегментации способом на основе анализа ODMD исходного и информативного речевого сигнала № 209 (полученного методом МДЭМ)

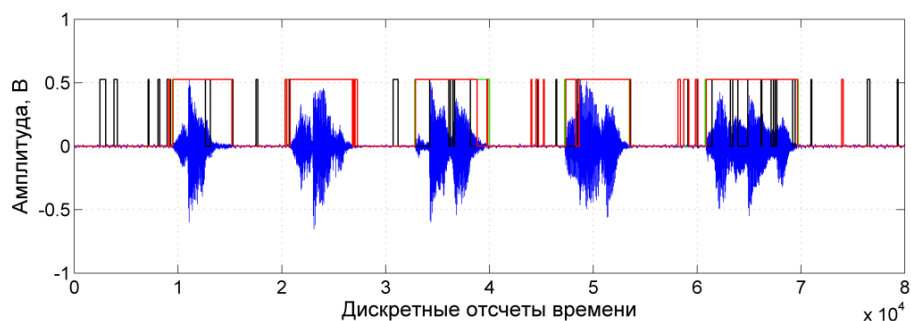


Рис. 3. Результат сегментации способом на основе анализа ZCR исходного и информативного речевого сигнала № 80 (полученного методом МДЭМ)

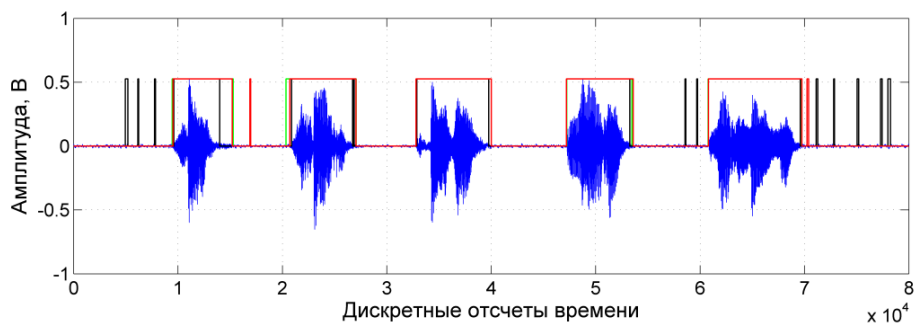


Рис. 4. Результат сегментации способом на основе анализа STE исходного и информативного речевого сигнала № 145 (полученного методом МДЭМ)

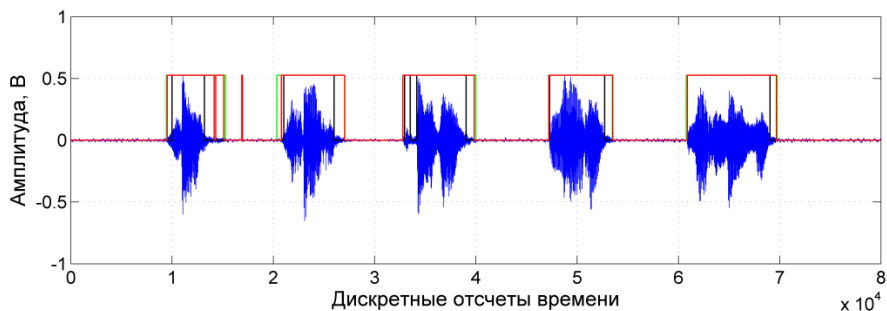


Рис. 5. Результат сегментации способом на основе анализа ODMD исходного и информативного речевого сигнала № 209 (полученного методом улучшенной ПМДЭМАШ)

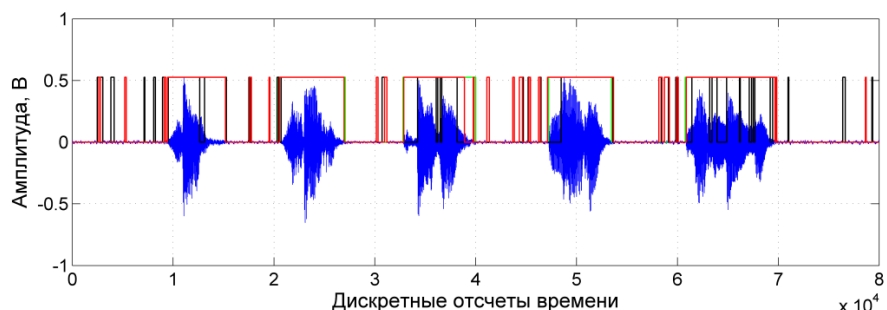


Рис. 6. Результат сегментации способом на основе анализа ZCR исходного и информативного речевого сигнала № 79 (полученного методом улучшенной ПМДЭМАШ)

