

УДК 004.5

ОБОБЩЕНИЕ МЕТОДА ЦЕПОЧЕК ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ ДЛЯ АНАЛИЗА СИГНАЛОВ РАЗЛИЧНОЙ ПРИРОДЫ

Туровский Я.А., к.м.н., зав. лабораторией медицинской кибернетики Воронежского государственного университета (ВГУ), e-mail: yaroslav_turovsk@mail.ru;

Кургалин С.Д., д.ф.-м.н., зав. кафедрой цифровых технологий ВГУ, e-mail: kurgalin@bk.ru;

Вахтин А.А., к.ф.-м.н., доцент кафедры программирования и информационных технологий ВГУ, e-mail: alvahtin@gmail.com;

Борзунов С.В., к.ф.-м.н., доцент кафедры цифровых технологий ВГУ, e-mail: kurgalin@bk.ru;

Белобродский В.А., аспирант кафедры цифровых технологий ВГУ, e-mail: belobrodsky@yandex.ru.

Ключевые слова: биомедицинские сигналы, спектральный анализ, вейвлет-преобразование, цепочки локальных максимумов и минимумов, рибонуклеиновая кислота, метеорология, биржевые котировки.

Введение

Методы спектрального анализа одномерных биомедицинских сигналов достаточно давно и прочно вошли не только в научные исследования, но и в клиническую практику [1, 2]. Они широко используются в анализе электроэнцефалограмм (ЭЭГ), являющихся результатом регистрации электрической активности мозга через неповреждённые покровы головы [3], электромиограмм – регистрации электрической активности мышц [4], variability сердечного ритма – изменения частоты пульса во времени [5] и т.д. Одним из важных направлений совершенствования теории и практики спектрального оценивания применительно к одномерным сигналам, регистрируемым с человека, является создание новых методов выделения научно-значимой информации.

Большое распространение получили в настоящее время подходы, осуществляющие развитие уже хорошо зарекомендовавших себя преобразования Фурье или вейвлет-преобразования для исследования разнообразных сигналов. При этом совершенствование методов анализа одномерных сигналов на основе указанных выше преобразований во многом не затрагивает методов расчёта их коэффициентов. Основной упор разработчики новых методов делают на улучшение подходов к обработке коэффициентов, полученных в ходе преобразования, и на извлечение из них интересующей исследователей информации.

В работах [6-9] был разработан «метод цепочек локальных экстремумов» – метод построения цепочек локальных максимумов (ЦЛМ) и минимумов (ЦЛМин) на основе матрицы квадратов коэффициентов вейвлет-преобразования для исследования сигналов ЭЭГ. Полученные ЦЛМ в рамках принятой интерпретации могут рассматриваться как частотно-временные области с максимальной энергией сигнала в заданном диапазоне

Обобщен метод построения цепочек локальных экстремумов матрицы квадратов коэффициентов вейвлет-преобразования, разработанный и использованный ранее для анализа сигналов ЭЭГ («метод цепочек локальных экстремумов»). Существенным моментом развития метода является формулировка общего правила для присоединения имеющегося экстремума локального спектра к уже сформированной цепочке экстремумов. Предлагаемый подход позволяет избежать чрезмерного объединения экстремумов в цепочки с большими разрывами во временной области. Обобщенный метод цепочек локальных экстремумов был проверен на данных разной природы: перекодированные РНК-последовательности генома вируса гепатита С, метеорологические данные – среднемесячная температура, сведения о биржевых котировках акций.

времени и частоты. Для ЦЛМин верно обратное – это частотно-временные области с минимальной энергией сигнала в заданном диапазоне времени и частоты. Данный метод был апробирован на сигналах ЭЭГ и рассчитанных на их основе т.н. «вызванных потенциалах головного мозга» [10], а также на ряде других сигналов, зарегистрированных с человека [11]. Однако, для эффективного внедрения этого метода в разные области исследований необходимо провести и обосновать его обобщение на сигналы различной природы.

Целью данной работы является разработка наиболее общего, не зависящего от природы одномерного сигнала метода построения цепочек локальных экстремумов матрицы квадратов коэффициентов вейвлет-преобразования для извлечения из данных дополнительной значимой информации.

