

Способ повышения эффективности сегментации речь/пауза на основе метода декомпозиции на эмпирические моды

А. К. Алимуратов¹, А. Ю. Тычков², П. П. Чураков³,
А. В. Агейкин⁴, А. В. Кузьмин⁵, М. А. Митрохин⁶, И. А. Чернов⁷

^{1,2,3,4,5,6,7} Пензенский государственный университет, Пенза, Россия

¹alansapfir@yandex.ru, ²tychkov-a@mail.ru, ³churakov-pp@mail.ru,

⁴ageykinav@yandex.ru, ⁵a.v.kuzmin@pnzgu.ru, ⁶mmax83@mail.ru,

⁷igorchernov999@mail.ru

Аннотация. *Актуальность и цели.* Сегментация речь/пауза является одной из важнейших задач обработки в речевых приложениях и представляет собой точное обнаружение границ начала и окончания вокализованной, невокализованной речи и пауз. Особенно это важно при анализе скорости, ускорения и энтропии распределения вокализованных, невокализованных участков речи и пауз, а также при анализе средней продолжительности пауз. Целью работы является повышение эффективности сегментации речь/пауза на основе метода декомпозиции на эмпирические моды. *Материалы и методы.* В работе использовалась уникальная технология адаптивного разложения нестационарных сигналов – улучшенная полная множественная декомпозиция на эмпирические моды с адаптивным шумом. Программная реализация способа была выполнена в среде математического моделирования © Matlab (MathWorks). *Результаты.* Разработан способ, основанный на применении декомпозиции на этапе предварительной обработки исходных речевых сигналов для формирования набора новых исследуемых сигналов, содержащих максимально достоверную информацию о границах начала и окончания участков вокализованной, невокализованной речи и пауз. Проведено исследование, в рамках которого оценивалось влияние метода декомпозиции и длительности исследуемых фрагментов сигналов на эффективность сегментации речь/пауза. Для сегментации использовались способы, основанные на анализе количества пересечения сигнала через нулевую ось, кратковременной энергии, а также на основе анализа одномерного расстояния Махаланобиса. *Выводы.* В соответствии с результатами исследований выявлено, что предложенный способ обеспечивает повышение эффективности сегментации участков вокализованной и невокализованной речи: для способа на основе анализа количества пересечения сигнала через нулевую ось – на 13,96 %; для способа на основе анализа кратковременной энергии – на 8,24 %; для способа на основе совместного анализа количества пересечения и кратковременной энергии – на 5,72 %; для способа на основе анализа одномерного расстояния Махаланобиса – на 17,85 %.

Ключевые слова: обработка речевых сигналов, сегментация речи, вокализованная и невокализованная речь, декомпозиция на эмпирические моды

Финансирование: работа выполнена при финансовой поддержке Совета по грантам Президента РФ, проект № МК-490.2020.8.

Для цитирования: Алимуратов А. К., Тычков А. Ю., Чураков П. П., Агейкин А. В., Кузьмин А. В., Митрохин М. А., Чернов И. А. Способ повышения эффективности сегментации речь/пауза на основе метода декомпозиции на эмпирические моды // Известия высших учебных заведений. Поволжский регион. Технические науки. 2021. № 2. С. 24–43. doi:10.21685/2072-3059-2021-2-3

EMD-based method to improve the efficiency of speech/pause segmentation

A.K. Alimuradov¹, A.Yu. Tychkov², P.P. Churakov³,
A.V. Ageykin⁴, A.V. Kuz'min⁵, M.A. Mitrokhin⁶, I.A. Chernov⁷

^{1,2,3,4,5,6,7}Penza State University, Penza, Russia

¹alansapfir@yandex.ru, ²tychkov-a@mail.ru, ³churakov-pp@mail.ru,

⁴ageykinav@yandex.ru, ⁵a.v.kuzmin@pnzgu.ru, ⁶mmax83@mail.ru,

⁷igorchernov999@mail.ru

Abstract. *Background.* Speech/pause segmentation is one of the most important tasks in speech applications being accurate detection of the boundaries of the beginning and the end of voiced and unvoiced speech, and pauses. This is especially important both when analyzing distribution speed, acceleration, and entropy of voiced and unvoiced speech sections, and pauses, and analyzing the average duration of pauses. The aim of the work is to improve the efficiency of speech/pause segmentation based on the method of empirical mode decomposition. *Materials and methods.* A unique technology for adaptive decomposition of non-stationary signals, namely, the improved complete ensemble empirical mode decomposition with adaptive noise, has been used in the work. The software implementation of the method was performed in ©MATLAB (MathWorks) mathematical modeling environment. *Results.* A decomposition-based method has been developed to be used at the preprocessing stage of the original speech signals to form a set of new investigated signals containing the most reliable information about the boundaries of the beginning and the end of the voiced and unvoiced speech, and pauses. The research to assess the influence of the decomposition method, and the duration of the studied signal fragments on the efficiency of speech/pause segmentation has been done. We have used the methods based on the analysis of zero-crossing rate, short-term energy, and one-dimensional Mahalanobis distance. *Conclusions.* Based on the research results, it was found that the proposed method provides an increase in the efficiency of segmentation of voiced and unvoiced speech sections: by 13.96% for the method based on the analysis of zero-crossing rate; by 8.24% for the method based on the analysis of short-term energy; by 5.72% for the method based on the combined analysis of zero-crossing rate and short-term energy; by 17.85% for the method based on the analysis of one-dimensional Mahalanobis distance.

Keywords: speech signal processing, speech segmentation, voiced and unvoiced speech, empirical mode decomposition

Acknowledgments: the research was financed by the Council for Grants of the President of the Russian Federation, project No. MK-490.2020.8

For citation: Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Ageykin A.V., Kuz'min A.V., Mitrokhin M.A., Chernov I.A. EMD-method to improve the efficiency of speech/pause segmentation. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki = University proceedings. Volga region. Engineering sciences.* 2021;2:24–43. (In Russ.). doi:10.21685/2072-3059-2021-2-3

Введение

Точное обнаружение границ начала и окончания информативных участков речевых сигналов является одной из важнейших задач в речевых приложениях. В соответствии с физиологией речевого аппарата речь человека содержит три вида информативных участков: пауза, или тишина; невокализованная речь (апериодическая речь случайного характера); вокализованная речь (квазипериодическая речь). Задача сегментации речь/пауза пред-

ставляет собой классификацию информативных участков на вокализованную, невокализованную речь и паузы.

В зависимости от назначения речевых приложений информативные участки речи имеют разную значимость для дальнейшей обработки. В приложениях распознавания речи, голосового управления и преобразования речи в текст паузы в анализируемых речевых сигналах являются неинформативными и удаляются на этапе предварительной обработки. Объясняется это тем, что основной набор информативных параметров дикторов (амплитудно-частотных, спектрально-временных, кепстральных и др.) сосредоточен в вокализованной и невокализованной речи. В приложениях, предназначенных для выявления патологии голоса человека, нарушений моторики речевого аппарата или оценки психоэмоционального состояния паузы в речевых сигналах, максимально информативны. Например: при анализе скорости, ускорения и энтропии распределения вокализованных, невокализованных участков и пауз, а также при анализе средней продолжительности пауз в речи.

