



## ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

---

А.К. Алимуратов, А.Ю., Тычков, П.П. Чураков, Д.С. Дудников

### НОВЫЙ ПОДХОД К ВЫЯВЛЕНИЮ СКРЫТЫХ ИНФОРМАТИВНЫХ ПАРАМЕТРОВ РЕЧЕВЫХ СИГНАЛОВ НА ОСНОВЕ МЕТОДА ДЕКОМПОЗИЦИИ НА ЭМПИРИЧЕСКИЕ МОДЫ

(Пензенский государственный университет)

Контроль психоэмоционального состояния населения на сегодняшний день 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].

Анализ открытых источников научной литературы в период с 2005 по 2021 гг. 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 автоматическому распознаванию эмоций человека по речи и др.

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

В данной статье представлен новый подход к выявлению 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], посвященной исследованию и поиску скрытых особенностей речевых сигналов, релевантных естественно выраженным эмоциям человека. Научные исследования выполняются при финансовой поддержке Совета по грантам Президента РФ, проект, № МД-1066.2022.4.

На рисунке 1 представлен пример, иллюстрирующий функционал предлагаемого подхода (блоки 1-4), который основан на равномерном делении исходного речевого сигнала на фрагменты, декомпозиции фрагментов на ЭМ, фор-



мировании набора информативных (модовых и комбинированных) речевых сигналов.

Фрагментирование представляет собой процесс линейного разделения исходного речевого сигнала  $x(n)$  на отрезки одинаковой длительности, которые записываются в отдельные переменные  $x_s(n)$ :

$$S = \frac{N}{L},$$

где  $S$  - количество фрагментов в исходном речевом сигнале;  $N$  - количество дискретных отсчетов времени в исходном сигнале;  $L$  - количество дискретных отсчетов времени в одном фрагменте.

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

где  $s=0, 1, 2, \dots S$  - номер фрагмента.

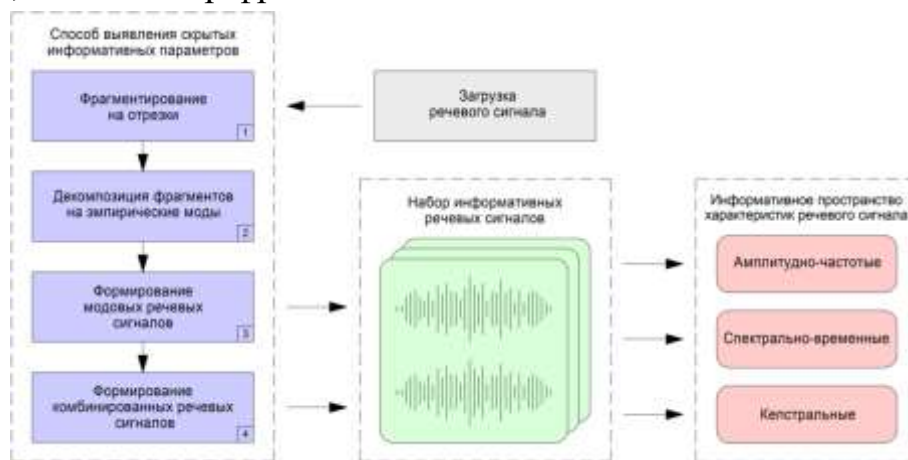


Рис. 1. Функционал предлагаемого подхода для выявления скрытых информативных параметров речи

Наиболее адаптивными к нестационарной речи являются методы МДЭМ [5] и улучшенной ПМДЭМАШ [6]. Отличительной особенностью метода улучшенной ПМДЭМАШ от метода МДЭМ является возможность локального разложения белого шума на шумовые ЭМ параллельно с разложением исходного сигнала. Использование шумовых мод в качестве добавляемого контролируемого белого шума на каждом этапе декомпозиции обеспечивает полноту разложения.

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 информативных ЭМ фрагментов исходного речевого сигнала:



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

где  $xmode_i(n)$  - модовый речевой сигнал;  $i = 1, 2, \dots, I$  - номер ЭМ для каждого фрагмента.

Количество сформированных модовых речевых сигналов зависит от количества используемых информативных ЭМ, полученных для каждого фрагмента.

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

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

где  $a_i$  - коэффициенты, определяющие участие модовых речевых сигналов в формировании комбинированных сигналов (принимают два значения: 0 и 1);  $I$  - 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  $m=k+i$  - количество информативных сигналов.

Суть исследования предлагаемого нового подхода 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).

В таблице 1 представлены результаты исследования - усредненные значения ошибок 1-го и 2-го рода, полученные по результатам сегментации речевых сигналов с помощью способов на основе анализа 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, %	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 анализа ZCR	24,256	8,881	7,322	0,888	5,263	8,703
Способ на основе анализа STE	8,466	3,907	2,059	1,598	2,059	1,776

В соответствии с данными в таблице 1 можно отметить следующие закономерности:



Повышение эффективности сегментации вокализованных и невокализованных фрагментов речи (уменьшение ошибки 1-ого) 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;

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;

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

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 новым подходом.
- 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.

### Литература

1. Schuller B.W. Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing / B.W. Schuller, A.M. Batliner. - New York: Wiley, 2013. - 344 p.
2. Schuller B.W. Speech emotion recognition / B.W. Schuller // Communications of the ACM. - 2018. - Vol. 61. - № 5. - P. 90-99.
3. Huang, N.E. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis / N.E. Huang, Sh. Zheng, R.L. Steven // 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 / А.К. Алимуратов // Измерение. Мониторинг. Управление. Контроль. - 2021. - № 3 (37). - С. 80-92.



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

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

Н.С. Бужлаков, Е.В. Сопченко

## РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПЕРЕДАЧИ ДАННЫХ ОБ УРОВНЕ ГЛЮКОЗЫ В ПРИЛОЖЕНИИ «XDRIP+»

(Самарский университет)

Сахарный диабет является одним из самых распространенных заболеваний в мире. Это группа эндокринных заболеваний, которые связаны с нарушением усвоения глюкозы и абсолютной или относительной недостаточности гормона инсулина, вследствие чего развивается гипергликемия – увеличенное содержание глюкозы в крови [1].

Наиболее обременительным для больных является диабет первого типа, который обусловлен абсолютной инсулиновой недостаточности. Для таких больных гипогликемия – пониженное содержание глюкозы в крови настолько же опасна, как и гипергликемия. Вследствие чего больному необходимо регулярно знать уровень своего сахара в крови. Для этого используется регулярное измерение сахара глюкометром при помощи прокола пальца или системы непрерывного мониторинга глюкозы. Наиболее удобной и распространённым решением для работы с системами непрерывного мониторинга глюкозы на данный момент является приложение xDrip+ [2].

Если больной диабетом является несовершеннолетним существует необходимость просматривать показатели сахара больного на более чем одном устройстве, удалённо и на любом устройстве. На данный момент это направление слабо развито.

Передо мной была поставлена задача разработать автоматизированной системы передачи данных об уровне глюкозы в приложении «xDrip+». Требовалось разработать систему простую в настройке и доступную на как можно большем числе устройств. Для этого я проанализировал системы-аналоги, изучил принцип работы сенсора непрерывного мониторинга уровня глюкозы и приложение xDrip+ и разработал систему, которая состоит из клиентской и серверной части (рисунок 1). Серверная часть осуществляет функции авторизации и регистрации пользователя в системе, а также работы с базой данных системы, включающая функции работы с пользователями и границами гликемии в БД. Клиентская часть, в свою очередь, осуществляет визуализацию графика уровня сахара, среднего значения за период и рассчитанной тенденции к росту или па-