Метод построения ЦЛМ И ЦЛМин на основе матриц вейлет-преобразования

В предлагаемом методе используется тот факт, что после непрерывного вейвлет-преобразования квадраты элементов полученной матрицы $W^2(a, b)$ отражают распределение энергии сигнала по масштабам вейвлет-преобразования. Величины $W^2(a, b)$ формируют трёхмерную область, содержащую как локальные максимумы энергии сигнала на определённом масштабе a и сдвиге вейвлет-функции b , отражающем сдвиг вейвлета по оси времени t , так и локальные минимумы. Очевидно, что локальные максимумы и минимумы величин $W^2(a, b)$

несут информацию о свойствах исследуемых сигналов. Эти свойства могут быть содержательно интерпретированы, исходя из современных представлений о причинах появления в анализируемых сигналах тех или иных частотно-временных особенностей и их характеристиках. Для значительной части сигналов различной природы свойственно формирование достаточно близко расположенных в частотно-временном пространстве последовательностей локальных максимумов и минимумов.

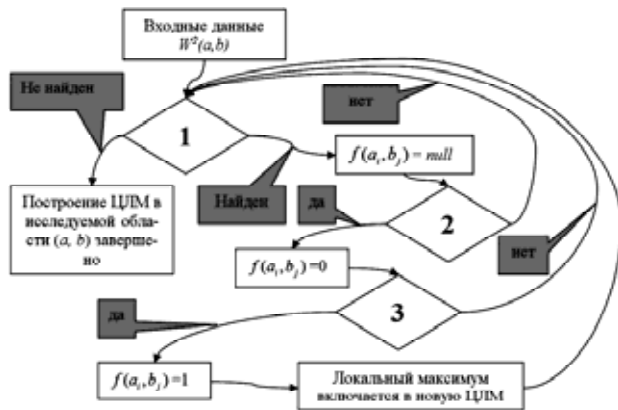


Рис. 1. Схема формирования цепочек локальных максимумов. На схеме: 1. Нахождение локального максимума с координатой b_j , не включенного ни в одну из уже построенных ЦПМ и не проверенного для текущей конфигурации построенных и строящихся ЦПМ

2. $(a_i \in [a_{i-1} - (\nu a_s + u), a_{i-1} + (\nu a_s + u)]) \vee (b_j \in [b_{j-1}, b_{j-1} + (\nu a_s + u)])$
3. $(a_i \in [a_{i-1} - (\nu a_s + u), a_{i-1} + (\nu a_s + u)]) \wedge (b_j \in [b_{j-1}, b_{j-1} + (\nu a_s + u)])$

Схема формирования таких цепочек на примере локальных максимумов представлена на рис. 1, где использовано: a_{i-1} – масштаб вейвлет-преобразования для локального максимума концевой точки ЦПМ, ближайшей по величине масштаба к предполагаемому для включения в ЦПМ локальному максимуму; b_{j-1} – номер отсчёта (номер скейлграммы) [6-10] концевой точки ЦПМ, ближайшей к аналогичному значению предполагаемого для включения в ЦПМ локального максимума; a_s – значение масштаба первой точки формирующейся ЦПМ; индексы $i = 1, 2, \dots$ и $j = 1, 2, \dots$ характеризуют текущие номера локальных экстремумов, потенциально принадлежащих той или иной ЦПМ (или ЦПМин); $f(a_i, b_j)$ – логическая функция, описывающая решение о включении или не включении новой точки локального экстремума в уже существующие ЦПМ или ЦПМин (для локальных максимумов и минимумов, соответственно).

Константы u и ν получены эмпирически [6-10] для данного типа сигналов. Параметр b в нашем методе расчета является безразмерным, он представляет собой номер отсчёта АЦП, а в размерность времени его значения переводятся путём умножения на частоту дискретизации исследуемого сигнала. Зная частоту дискретизации сигнала, нетрудно получить размерность частоты (в Гц) тех или иных его компонентов.

Из представленной на рис. 1 схемы видно, что формирование ЦПМ и ЦПМин определяется константами u и ν . Для расчёта констант используются два параметра: k – число точек в ЦПМ (или ЦПМин) и Δb – разность первого и последнего порядковых номеров локальных спектров матрицы $W^2(a, b)$, в которых находились локальные максимумы (минимумы). В случае если локальный спектр строился при каждом временном отсчёте, то Δb будет равным разности порядковых номеров первого и последнего отсчёта ЦПМ (или ЦПМин). Величина $k/\Delta b$ принадлежит диапазону значений $(0, 1]$ и, по сути, выражает долю насыщения ЦПМ определённой длины локальными максимумами. Очевидно, что в случае, если число локальных максимумов (или минимумов) равно длине цепочки Δb , то $k/\Delta b = 1$. Такую цепочку следует признать удачно построенной, т.к. отсутствуют локальные спектры, принадлежащие отрезку Δb , но не содержащие локальных максимумов (минимумов) для включения в цепочку. Поскольку в каждый момент времени (или для конкретного порядкового номера отсчёта) b в ЦПМ (ЦПМин) включается только один локальный экстремум, отношение $k/\Delta b$ принципиально не может быть большим единицы. В случае же, если происходит объединение двух далеко отстоящих по времени ЦПМ (ЦПМин) в одну, то отношение $k/\Delta b \ll 1$. Подобное объединение следует признать «искусственным», не отражающим реальное распределение энергии сигнала по элементам матрицы $W^2(a, b)$. Таким образом, задача поиска оптимальных значений u и ν для корректного построения ЦПМ (ЦПМин) сводится к максимизации отношения $k/\Delta b$ (т.е. $k/\Delta b \rightarrow \max$) в пространстве изменения параметров (u, ν) .