На сегодня задача сегментации речь/пауза решается разными способами, которые можно разделить на частотные и временные. Частотные способы основаны на анализе мел-частотных кепстральных коэффициентов (Mel-Frequency Cepstral Coefficients, MFCC) [1] и линейно-частотных кепстральных коэффициентов (Linear-Frequency Cepstral Coefficients, LFCC) [2]. Временные способы основаны на анализе количества пересечения сигнала через нулевую ось (Zero-Crossing Rate, ZCR) [3], отклонения автокорреляционной функции (Autocorrelation Function, ACR) [4], кратковременной энергии (Short Time Energy, STE) [5], а также одномерного расстояния Махаланобиса (One Dimensional Mahalanobis Distance, ODMD) [6].

В данной статье представлен способ повышения эффективности сегментации речь/пауза на основе метода декомпозиции на эмпирические моды (ДЭМ). ДЭМ применяется на этапе предварительной обработки исходного речевого сигнала для формирования набора новых исследуемых сигналов, содержащих максимально достоверную информацию о границах начала и окончания участков вокализованной, невокализованной речи и пауз.

Данная статья является результатом научной работы, посвященной исследованию и разработке эффективных алгоритмов и способов обработки речевых сигналов [7–9].

1. Декомпозиция на эмпирические моды

ДЭМ представляет собой адаптивный метод разложения нестационарных данных [10], основным преимуществом которого является полная адаптивность. Базисные функции, используемые для разложения, извлекаются непосредственно из исследуемого сигнала и позволяют учитывать только ему свойственные особенности. В рамках декомпозиции многократно осуществляется процесс просеивания, в результате которого исследуемый сигнал раскладывается на эмпирические моды (ЭМ) и конечный неделимый остаток. Процесс просеивания включает в себя обязательное решение следующих задач:

Задача 1. Определение среднего значения верхней и нижней огибающих исследуемого сигнала.

Задача 2. Вычитание среднего значения верхней и нижней огибающих из исследуемого сигнала.

Задача 3. Многократное повторение первой и второй задачи до тех пор, пока среднее значение не будет близко к нулю.

Задача 4. Выделение сигнала первой ЭМ, у которого среднее значение верхней и нижней огибающих максимально приблизилось к нулю в соответствии с критерием останова.

Задача 5. Вычитание первой ЭМ из исследуемого сигнала.

Задача 6. Повторение с первой по пятую задачи до тех пор, пока не будет получен монотонный сигнал (конечный неделимый остаток), из которого невозможно извлечь ни одну ЭМ.

Аналитическое выражение ДЭМ выглядит следующим образом:

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

где $x(n)$ – исследуемый сигнал; n – дискретный отсчет времени; $IMF_i(n)$ – сигнал ЭМ; i – номер ЭМ; I – количество ЭМ; $r_I(n)$ – конечный неделимый остаток.

Метод ДЭМ впервые был представлен в 1998 г. [10]. На сегодня известны различные методы декомпозиций: множественная ДЭМ – МДЭМ (2009) [11], комплементарная МДЭМ – КМДЭМ (2010) [12], полная МДЭМ с адаптивным шумом – ПМДЭМАШ (2011) [13] и улучшенная ПМДЭМАШ (2014) [14]. Наиболее адаптивным методом декомпозиции для обработки речи является метод улучшенной ПМДЭМАШ. Особенность улучшенной декомпозиции заключается в добавлении к исследуемому сигналу контролируемого белого шума малой амплитуды для создания новых нулей и экстремумов (локальных особенностей) функции сигнала. Создание новых локальных особенностей позволяет устранить известные недостатки предыдущих методов декомпозиции: эффект смешивания ЭМ; остаточный шум; неполное разложение; неинформативные «паразитные» ЭМ, выделяемые на ранних этапах декомпозиции.

Математический аппарат методов декомпозиции с добавлением шума выглядит следующим образом:

$$x_j(n) = x(n) + w_j(n), \quad (2)$$

где $x_j(n)$ – зашумленный речевой сигнал белым шумом; $w_j(n)$ – белый шум; $j = 1, 2, \dots, J$ – реализации белого шума;

$$x_j(n) = \sum_{i=1}^I IMF_{j,i}(n) + r_{j,I}(n), \quad (3)$$

$$IMF_i(n) = \sum_{j=1}^J \frac{IMF_{j,i}(n)}{J}, \quad (4)$$

$$r_I(n) = \sum_{j=1}^J \frac{r_{j,I}(n)}{J}. \quad (5)$$

2. Способы сегментации речь/пауза

Как отмечалось ранее, способы на основе анализа ZCR , STE , а также на основе совместного анализа ZCR и STE являются самыми распространенными в задачах сегментации речи на информативные участки. Однако в упомянутых способах до сих пор существует основная проблема, ограничивающая их массовое применение в речевых приложениях. Ограничения связаны с выбором корректных пороговых значений, соответствующих вокализованной, невокализованной речи и паузам. Ниже представлены аналитические выражения, используемые для вычисления значений ZCR и STE .

Функция ZCR основана на сравнении знаков соседних дискретных отсчетов времени и определяется по следующей формуле:

$$ZCR_s = 0,5 \sum_{n=-\infty}^{\infty} \left| \operatorname{sgn}(x(n)) - \operatorname{sgn}(x(n-1)) \right| w(s-n), \quad (6)$$

где s – номер фрагмента; $\operatorname{sgn}(x)$ – знаковая функция ($\operatorname{sgn}(x) = 1$ при $x \geq 0$ и $\operatorname{sgn}(x) = -1$ при $x < 0$); w – функция анализируемого окна.

Для прямоугольного анализируемого окна формула (6) принимает следующий вид:

$$ZCR_s = 0,5 \sum_{n=1}^{N-1} \left| \operatorname{sgn}(x(s-1)N+n+1) - \operatorname{sgn}(x(s-1)N+n) \right|, \quad (7)$$

где N – количество дискретных отсчетов в исследуемом фрагменте.

Функция STE представляет собой сумму квадратов амплитуд дискретных отсчетов времени и определяется по следующей формуле:

$$E_s = \sum_{n=-\infty}^{\infty} [x(n)w(s-n)]^2. \quad (8)$$

Для прямоугольного анализируемого окна формула (8) принимает следующий вид:

$$E_s = \sum_{n=1}^N [x(s-1)N+n]^2. \quad (9)$$

Способ сегментации речь/пауза на основе анализа ZCR построен на предположении, что количество пересечений функции сигнала через нулевую ось для пауз с фоновым шумом больше по сравнению с вокализованной, невокализованной речью. Аналогично построен способ на основе анализа STE – кратковременная энергия вокализованной, невокализованной речи больше, чем энергия пауз с фоновым шумом. Однако данные предположения не совсем корректные, так как остается нерешенным вопрос – насколько текущие значения ZCR и STE должны быть больше, чем пороговые для корректной сегментации информативных участков. Кроме того, известно, что пороговые значения могут варьировать для каждого конкретного анализируемого речевого сигнала. В работе [15] авторами была предпринята попытка выбрать и обосновать пороговые значения ZCR и STE , соответствующие вокализован-

ной, невокализованной речи и паузам. В соответствии с выводами в работе [15] точность составила 65 % в сравнении с сегментацией, осуществленной вручную.