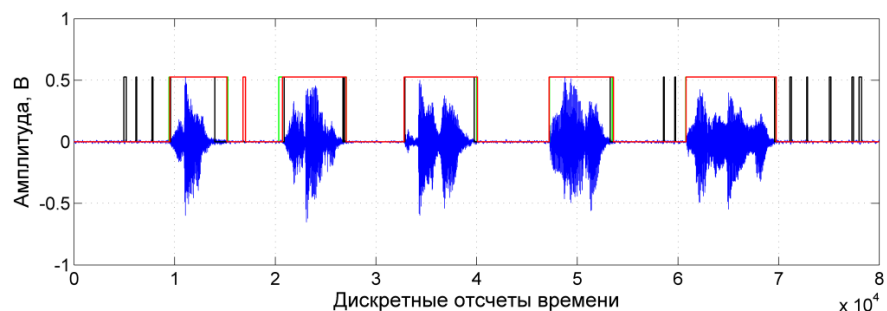


Рис. 7. Результат сегментации способом на основе анализа STE исходного и информативного речевого сигнала № 209 (полученного методом улучшенной ПМДЭМАШ)

В табл. 2 представлены данные об информативных сигналах, обеспечивающих повышение эффективности сегментации речевого сигнала из сформированной речевой базы данных (см. рис. 2–7).

Таблица 2

Данные об информативных сигналах

Номер информативного сигнала	Значение коэффициентов	Модовые речевые сигналы
$xinfo_{209}(n)$	$a_1 = a_2 = a_4 = a_8 = 1,$ $a_3 = a_5 = a_6 = a_7 = 0$	$xmode_1(n) + xmode_2(n) + xmode_4(n) +$ $+xmode_8(n)$
$xinfo_{80}(n)$	$a_2 = a_4 = 1,$ $a_1 = a_3 = a_5 = a_6 = a_7 = a_8 = 0$	$xmode_2(n) + xmode_4(n)$
$xinfo_{145}(n)$	$a_1 = a_4 = a_8 = 1,$ $a_2 = a_3 = a_5 = a_6 = a_7 = 0$	$xmode_1(n) + xmode_4(n) + xmode_8(n)$
$xinfo_{79}(n)$	$a_2 = a_5 = a_6 = a_7 = a_8 = 1,$ $a_1 = a_3 = a_4 = 0$	$xmode_2(n) + xmode_5(n) + xmode_6(n) +$ $+xmode_7(n) + xmode_8(n)$

**Заключение**

В соответствии с результатами анализа исследований можно сделать следующие выводы:

1. Повышение эффективности сегментации речь/пауза подтверждает расширение информативного пространства характеристик исходного речевого сигнала и выявление скрытых информативных параметров предлагаемым способом.

2. Уменьшение ошибок 1-го и 2-го рода демонстрирует уровень информативности выявленных скрытых параметров и их релевантность задаче сегментации речь/пауза.

3. Достигнутая точность сегментации позволит повысить эффективность оценки психоэмоциональных состояний человека посредством анализа скорости, ускорения и энтропии распределения фрагментов речи и пауз, а также средней продолжительности пауз в слитной речи.

**Библиографический список**

1. Schuller B.W., Batliner A.M. Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing. – New York: Wiley, 2013. – 344 p.

2. Schuller B.W. Speech emotion recognition // Communications of the ACM. – 2018. – Vol. 61. – № 5. – P. 90–99.

3. Huang, N.E., Zheng Sh., Steven R.L. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis // Proceedings of the Royal Society of London. – 1998. – A 454. – P. 903–995.