Рис. 2. Схема выбора параметров u и ν для расчета ЦПМ и ЦПМин

В настоящее время не существует теоретических подходов, позволяющих корректно рассчитать значения u и ν для получения максимальных величин $k/\Delta b$ для всего пула ЦПМ (или ЦПМин) в пространстве параметров (a, b) . Простейшим решением этой проблемы является перебор с определённым шагом всех возможных значений u и ν для поиска максимума отношения $k/\Delta b$ (см. рис. 2).

Апробация метода на сигналах различной природы

Подход к выбору параметров формирования ЦПМ, схематически представленный на рис. 2, ранее был использован [6-10] для анализа ЭЭГ вызванных потенциа-

лов. Для этого класса сигналов были получены значения $u=3$ и $v=0.05$.

Рассмотрим применение данного подхода для исследования различных одномерных данных. Возьмём, во-первых, данные о последовательности нуклеотидов в нуклеиновых кислотах вируса гепатита С одного из изолятов [12].

Произведём перекодировку нуклеотидов:

$$T(d) = \begin{cases} T(d-1)-1, P_i \\ T(d-1)+1, P_u \end{cases} \quad (1)$$

где $T(d)$ – текущие значения перекодированной функции последовательности РНК; $d=1,2,3...$ – порядковый номер (позиция) нуклеотида в исследуемом ряду; присутствие P_i – означает наличие в позиции d нуклеотида с пиримидиновым основанием, а присутствие P_u – наличие в позиции d нуклеотида с пуриновым основанием. При $d=1$ $T(d)=0$.

На рис. 3 представлены результаты оценки структур ЦЛМ для перекодированных согласно (1) РНК-последовательностей вируса гепатита С. Глобальный максимум величины $k/\Delta b$ при $v \in [0, 0.1]$ и $u \in [3, 7]$ имеет координаты (0,3), что дает возможность считать координаты этой области пространства (u, v) как параметры для формул, изложенных на рис.1. Однако при $v \in [0, 0.1]$ и $u \in [6, 7]$ выполняется условие $k/\Delta b > 0.97$, что также позволяет рассматривать и эту область пространства (u, v) в качестве приемлемого варианта для использования констант u и v при построении ЦЛМ (ЦЛМин), согласно схеме формирования цепочек, представленной на рис. 1.

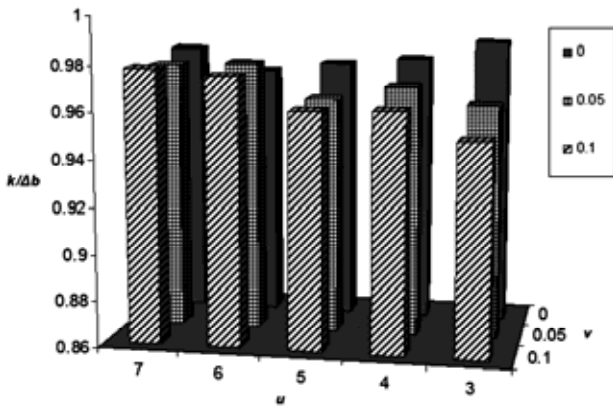


Рис. 3. Результаты оценки величин $k/\Delta b$ в пространстве (u, v)

для РНК- последовательностей вируса гепатита С

Во-вторых, исследуем информацию о среднемесячной температуре по Москве [13], начиная с 1989 г. В этом случае выявляется глобальный максимум функции $k/\Delta b$ с координатами (0,3). Несколько уступают ему значения $k/\Delta b$ в координатах (0,4) и (0.05,4). Остальные значения $k/\Delta b$, рассчитанные для других координат v и u , существенно ниже (рис. 4). Следовательно, для этого случая выделяется только одна область параметров (u, v) , в отличие от примера с РНК вируса гепатита С, где таких областей было выделено две.