Способ сегментации речь/пауза на основе анализа *ODMD* построен на статистических свойствах фонового шума и физиологии речевого аппарата человека [6]. В соответствии с физиологией воспроизведения речи человек перед произношением выдерживает вынужденную начальную паузу, длительностью не менее 200 мс, которая соответствует фоновому шуму. Предполагается, что фоновый шум, регистрируемый во время начальной паузы, имеет Гауссовский характер, а остальные информативные участки вокализованной и невокализованной речи имеют другое распределение. В этом случае функция плотности вероятности распределения фонового шума является критерием для сегментации речь/пауза. Таким образом, решается основная проблема – выбор корректных пороговых значений.

В основе вычисления *ODMD* лежит функция плотности вероятности нормального распределения:

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}, \quad (10)$$

где μ и σ – математическое ожидание и стандартное отклонение независимых случайных величин y .

Как известно, кривая функции плотности вероятности нормального распределения имеет форму симметричного колоколообразного импульса. Независимые случайные величины имеют тенденцию группироваться около среднего значения. Пик нормального распределения соответствует $y = \mu$, а ширина пропорциональна стандартным отклонениям σ . Аналитическое выражение *ODMD* имеет следующий вид:

$$r = \frac{|y - \mu|}{\sigma}, \quad (11)$$

где выражение $|y - \mu|$ является естественной мерой расстояния от y к среднему значению μ (численно вероятности подчиняются следующим выражениям:

$$|y - \mu| \leq \sigma = 0,68, \quad |y - \mu| \leq 2\sigma = 0,95 \quad \text{и} \quad |y - \mu| \leq 3\sigma = 0,997.$$

В работе [16] представлен подробный сравнительный анализ результатов сегментации речь/пауза, полученных с помощью способов на основе анализа *ZCR*, *STE* и *ODMD*. В соответствии с выводами в работе [16] способ на основе анализа *ODMD* эффективнее для отдельных словосочетаний, чем способы на основе анализа *ZCR* и совместного анализа *ZCR* и *STE* на 5,6 и 13,18 % соответственно, для слитной речи – на 8,88 и 9,59 % соответственно.

3. Способ повышения эффективности сегментации речь/пауза

Способ повышения эффективности сегментации речь/пауза основан на применении ДЭМ на этапе предварительной обработки исходного речевого

сигнала. Результатом предварительной обработки является набор новых исследуемых сигналов, содержащих максимально достоверную информацию о границах начала и окончания участков вокализованной, невокализованной речи и пауз.

На рис. 1 представлена структура алгоритма предлагаемого способа. Блоки зеленого цвета представляют собой предварительную обработку на основе ДЭМ. Блоки фиолетового, красного и голубого цветов представляют способы сегментации на основе анализа *ODMD* и *ZCR*, *STE*. Блоки серого цвета не относятся к предлагаемому способу и предназначены для постобработки ошибок сегментации речь/пауза и сравнения результатов с сегментацией, осуществленной вручную.

Предварительная обработка на основе ДЭМ

Фрагментирование представляет собой линейное разделение речевого сигнала на отрезки (фрагменты) равной длительности. Фрагментирование основано на кратковременном анализе, в рамках которого фрагменты обрабатываются так, как если бы они были короткими речевыми сигналами с отличающимися свойствами. В соответствии с предлагаемым способом от длительности исследуемых фрагментов зависит результат последующей декомпозиции.

В работе [17] авторами представлены результаты исследований влияния длительности анализируемых речевых сигналов на частотно-избирательные свойства различных методов декомпозиции. В соответствии с полученными результатами в работе [17] сделан вывод, что для корректного частотно-временного анализа длительность исследуемых фрагментов должна быть от 10 до 50 мс.

Фрагментирование речевого сигнала осуществляется по следующим формулам:

$$S = \frac{x(n)}{L}, \quad (12)$$

где S – количество фрагментов в исследуемом речевом сигнале $x(n)$ (с округлением в меньшую сторону); L – количество дискретных отсчетов времени в одном фрагменте;

$$x_{s+1}(n) = x[(s \cdot L) + 1 : (s + 1) \cdot L]. \quad (13)$$

В предлагаемом способе повышения эффективности сегментации речь/пауза применяются следующие методы декомпозиции: ДЭМ, МДЭМ и улучшенная ПМДЭМАШ. Использование двух методов декомпозиции с добавлением шума объясняется тем, что МДЭМ и КМДЭМ, а также ПМДЭМАШ и улучшенная ПМДЭМАШ аналогичны с точки зрения просеивания ЭМ. Отличительной особенностью улучшенной ПМДЭМАШ от других методов декомпозиции с добавлением шума является локальное разложение белого шума на шумовые моды параллельно с разложением исследуемого сигнала. Использование шумовых мод в качестве добавляемого контролируемого белого шума на каждом этапе процесса просеивания обеспечивает полноту разложения [14].