4. Алимуратов А.К. Повышение эффективности сегментации речевых сигналов на основе энергетического оператора Тигера // Измерение. Мониторинг. Управление. Контроль. – 2021. – № 3 (37). – С. 80–92.

5. Способ повышения эффективности сегментации речь/пауза на основе метода декомпозиции на эмпирические моды / А.К. Алимуратов, А.Ю. Тычков, П.П. Чураков, А.В. Агейкин, А.В. Кузьмин, М.А. Митрохин, И.А. Чернов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2021. – № 2 (58). – С. 24–43.

6. Mi X., Liu H., Li Y. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine // Energy Conversion and Management. – 2019. – Vol. 180. – P. 196–205.

7. The impact of global warming on ENSO variability in climate records / Z. Wu, E. Schneider, Z. Hu, L. Cao // Tech. report, Cent. Ocean. Stud. – 2001. – Vol. 110. – 25 p.

8. Lee T., Ouarda T.B.M.J. Multivariate Nonstationary Oscillation Simulation of Climate Indices With Empirical Mode Decomposition // Water Resources Research. – 2019. – Vol. 55. – № 6. – P. 5033–5052.

9. Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: A comparison / D. Labate, F. La Foresta, G. Occhiuto, F.C. Morabito, A. Lay-Ekuakille, P. Vergallo // IEEE Sensors Journal. – 2013. – Vol. 13, № 7. – P. 2666–2674.

10. The removal of wall components in doppler ultrasound signals by using the empirical mode decomposition algorithm / Y. Zhang, Y. Gao, L. Wang, J. Chen, X. Shi // IEEE Transactions on Biomedical Engineering. – 2007. – Vol. 54, № 9. – P. 1631–1642.

11. EMG signal filtering based on Empirical Mode Decomposition / A.O. Andrade, S. Nasuto, P. Kyberd, C.M. Sweeney-Reed, F.R. Van Kanijn // Biomedical Signal Processing and Control. – 2006. – Vol. 1, № 1. – P. 44–55.

12. Schlotthauer G., Torres M.E., Rufiner H.L. A new algorithm for instantaneous F0 speech extraction based on ensemble empirical mode decomposition // 2009 17th European Signal Processing Conference (Aug. 24–28, 2009). – UK, Glasgow. – P. 2347–2351.

13. Bouzid A., Ellouze N. Empirical mode decomposition of voiced speech signal // First International Symposium on Control, Communications and Signal Processing (March 21–24, 2004). – Tunisia, Hammamet. – P. 603–606.

14. Sethu V., Ambikairajah E., Epps J. Empirical mode decomposition based weighted frequency feature for speech-based emotion classification // 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (31 March–4 April, 2008). – Las Vegas, NV, USA. – P. 5017–5020.

15. Ali M., Prasad R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition // Renewable and Sustainable Energy Reviews. – 2019. – Vol. 104. – P. 281–295.

16. Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation / R. Prasad, M. Ali, P. Kwan, H. Khan // Applied Energy. – 2019. – Vol. 236. – P. 778–792.

17. Chu Y.Y., Xiong W.H., Chen Wei. Speech Emotion Recognition Based on EMD in Noisy Environments // Advanced Materials Research. – 2013. – Vol. 831. – P.460–466.

18. Automatic speech emotion recognition using an optimal combination of features based on EMD-TKEO / L. Kerkeni, Y. Serrestou, K. Raoof, M. Mbarki, M.A. Mahjoub, C. Cleder // Speech Communication. – 2019. – Vol. 114. – P. 22–35.

19. Emotion recognition based on EMD-Wavelet analysis of speech signals / C. Shahnaz, S. Sultana, S.A. Fattah, R. H.M. Rafi, I. Ahmmed, W.-P. Zhu, M.O. Ahmad // 2015 IEEE International Conference on Digital Signal Processing (DSP) (July 21–24, 2015). – Singapore. – P. 307–310.

20. EMD-TEO Based Speech Emotion Recognition / X. Li, Xin Li, X. Zheng, D. Zhang // Life System Modeling and Intelligent Computing. – 2010. – P. 180–189.

21. Zhaohua W., Huang N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method // Advances in Adaptive Data Analysis. – 2009. – № 1 (1). – P. 1–41.