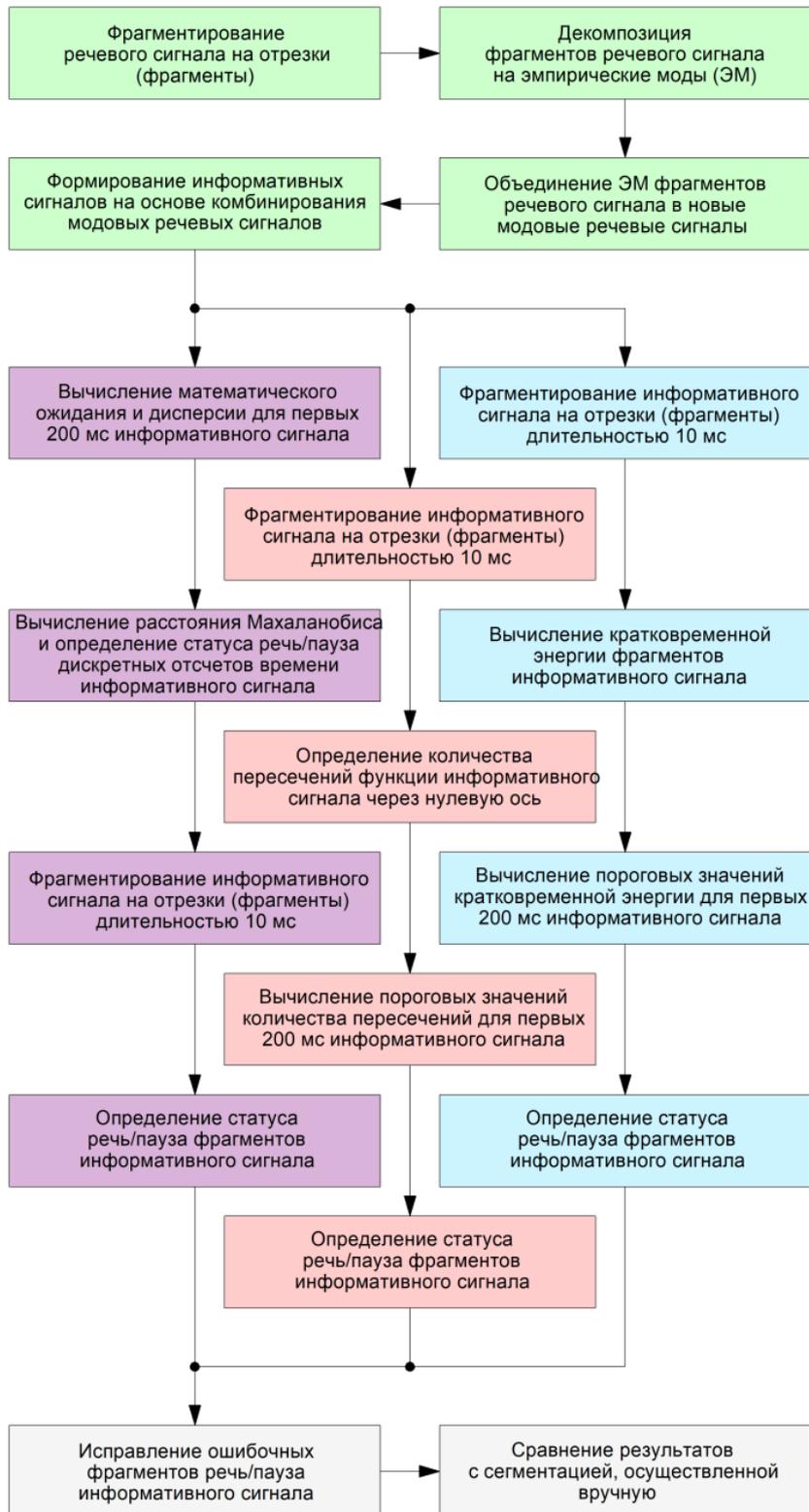


Рис. 1. Структурная схема способа повышения эффективности сегментации речь/пауза на основе метода ДЭМ

Для методов декомпозиции с добавлением шума применяются следующие настройки: стандартное отклонение амплитуды шума от амплитуды исследуемого сигнала – не более 20 %; количество реализаций белого шума – 100; допустимое максимальное количество просеивающих итераций – 50; отношение стандартных отклонений сигнала и шума на всех этапах процесса просеивания ЭМ неизменное (для метода улучшенной ПМДЭМАШ).

В соответствии с результатом декомпозиции каждый фрагмент исходного речевого сигнала представлен набором ЭМ. Для оптимизации вычислительных затрат предлагаемого способа осуществляется объединение ЭМ фрагментов исходного речевого сигнала в новые модовые речевые сигналы:

$$xmode_i(n) = \sum_{s=1}^S IMF_{s,i}[(s \cdot L) + 1 : (s + 1) \cdot L], \quad (14)$$

где $xmode_i(n)$ – модовый речевой сигнал; $i = 1, 2, \dots, I$ – количество ЭМ в наборах для каждого фрагмента s .

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

Формирование информативных сигналов на основе комбинирования четырех модовых речевых сигналов осуществляется по следующей формуле:

$$xinfo(n) = a \cdot x(n) + b \cdot xmode_1(n) + c \cdot xmode_2(n) + d \cdot xmode_3(n) + e \cdot xmode_4(n), \quad (15)$$

где a, b, c, d, e – коэффициенты, определяющие участие исходного и модовых речевых сигналов в формировании информативных сигналов (коэффициенты принимают только три значения: $-1, 0, 1$).

Суть формирования информативных сигналов заключается в поиске уникального сигнала, содержащего максимально достоверную информацию о границах начала и окончания участков вокализованной, невокализованной речи и пауз. В соответствии с ранее проведенными исследованиями [18] в табл. 1 представлены оптимальные информативные сигналы.

Таблица 1

Оптимальные информативные сигналы

Информативный сигнал	Значение коэффициентов	Комбинирование модовых речевых сигналов
1	2	3
$xinfo_1$	$a = 0, b = 1, c = 1, d = 0, e = 0$	$xmode_1 + xmode_2$
$xinfo_2$	$a = 0, b = 1, c = 1, d = 1, e = 0$	$xmode_1 + xmode_2 + xmode_3$
$xinfo_3$	$a = 0, b = 1, c = 1, d = 1, e = 1$	$xmode_1 + xmode_2 + xmode_3 + xmode_4$
$xinfo_4$	$a = 0, b = 0, c = 1, d = 1, e = 0$	$xmode_2 + xmode_3$
$xinfo_5$	$a = 0, b = 0, c = 1, d = 1, e = 1$	$xmode_2 + xmode_3 + xmode_4$
$xinfo_6$	$a = 0, b = 0, c = 1, d = 0, e = 1$	$xmode_2 + xmode_4$
$xinfo_7$	$a = 0, b = 1, c = 0, d = 1, e = 1$	$xmode_1 + xmode_3 + xmode_4$
$xinfo_8$	$a = 0, b = 1, c = 1, d = 0, e = 1$	$xmode_1 + xmode_2 + xmode_4$
$xinfo_9$	$a = 1, b = -1, c = 0, d = 0, e = 0$	$x - xmode_1$
$xinfo_{10}$	$a = 1, b = 0, c = -1, d = 0, e = 0$	$x - xmode_2$

Окончание табл. 1

1	2	3
$xinfo_{11}$	$a = 1, b = 0, c = 0, d = -1, e = 0$	$x - xmode_3$
$xinfo_{12}$	$a = 1, b = 0, c = 0, d = 0, e = -1$	$x - xmode_4$
$xinfo_{13}$	$a = 1, b = -1, c = -1, d = 0, e = 0$	$x - xmode_1 - xmode_2$
$xinfo_{14}$	$a = 1, b = 0, c = -1, d = -1, e = -1$	$x - xmode_2 - xmode_3 - xmode_4$
$xinfo_{15}$	$a = 1, b = 0, c = 0, d = -1, e = -1$	$x - xmode_3 - xmode_4$
$xinfo_{16}$	$a = 1, b = 0, c = -1, d = 0, e = -1$	$x - xmode_2 - xmode_4$
$xinfo_{17}$	$a = 1, b = 0, c = -1, d = -1, e = 0$	$x - xmode_1 - xmode_3$
$xinfo_{18}$	$a = 1, b = -1, c = 0, d = 0, e = -1$	$x - xmode_1 - xmode_4$

Сегментация на основе анализа ODMD

В соответствии с методикой [6] осуществляется вычисление математического ожидания и дисперсии для начальных 200 мс, соответствующих вынужденной паузе (фоновому шуму). При частоте дискретизации 8000 Гц начальная пауза составляет 1600 дискретных отсчетов:

$$\mu = \frac{1}{1600} \sum_{n=1}^{1600} xinfo_i(n), \quad (16)$$

$$\sigma = \sqrt{\frac{1}{1600} \sum_{n=1}^{1600} (xinfo_i(n) - \mu)^2}, \quad (17)$$

где $xinfo_i(n)$ – информативный сигнал.

Определение статуса речь/пауза дискретных отсчетов времени информативного сигнала заключается в вычислении и анализе значений ODMD:

$$\frac{|xinfo_i(n) - \mu|}{\sigma} \geq 3, \quad (18)$$

где $n = 1, 2, \dots, N$ – номер дискретного отсчета времени.

При выполнении условия дискретный отсчет времени соответствует речи, и наоборот, если условие не выполняется, то отсчет соответствует паузе.

Фрагментирование информативного сигнала осуществляется аналогично, как и фрагментирование исходного речевого сигнала по формулам (12) и (13). Длительность анализируемого фрагмента 10 мс подобрана в соответствии с минимальной длительностью смысловой речевой единицы – фонемы.

Определение статуса речь/пауза фрагментов информативного сигнала заключается в проверке следующего условия: количество дискретных отсчетов в анализируемом фрагменте, соответствующих речи или паузе, должно быть больше или равно половине общего числа отсчетов во фрагменте. То есть не менее 40 дискретных отсчетов времени при длительности фрагмента 10 мс и частоте дискретизации 8000 Гц.

Сегментация на основе анализа ZCR и STE

Фрагментирование информативного сигнала осуществляется по формулам (12) и (13). Вычисление значений ZCR и STE фрагментов информативного сигнала осуществляется по формулам (7) и (9) соответственно.

Для корректного анализа в предлагаемом способе представлено решение проблемы выбора пороговых значений ZCR и STE . По аналогии со способом сегментации на основе анализа $ODMD$ для формирования пороговых значений ZCR и STE используется начальная пауза. В соответствии с формулами (16) и (17) вычисляются математические ожидания и дисперсии значений ZCR и STE для начальных фрагментов, соответствующих начальной паузе 200 мс (фоновому шуму):

$$\mu_{ZCR} = \frac{1}{S} \sum_{s=1}^S ZCRxinfo_s, \quad (19)$$

$$\sigma_{ZCR} = \sqrt{\frac{1}{S} \sum_{s=1}^S (ZCRxinfo_s - \mu_{ZCR})^2}, \quad (20)$$

$$\mu_E = \frac{1}{S} \sum_{s=1}^S Exinfo_s, \quad (21)$$

$$\sigma_E = \sqrt{\frac{1}{S} \sum_{s=1}^S (Exinfo_s - \mu_E)^2}, \quad (22)$$

где $ZCRxinfo_s$, $Exinfo_s$ – значения ZCR и STE исследуемого фрагмента информативного сигнала.

Определение статуса речь/пауза фрагментов информативного сигнала заключается в проверке следующих условий:

$$|ZCRxinfo_s - \mu_{ZCR}| \geq \sigma_{ZCR}, \quad (23)$$

$$|Exinfo_s - \mu_E| \geq \sigma_E. \quad (24)$$

Если разница между текущим и средним значениями ZCR больше или равна стандартному отклонению, то фрагмент соответствует паузе. И наоборот, если условие не выполняется, то фрагмент соответствует речи. Аналогично: если разница между текущим и средним значениями STE больше или равна стандартному отклонению, то фрагмент соответствует речи. И наоборот, если условие не выполняется, то фрагмент соответствует паузе.

Сегментация на основе совместного анализа ZCR и STE осуществляется аналогично анализу значений ZCR , STE по отдельности с учетом сопоставления результатов.

Постобработка и сравнение результатов сегментации

В соответствии с физиологией речевого аппарата человек не может кратковременно (в течение 20 мс) изменить воспроизводимую речь на паузу и наоборот [4]. Данный физиологический аспект заложен в основу исправления ошибок сегментации – поиске некорректно определенных фрагментов речь/пауза. На рис. 2 представлен пример, иллюстрирующий ошибки сегментации речь/пауза.

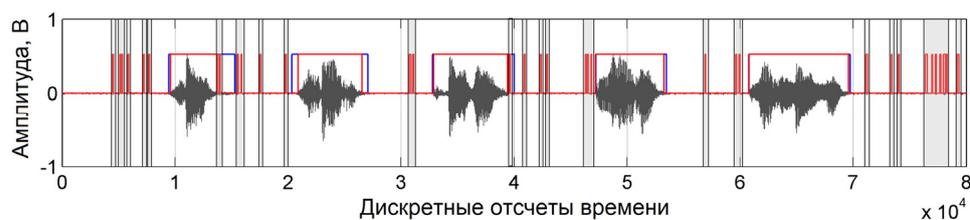


Рис. 2. Ошибки сегментации речь/пауза (линией красного цвета обозначен результат сегментации, линией синего цвета – результат сегментации, осуществленной вручную, фоном серого цвета обозначены некорректно определенные фрагменты)

Сегментация речевых сигналов на информативные участки вручную осуществлялась специалистами с помощью многоплатформенного аудиоредактора звуковых файлов Audacity, производитель Audacity Team (audacityteam.org).

4. Исследование способа повышения эффективности

Для оценки эффективности сегментации речь/пауза была сформирована база, состоящая из 600 речевых сигналов. Запись сигналов осуществлялась посредством специально разработанной методики, в рамках которой дикторы зачитывали следующий текстовый материал:

- статья из рекламно-информационной газеты, включающая публицистический текст на русском языке (не менее 200 слов);
- краткое детское литературное произведение, включающее фонетически сбалансированный текст на русском языке (не менее 200 слов);
- счет чисел от 0 до 99 на русском языке (100 слов).

Запись осуществлялась в обычном лабораторном помещении с естественным фоновым шумом. Произношение диктором – размеренное. Количество дикторов – 20 человек из числа мужчин и женщин. Параметры и пространственные характеристики микрофона не изменялись для всех регистрируемых дикторов.

Эффективность сегментации речь/пауза оценивалась посредством определения ошибок первого (α) и второго (β) рода. Основной задачей сегментации является точное обнаружение границ начала и окончания вокализованной и невокализованной речи, поэтому ошибкой первого рода считалось ошибочное присваивание речевому фрагменту статуса «пауза». Ошибкой второго рода считалось ошибочное присваивание фрагменту паузы статуса «речь». Ошибки первого и второго рода определялись в сравнении с результатом сегментации, осуществленной вручную.

В рамках исследования оценивалось влияние метода декомпозиции (ДЭМ, МДЭМ и улучшенная ПМДЭМАШ) и длительности (10, 30 и 50 мс) исследуемого фрагментов на эффективность сегментации речь/пауза способов, основанных на анализе ZCR , STE , совместном анализе ZCR и STE , а также анализе $ODMD$.

Программная реализация способа повышения эффективности сегментации речь/пауза выполнена в среде математического моделирования ©Matlab, производитель MathWorks (www.mathworks.com). Программная реализация метода улучшенной ПМДЭМАШ была заимствована из работы [19].

В табл. 2 представлены усредненные данные ошибок первого и второго рода, полученные по результатам сегментации с помощью упомянутых выше способов.

Таблица 2

Усредненные данные ошибок первого и второго рода, полученные по результатам сегментации способами на основе анализа *ZCR*, *STE*, совместного анализа *ZCR* и *STE*, а также анализа *ODMD*

Способ сегментации речь/пауза	Ошибки первого и второго рода, %	
	α	β
Способ на основе анализа <i>ZCR</i>	23,11	3,02
Способ на основе анализа <i>STE</i>	10,53	3,2
Способ на основе совместного анализа <i>ZCR</i> и <i>STE</i>	7,32	5,33
Способ на основе анализа <i>ODMD</i>	21,97	0,89

В табл. 3 представлены усредненные данные наименьших среднеарифметических значений ошибок первого и второго рода, полученных с помощью предлагаемого способа и соответствующих наилучшему результату сегментации речь/пауза. Данные в табл. 3 представлены в соответствии с результатами исследования трех модовых речевых сигналов и 18 информативных сигналов. Для каждого исследуемого сигнала данные ошибок первого и второго рода представлены при разных значениях длительности фрагментов 10, 30 и 50 мс.