22. Colominasa M.A., Schlotthauera G., Torres M.E. Improved complete ensemble EMD: a suitable tool for biomedical signal processing // Biomed. Signal Proces. – 2014. – Vol. 14. – P. 19–29.

23. Алимуратов А.К., Фокина Е.А., Журина А.Е. Исследование влияния длительности анализируемых речевых сигналов на частотно-избирательные свойства декомпозиции на эмпирические моды // Новые информационные технологии и системы: сб. науч. ст. XVI Междунар.

науч.-техн. конф. (Пенза, 27–29 ноября 2019 г.). – Пенза: Изд-во ПГУ, 2019. – С. 201–205.

24. Алгоритм сегментации речь/пауза на основе декомпозиции на эмпирические моды и одномерного расстояния Махаланобиса / А.К. Алимуратов, А.Ю. Тычков, П.П. Чураков, А.В. Агейкин, А.П. Кулешов, И.А. Чернов // Труды МФТИ. – 2021. – Т. 13, № 3 (51). – С. 4–22.

25. National University of Entre Ríos. The Laboratory of Signals and Nonlinear Dynamics, Faculty of Engineering [Электронный ресурс]. – URL: <http://www.bioingenieria.edu.ar/grupos/ldnlys> (дата обращения: 01.05.2021).

## Reference

1. Schuller B.W., Batliner A.M. Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing. New York: Wiley, 2013, 344 p.

2. Schuller B.W. Speech emotion recognition. *Communications of the ACM*, 2018, vol. 61, no. 5, pp. 90-99.

3. Huang N.E., Zheng Sh., Steven R.L. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London*, 1998, A 454, pp. 903-995.

4. Alimuradov A.K. Povysheniye effektivnosti segmentatsii rechevykh signalov na osnove energeticheskogo operatora Tigera [Enhancement of speech signal segmentation using teager energy operator]. *Izmereniye. Monitoring. Upravleniye. Kontrol'*, 2021, no. 3(37), pp. 80-92.

5. Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Ageykin A.V., Kuzmin A.V., Mitrokhin M.A., Chernov I.A. Sposob povysheniya effektivnosti segmentatsii rech'/pauza na osnove metoda dekompozitsii na empiricheskiye mody [EMD-based method to improve the efficiency of speech/pause segmentation]. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskiye nauki*, 2021, no. 2(58), pp. 24-43.

6. Mi X., Liu H., Li Y. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Conversion and Management*, 2019, vol. 180, no. 1, pp. 196-205.

7. Wu Z., Schneider E.K., Hu Z.-Z., Cao L. The impact of global warming on ENSO variability in climate records. *COLA Technical Report, CTR 110*, 2001.

8. Lee T., Ouarda T.B.M.J. Multivariate nonstationary oscillation simulation of climate indices with empirical mode decomposition. *Water Resources Research*, 2019, vol. 55, no. 6, pp. 5033-5052.

9. Labate D., La Foresta F., Occhiuto G., Morabito F.C., Lay-Ekuakille A., Vergallo P. Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: A comparison. *IEEE Sensors Journal*, 2013, vol. 13, no. 7, pp. 2666-2674.

10. Zhang Y., Gao Y., Wang L., Chen J., Shi X. The removal of wall components in doppler ultrasound signals by using the empirical mode decomposition algorithm. *IEEE Transactions on Biomedical Engineering*, 2007, vol. 54, no. 9, pp. 1631-1642.

11. Andrade A.O., Nasuto S., Kyberd P., Sweeney-Reed C.M., Van Kanijn F.R. EMG signal filtering based on Empirical Mode Decomposition. *Biomedical Signal Processing and Control*, 2006, vol. 1, no. 1, pp. 44-55.

12. Schlotthauer G., Torres M.E., Rufiner H.L. A new algorithm for instantaneous F0 speech extraction based on ensemble empirical mode decomposition. *2009 17th European Signal Processing Conference (August 24-28, 2009)*. Glasgow, UK, pp. 2347-2351.

13. Bouzid A., Ellouze N. Empirical mode decomposition of voiced speech signal. *First International Symposium on Control, Communications and Signal Processing (March 21-24, 2004)*. Hammamet, Tunisia, pp. 603-606.