Таблица 3

Усредненные данные наименьших среднеарифметических значений ошибок первого и второго рода, соответствующих наилучшему результату сегментации речь/пауза

Способ сегментации речь/пауза	Длительность исследуемых фрагментов, мс	Среднеарифметическое значение ошибок первого и второго рода, %		
		ДЭМ	МДЭМ	Улучшенная ПМДЭМАШ
Способ на основе анализа <i>ZCR</i>	10	8,04	8,59	8,71
	30	8,08	9,17	8,34
	50	8,07	8,45	8,48
Способ на основе анализа <i>STE</i>	10	3,05	3,89	4,66
	30	4,74	2,8	4,22
	50	3,43	2,03	2,35
Способ на основе совместного анализа <i>ZCR</i> и <i>STE</i>	10	4,38	5,84	5,62
	30	5,12	4,05	3,69
	50	3,57	2,31	2,62
Способ на основе анализа <i>ODMD</i>	10	2,91	2,66	3,0
	30	2,5	2,5	2,73
	50	2,96	2,73	2,73

Наилучший результат сегментации речь/пауза способом на основе анализа *ZCR* со среднеарифметическим значением ошибок первого и второго рода 8,04 % достигается при исследовании информативного сигнала 6. Ин-

формативный сигнал 6 представляет собой сумму второго и четвертого модовых речевых сигналов. Среднеарифметическое значение 8,04 % соответствует оптимальным значениям 9,15 и 6,93 % ошибок первого и второго рода соответственно. Исследование осуществлялось методом ДЭМ с длительностью фрагментов 10 мс.

Наилучший результат сегментации речь/пауза способом на основе анализа *STE* со среднеарифметическим значением ошибок первого и второго рода 2,03 % достигается при исследовании информативного сигнала 3. Информативный сигнал 3 представляет собой сумму всех четырех модовых речевых сигналов. Среднеарифметическое значение 2,03 % соответствует оптимальным значениям 2,29 и 1,78 % ошибок первого и второго рода соответственно. Исследование осуществлялось методом МДЭМ с длительностью фрагментов 50 мс.

Наилучший результат сегментации речь/пауза способом на основе совместного анализа *ZCR* и *STE* со среднеарифметическим значением ошибок первого и второго рода 2,31 % достигается при исследовании информативного сигнала 3. Среднеарифметическое значение 2,31 % соответствует оптимальным значениям 1,6 и 3,02 % ошибок первого и второго рода соответственно. Исследование осуществлялось методом МДЭМ с длительностью фрагментов 50 мс.

Наилучший результат сегментации речь/пауза способом на основе анализа *ODMD* со среднеарифметическим значением ошибок первого и второго рода 2,5 % достигается при исследовании информативного сигнала 2. Информативный сигнал 2 представляет собой сумму первых трех модовых речевых сигналов. Среднеарифметическое значение 2,5 % соответствует оптимальным значениям 4,12 и 0,89 % ошибок первого и второго рода соответственно. Исследование осуществлялось методом МДЭМ с длительностью фрагментов 30 мс.

В табл. 4 представлены исходные и достигнутые усредненные данные ошибок первого и второго рода, полученные по результатам сегментации речь/пауза с помощью способов на основе анализа *ZCR*, *STE*, совместного анализа *ZCR* и *STE*, а также анализа *ODMD*.

Таблица 4

Исходные и достигнутые усредненные данные ошибок первого и второго рода, полученные по результатам сегментации речь/пауза способами на основе анализа *ZCR*, *STE*, совместного анализа *ZCR* и *STE*, а также анализа *ODMD*

Способ сегментации речь/пауза	Ошибки первого и второго рода, %			
	Исходные данные		Достигнутые данные	
	α	β	α	β
Способ на основе анализа <i>ZCR</i>	23,11	3,02	9,15	6,93
Способ на основе анализа <i>STE</i>	10,53	3,2	2,29	1,78
Способ на основе совместного анализа <i>ZCR</i> и <i>STE</i>	7,32	5,33	1,6	3,02
Способ на основе анализа <i>ODMD</i>	21,97	0,89	4,12	0,89

На рис. 3 представлен пример, иллюстрирующий исходные и достигнутые результаты сегментации речь/пауза сигнала длительностью 10 с, представляющего собой сочетание следующих слов на русском языке: шанс, шар,

баян, Лара, нормально. Слова подобраны таким образом, чтобы в них содержались разные по способу образования звуки: гласные, сонорные, шумные смычные (взрывные, фрикативные) и шумные щелевые.

Заключение

Подводя итоги анализа результатов исследований, можно сделать следующие выводы:

1. Сравнительный анализ исходных и достигнутых результатов демонстрирует повышение эффективности сегментации участков вокализованной и невокализованной речи за счет применения методов ДЭМ на этапе предварительной обработки (рис. 3).

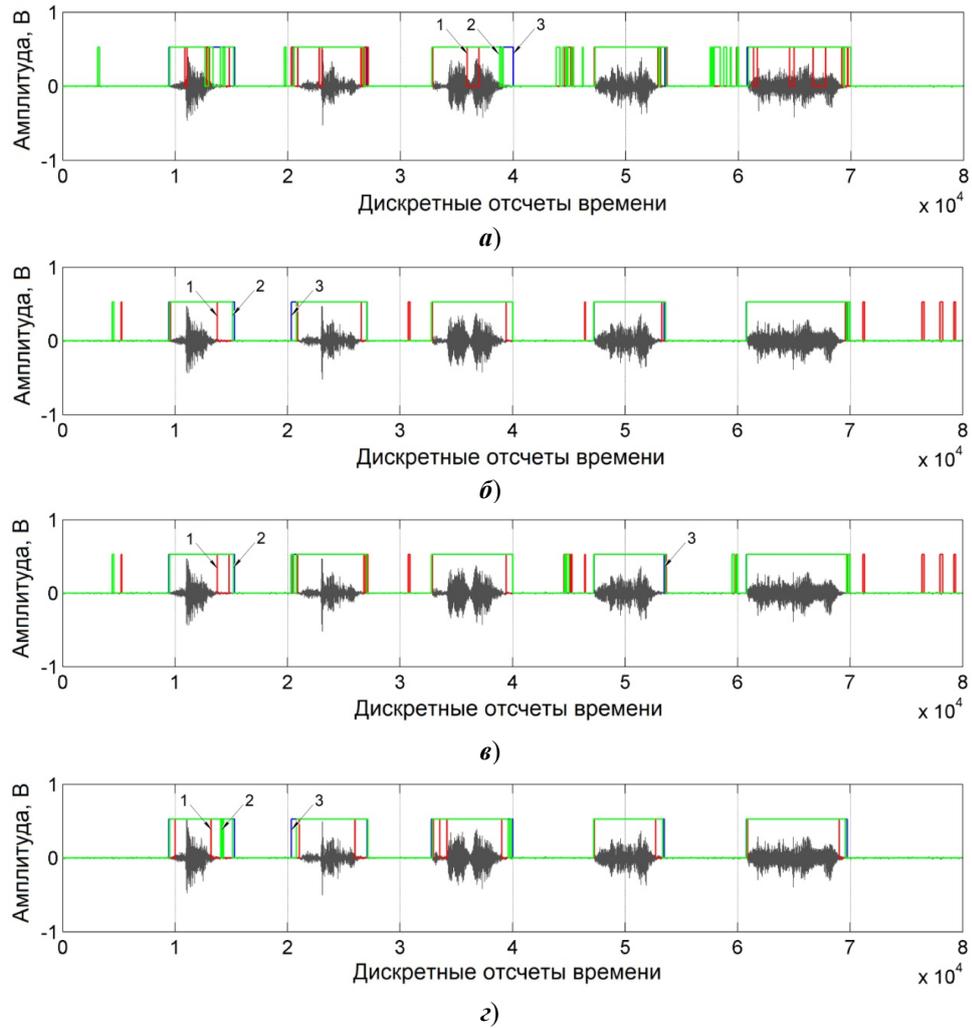


Рис. 3. Пример, иллюстрирующий исходные и достигнутые результаты сегментации речь/пауза (линией 1 обозначены исходные результаты сегментации, линией 2 – достигнутые результаты сегментации, линией 3 – результат сегментации, осуществленной вручную): *а* – способ на основе анализа ZCR; *б* – способ на основе анализа STE; *в* – способ на основе совместного анализа ZCR и STE; *г* – способ на основе анализа ODMD