14. Sethu V., Ambikairajah E., Epps J. Empirical mode decomposition based weighted frequency feature for speech-based emotion classification. *2008 IEEE International Conference on Acoustics, Speech and Signal Processing (March 31-April 4, 2008)*. Las Vegas, NV, USA, pp. 5017-5020.

15. Ali M., Prasad R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. *Renewable and Sustainable Energy Reviews*, 2019, vol. 104(C), pp. 281-295.

16. Prasad R., Ali M., Kwan P., Khan H. Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation. *Applied Energy*, 2019, vol. 236(C), pp. 778-792.

17. Chu Y.Y., Xiong W.H., Chen W. Speech emotion recognition based on EMD in noisy environments. *Advanced Materials Research*, 2013, vol. 831, pp. 460-464.



18. Kerkeni L., Serrestou Y., Raouf K., Mbarki M., Mahjoub M.A., Cleder C. Automatic speech emotion recognition using an optimal combination of features based on EMD-TKEO. *Speech Communication*, 2019, vol. 114, pp. 22-35.

19. Shahnaz C., Sultana S., Fattah S.A., Rafi R.H.M., Ahmmed I., Zhu W.-P., Ahmad M.O. Emotion recognition based on EMD-Wavelet analysis of speech signals. *2015 IEEE International Conference on Digital Signal Processing (DSP) (July 21-24, 2015)*, Singapore, pp. 307-310.

20. Li X., Li X., Zheng X., Zhang D. EMD-TEO Based Speech Emotion Recognition. In *Life System Modeling and Intelligent Computing – Lecture Notes in Computer Science*, Li K., Fei K., Jia L., Irwin G., Eds. Berlin: Springer/Heidelberg, 2010, vol. 6329, pp. 180-189.

21. Zhaohua W., Huang N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 2009, no. 1(1), pp. 1-41.

22. Colominasa M.A., Schlotthauera G., Torres M. E. Improved complete ensemble EMD: a suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 2014, vol. 14, pp. 19-29.

23. Alimuradov A.K., Fokina E.A., Zhurina A.E. Issledovaniye vliyaniya dlitel'nosti analiziruyemykh rechevykh signalov na chastotno-izbiratel'nyye svoystva dekompozitsii na empiricheskiye mody [Investigation of the effect of the analyzed speech signal duration on the frequency-selective properties of empirical mode decomposition]. *Novye informatsionnye tekhnologii i sistemy. Sbornik nauchnykh statei XVI Mezhdunarodnoi nauchno-tekhnicheskoi konferentsii (November 27-29, 2019)*. Penza, Russia, pp. 201-205.

24. Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Ageykin A.V., Kuleshov A.P., Chernov I.A. Algoritm segmentatsii rech'/pauza na osnove dekompozitsii na empiricheskiye mody i odnomernogo rasstoyaniya Makhalanobisa [Speech/pause segmentation algorithm based on empirical mode decomposition and one-dimensional Mahalanobis distance]. *Trudy Moskovskogo fiziko-tekhnicheskogo instituta*, 2021, vol. 13, no. 3(51), pp. 4-22.

25. National University of Entre Ríos. The Laboratory of Signals and Nonlinear Dynamics, Faculty of Engineering. [Online] Available: <http://www.bioingenieria.edu.ar/grupos/ldnlys> (accessed 01 March 2022).

### **Сведения об авторах**

**Алимуратов Алан Казанферович** (Пенза, Россия) – кандидат технических наук, директор студенческого научно-производственного бизнес-инкубатора Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: alansapfir@yandex.ru).

**Тычков Александр Юрьевич** (Пенза, Россия) – доктор технических наук, профессор кафедры «Радиотехника и радиоэлектронные системы» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: tychkov-a@mail.ru).

**Чураков Пётр Павлович** (Пенза, Россия) – доктор технических наук, профессор, профессор кафедры «Информационно-измерительная техника и метрология» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: churakov-pp@mail.ru).

**Порезанов Богдан Андреевич** (Пенза, Россия) – студент (специалитет) кафедры «Радиотехника и радиоэлектронные системы» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: bogdan.porezanov@yandex.ru).

**Стешкин Илья Олегович** (Пенза, Россия) – студент (специалитет) кафедры «Радиотехника и радиоэлектронные системы» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: ilya\_steshkin@mail.ru).