Эффективность сегментации (см. рис. 3):

- для способа на основе анализа *ZCR* уменьшение ошибки первого рода на 13,96 %;
- для способа на основе анализа *STE* уменьшение ошибки первого рода на 8,24 %;
- для способа на основе совместного анализа *ZCR* и *STE* уменьшение ошибки первого рода на 5,72 %;
- для способа на основе анализа *ODMD* уменьшение ошибки первого рода на 17,85 %.

Повышение эффективности сегментации участков пауз отмечается только у способов на основе анализа *STE* и совместного анализа *ZCR* и *STE* – уменьшение ошибки второго рода на 1,42 и 2,31 % соответственно. Для способа на основе анализа *ZCR* отмечается ухудшение – увеличение ошибки на 3,91 %. Для способа на основе анализа *ODMD* точность сегментации не изменилась.

2. Наилучший результат сегментации речи на информативные участки с ошибками первого и второго рода 1,6 и 3,02 % соответственно достигается способом на основе совместного анализа *ZCR* и *STE* (при анализе информативного сигнала 3 методом МДЭМ с длительностью исследуемых фрагментов 50 мс) (см. табл. 2–4).

3. Наилучший результат сегментации речи на информативные участки для всех способов достигается при использовании метода МДЭМ и длительности исследуемых фрагментов 50 мс. Важно отметить, что применение методов ДЭМ и улучшенной ПМДЭМАШ, а также длительностей фрагментов 10 и 30 мс также обеспечивает приемлемые результаты сегментации речь/пауза (в пределах 4 % для среднеарифметических значений ошибок 1-го и 2-го рода).

В перспективе коллективом авторов планируется провести исследование помехоустойчивости и быстродействия способов на основе анализа *ZCR*, *STE*, совместного анализа *ZCR* и *STE*, а также анализа *ODMD* за счет применения методов ДЭМ на этапе предварительной обработки.

Список литературы

1. Martin A., Charlet D., Mauuary L. Robust speech/non-speech detection using LDA applied to MFCC // 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221) (ICASSP2001) (May 7–11, 2001). Salt Lake City, UT, USA. Vol. 1. P. 237–240.
2. Hlavnička J., Čmejla R., Tykalová T., Šonka K., Růžička E., Ruzs J. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder // Scientific Reports. 2017. Vol. 7 (12). 13 p.
3. Atal B., Rabiner L. R. A pattern recognition approach to voiced unvoiced-silence classification with applications to speech recognition // IEEE Trans. Acoust. Speech Signal Process. 1976. Vol. 24, № 3. P. 201–212.
4. Huang, X., Acero A., Hon H.-W. Spoken Language Processing. Guide to Algorithms and System Developmen. New Jersey : Prentice Hall, 2001. 980 p.
5. Childers D. G., Hand M., Larar J. M. Silent and voiced/unvoiced/ mixed excitation (four-way), classification of speech // IEEE Transaction on ASSP. 1989. Vol. 37, № 11. P. 1771–1774.
6. Duda R. O., Hart P. E., Strok D. G. Pattern Classification. 2nd ed. New Jersey : A Wiley-Interscience Publ. John Wiley & Sons, Inc., 2001. 688 p.

7. Алимуратов А. К., Тычков А. Ю., Чураков П. П., Султанов Б. В. Способ определения формантной разборчивости речи для оценки психоэмоционального состояния операторов систем управления с высокой степенью ответственности // Измерение. Мониторинг. Управление. Контроль. 2019. № 4 (30). С. 58–69.
8. Алимуратов А. К., Тычков А. Ю., Чураков П. П., Артамонов Д. В. Помехоустойчивый алгоритм определения просодических характеристик речевых сигналов для систем оценки психоэмоционального состояния человека // Известия высших учебных заведений. Поволжский регион. Технические науки. 2019. № 3 (51). С. 3–16.
9. Алимуратов А. К., Тычков А. Ю., Чураков П. П. Оценка психоэмоционального состояния человека на основе декомпозиции на эмпирические моды и кепстрального анализа речевых сигналов // Вестник Пензенского государственного университета. 2018. № 2. С. 89–95.
10. 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.
11. 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.
12. Yeh J.-R., Shieh J.-S., Huang N. E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method // Advances in Adaptive Data Analysis. 2010. № 2 (2). P. 135–156.
13. Torres M. E., Colominas M. A., Schlotthauer G., Flandrin P. A complete Ensemble Empirical Mode decomposition with adaptive noise // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-11) (May 22–27, 2011). Prague, Czech Republic, 2011. P. 4144–4147.
14. Colominas M. A., Schlotthauer G., Torres M. E. Improved complete ensemble EMD: a suitable tool for biomedical signal processing // Biomed. Signal Proces. 2014. Vol. 14. P. 19–29.
15. Greenwood M. A., Kinghorn A. SUVing: automatic silence/unvoiced/voiced classification of speech // Undergraduate Coursework, Department of Computer Science, The University of Sheffield, UK, 1999. 4 p.
16. Saha G., Chakroborty S., Senapat S. A new silence removal and endpoint detection algorithm for speech and speaker recognition applications // Eleventh National Conference on Communications (NCC-2005) (Jan. 28–30, 2005). Kharagpur, India, 2005. P. 51–61.
17. 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. (г. Пенза, 27 – 29 ноября 2019 г.). Пенза : Изд-во ПГУ, 2019. С. 201–205.
18. Alimuradov A. K., Churakov P. P., Tychkov A. Yu., Artemov I. I., Kuzmin A. V. Improvement of the Efficiency of Voice Control Based on the Complementary Ensemble Empirical Mode Decomposition // 2016 International Siberian Conference on Control and Communications (SIBCON 2016) (May 12–14, 2016). Moscow, Russia, 2016. 6 p.
19. National University of Entre Ríos. The Laboratory of Signals and Nonlinear Dynamics, Faculty of Engineering. URL: <http://www.bioingenieria.edu.ar/grupos/ldnlys>. (дата обращения: 01.05.2021).

References

1. Martin A., Charlet D., Mauuary L. Robust speech/non-speech detection using LDA applied to MFCC. 2001 IEEE International Conference on Acoustics, Speech, and Signal

- Processing. Proceedings (Cat. No.01CH37221) (ICASSP2001) (May 7–11, 2001)*. Salt Lake City, UT, USA. 2001;1:237–240.
2. Hlavnička J., Čmejla R., Tykalová T., Šonka K., Růžička E., Ruzs J. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific Reports*. 2017;7(12):13.
 3. Atal B., Rabiner L.R. A pattern recognition approach to voiced unvoiced-silence classification with applications to speech recognition. *IEEE Trans. Acoust. Speech Signal Process.* 1976;24(3):201–212.
 4. Huang, X., Acero A., Hon H.-W. *Spoken Language Processing. Guide to Algorithms and System Developmen*. New Jersey: Prentice Hall, 2001:980.
 5. Childers D.G., Hand M., Larar J.M. Silent and voiced/unvoiced/ mixed excitation (four-way), classification of speech. *IEEE Transaction on ASSP*. 1989;37(11):1771–1774.
 6. Duda R.O., Hart P.E., Strok D.G. *Pattern Classification. 2nd ed*. New Jersey: A Wiley-Interscience Publ. John Wiley & Sons, Inc., 2001:688.
 7. Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Sultanov B.V. A method for determining formatted speech intelligibility for assessing the psychoemotional state of control system operators with a high degree of responsibility. *Izmerenie. Monitoring. Upravlenie. Kontrol' = Measurement. Monitoring. Management. Control*. 2019;4(30):58–69. (In Russ.)
 8. Alimuradov A.K., Tychkov A.Yu., Churakov P.P., Artamonov D.V. Noise-resistant algorithm for determining the prosodic characteristics of speech signals for systems assessing the psychoemotional state of a person. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskije nauki = University proceedings. Volga region. Engineering sciences*. 2019;3(51):3–16. (In Russ.)
 9. Alimuradov A.K., Tychkov A.Yu., Churakov P.P. Assessment of the psychoemotional state of a person based on the decomposition into empirical modes and cepstral analysis of speech signals. *Vestnik Penzenskogo gosudarstvennogo universiteta = Bulletin of Penza State University*. 2018;2:89–95. (In Russ.)
 10. 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;A454:903–995.
 11. Zhaohua W., Huang N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*. 2009;1(1):1–41.
 12. Yeh J.-R., Shieh J.-S., Huang N.E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*. 2010;2(2):135–156.
 13. Torres M.E., Colominas M.A., Schlotthauer G., Flandrin P. A complete Ensemble Empirical Mode decomposition with adaptive noise. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-11) (May 22–27, 2011)*. Prague, Czech Republic, 2011:4144–4147.
 14. Colominasa M.A., Schlotthauera G., Torres M.E. Improved complete ensemble EMD: a suitable tool for biomedical signal processing. *Biomed. Signal Proces*. 2014;14:19–29.
 15. Greenwood M.A., Kinghorn A. SUVing: automatic silence/unvoiced/voiced classification of speech. *Undergraduate Coursework, Department of Computer Science, The University of Sheffield, UK*, 1999:4.
 16. Saha G., Chakroborty S., Senapat S. A new silence removal and endpoint detection algorithm for speech and speaker recognition applications. *Eleventh National Conference on Communications (NCC-2005) (Jan. 28–30, 2005)*. Kharagpur, India, 2005:51–61.
 17. Alimuradov A.K., Fokina E.A., Zhurina A.E. Studying the influence of the duration of the analyzed speech signals on the frequency-selective properties of the decomposition into empirical modes. *Novye informatsionnye tekhnologii i sistemy: sb. nauch. st. XVI Mezhdunar. nauch.-tekhn. konf. (g. Penza, 27 – 29 noyabrya 2019 g.) = New information technologies and systems: proceedings of the 16th International scientific and*

technical conference (Penza, November 27-29, 2019). Penza: Izd-vo PGU, 2019:201–205. (In Russ.)

18. Alimuradov A.K., Churakov P.P., Tychkov A.Yu., Artemov I.I., Kuzmin A.V. Improvement of the Efficiency of Voice Control Based on the Complementary Ensemble Empirical Mode Decomposition. *2016 International Siberian Conference on Control and Communications (SIBCON 2016) (May 12–14, 2016)*. Moscow, Russia, 2016:6.
19. *National University of Entre Ríos. The Laboratory of Signals and Nonlinear Dynamics, Faculty of Engineering*. Available at: <http://www.bioingenieria.edu.ar/grupos/ldnlys>. (accessed 01.05.2021).

Информация об авторах / Information about the authors

Алан Казанферович Алимуратов

кандидат технических наук, доцент
кафедры радиотехники
и радиоэлектронных систем,
директор студенческого научно-
производственного бизнес-инкубатора,
Пензенский государственный
университет (Россия, г. Пенза,
ул. Красная, 40)

E-mail: alansapfir@yandex.ru

Alan K. Alimuradov

Candidate of engineering sciences, associate
professor of the sub-department of radio
engineering and radioelectronic systems,
director of the student research
and production business incubator,
Penza State University (40 Krasnaya
street, Penza, Russia)

Александр Юрьевич Тычков

доктор технических наук, заведующий
кафедрой радиотехники и
радиоэлектронных систем,
заместитель директора научно-
исследовательского института
фундаментальных и прикладных
исследований, Пензенский
государственный университет
(Россия, г. Пенза, ул. Красная, 40)

E-mail: tychkov-a@mail.ru

Aleksandr Yu. Tychkov

Doctor of engineering sciences, head
of the sub-department of radio engineering
and radioelectronic systems, deputy
director of the Research Institute
for Basic and Applied Studies, Penza
State University (40 Krasnaya
street, Penza, Russia)

Петр Павлович Чураков

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

E-mail: churakov-pp@mail.ru

Petr P. Churakov

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

Алексей Викторович Агейкин

ассистент кафедры микробиологии,
эпидемиологии и инфекционных
болезней, Медицинский институт,
Пензенский государственный
университет (Россия, г. Пенза,
ул. Красная, 40)

E-mail: ageykinav@yandex.ru

Aleksey V. Ageykin

Assistant of the sub-department
of microbiology, epidemiology
and infectious diseases, Medical Institute,
Penza State University (40 Krasnaya
street, Penza, Russia)

Андрей Викторович Кузьмин

доктор технических наук, доцент,
профессор кафедры информационно-
вычислительных систем, Пензенский
государственный университет
(Россия, г. Пенза, ул. Красная, 40)

E-mail: a.v.kuzmin@pnzgu.ru

Максим Александрович Митрохин

доктор технических наук, доцент,
заведующий кафедрой вычислительной
техники, Пензенский государственный
университет (Россия, г. Пенза,
ул. Красная, 40)

E-mail: mmax83@mail.ru

Игорь Алексеевич Чернов

студент, Пензенский
государственный университет
(Россия, г. Пенза, ул. Красная, 40)

E-mail: igorchernov999@mail.ru

Andrey V. Kuz'min

Doctor of engineering sciences, associate
professor, professor of the sub-department
of information and computing systems,
Penza State University (40 Krasnaya street,
Penza, Russia)

Maksim A. Mitrokhin

Doctor of engineering sciences, associate
professor, head of the sub-department
of computing technology, Penza State
University (40 Krasnaya street,
Penza, Russia)

Igor' A. Chernov

Student, Penza State University
(40 Krasnaya street, Penza, Russia)

Поступила в редакцию / Received 12.05.2021

Поступила после рецензирования и доработки / Revised 25.05.2021

Принята к публикации / Accepted 07.06.2021