**Платонов Кирилл Егорович** (Пенза, Россия) – студент (специалитет) кафедры «Радиотехника и радиоэлектронные системы» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: platonov.1408@mail.ru).

**Баранова Анастасия Валерьевна** (Пенза, Россия) – студент (магистрант) кафедры «Документоведение и архивоведение» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: anastasiyabar@mail.ru).

**Бофанова Наталья Сергеевна** (Пенза, Россия) – кандидат медицинских наук, доцент кафедры «Неврология, нейрохирургия и психиатрия» Пензенского государственного университета (440026, Пенза, ул. Красная, 40, e-mail: neurology-pgu@mail.ru).

## **About the authors**

**Alan K. Alimuradov** (Penza, Russian Federation) – Ph. D. in Technical Sciences, Director of student research and production business incubator Penza State University (440026, Penza, 40, Krasnaya str., e-mail: alansapfir@yandex.ru).

**Alexander Yu. Tychkov** (Penza, Russian Federation) – Doctor of Technical Sciences, Professor of the department of radio engineering and radio electronic systems Penza State University (440026, Penza, 40, Krasnaya str., e-mail: tychkov-a@mail.ru).

**Pyotr P. Churakov** (Penza, Russian Federation) – Doctor of Technical Sciences, Professor of the department of information and measuring equipment and metrology Penza State University (440026, Penza, 40, Krasnaya str., e-mail: churakov-pp@mail.ru).

**Bogdan A. Porezanov** (Penza, Russian Federation) – Student (specialist) of department of radio engineering and radioelectronic systems Penza State University (440026, Penza, 40, Krasnaya str., e-mail: bogdan.porezanov@yandex.ru).

**Ilya O. Steshkin** (Penza, Russian Federation) – Student (specialist) of department of radio engineering and radioelectronic systems Penza State University (440026, Penza, 40, Krasnaya str., e-mail: ilya\_steshkin@mail.ru).

**Kirill E. Platonov** (Penza, Russian Federation) – Student (specialist) of department of radio engineering and radioelectronic systems Penza State University (440026, Penza, 40, Krasnaya str., e-mail: platonov.1408@mail.ru).

**Anastasia V. Baranova** (Penza, Russian Federation) – Student (undergraduate) of department of documentation and archiving Penza State University (440026, Penza, 40, Krasnaya str., e-mail: anastasiyabar@mail.ru).

**Natalia S. Bofanova** (Penza, Russian Federation) – Candidate of Medical Sciences, Associate Professor of the Department neurology, neurosurgery and psychiatry Penza State University (440026, Penza, 40, Krasnaya str., e-mail: neurology-pgu@mail.ru).

Поступила: 10.03.2022 Одобрена: 29.04.2022 Принята к публикации: 12.09.2022

**Финансирование.** Научные исследования выполнялись при финансовой поддержке Совета по грантам Президента РФ, проект «Исследование скрытых паттернов речевых сигналов и разработка способов обнаружения и классификации естественно выраженных психоэмоциональных состояний человека», № МД-1066.2022.4.

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов по отношению к статье.

**Вклад авторов.** Все авторы сделали эквивалентный вклад в подготовку публикации.

Просьба ссылаться на эту статью в русскоязычных источниках следующим образом:

Способ выявления скрытых информативных параметров речи на основе декомпозиции на эмпирические моды для оценки психоэмоционального состояния человека / А.К. Алимуратов, А.Ю. Тычков, П.П. Чураков, Б.А. Порезанов, И.О. Стешкин, К.Е. Платонов, А.В. Баранова, Н.С. Бофанова // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. – 2022. – № 42. – С. 91–110. DOI: 10.15593/2224-9397/2022.2.05

Please cite this article in English as:

Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Porezanov B.A., Steshkin I.O., Platonov K.E., Baranova A.V., Bofanova N.S. Emd-based method for revealing hidden informative speech parameters to assess human psycho-emotional state. *Perm national research polytechnic university bulletin. Electrotechnics, information technologies, control systems*, 2022, no. 42, pp. 91-110. DOI: 10.15593/2224-9397/2022.2.05