В-третьих, отличную от предыдущих двух случаев структуру распределения значений $k/\Delta b$ в пространстве (u, v) продемонстрировали данные биржевых котировок

обыкновенных акций Внешторгбанка [14] за период с 1 по 4 декабря 2014 г. Оценивалось время открытия каждого из одноминутных интервалов торгов на ММВБ. Полученные результаты (см. рис. 5), продемонстрировали наличие только одного максимума величины $k/\Delta b$. При этом, в отличие от приведённых выше примеров, остальные значения функции $k/\Delta b$ были существенно ниже глобального максимума. Следовательно, соответствующие этим значениям $k/\Delta b$ координаты в пространстве (u, v) не могут быть использованы в качестве параметров формул, указанных на рис. 1, т.к. для них не выполняется указанное ранее условие $k/\Delta b \rightarrow \max$.

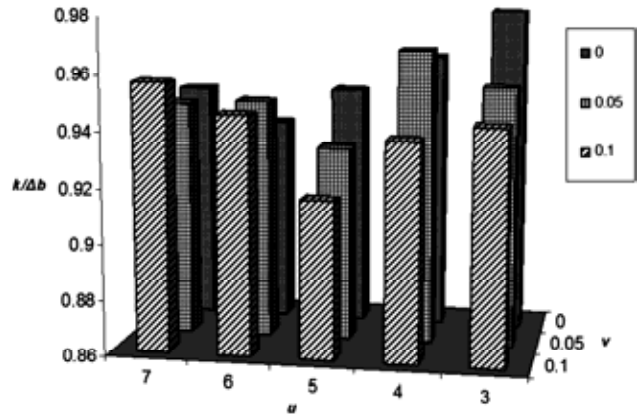


Рис. 4. Результаты оценки величин $k/\Delta b$ в пространстве (u, v)

для среднемесячной температуры по г. Москве

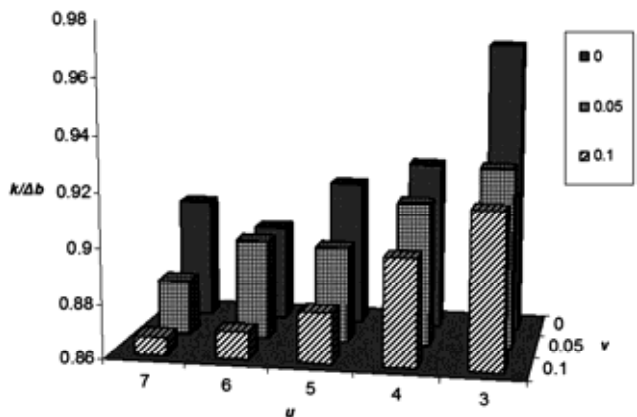
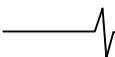


Рис. 5. Результаты оценки величин $k/\Delta b$ в пространстве параметров (u, v)

для биржевых котировок обыкновенных акций ВТБ за период с 1 по 3 декабря 2014 г.

Заключение

Метод построения цепочек локальных экстремумов матрицы квадратов коэффициентов вейвлет-преобразования, разработанный и использованный ранее для анализа сигналов ЭЭГ («метод цепочек локальных экстремумов»), обобщен в настоящей работе для исследования любых одномерных сигналов или данных, вне зависимости от их природы. Существенным моментом развития метода является формулировка общего правила для присоединения имеющегося экстремума локального спектра к уже сформированной цепочке экстремумов путем определения величин параметров u и v , от которых зависит характеристики построенной ЦЛМ или



ЦЛМин. Выбор параметров должен обеспечить максимизацию отношения числа экстремумов (только минимумов или только максимумов) ко времени существования цепочки. Данный подход позволяет избежать чрезмерного объединения экстремумов в цепочки с большими разрывами во временной области.

Предложенный универсальный метод был проверен на данных разной природы: на перекодированных РНК-последовательностях генома вируса гепатита С, на метеорологических данных – среднемесячных температурах, на сведениях о биржевых котировках акций. Продемонстрировано, что для каждого из этих типов данных величина $k/\Delta b$ имеет свои максимумы разной амплитуды в различных областях пространства параметров (u, v). Можно сделать вывод, что величина $k/\Delta b$ является достаточно чувствительным критерием для выбора значений u и v , необходимых для корректного построения ЦЛМ и ЦЛМин. Следовательно, данный подход позволяет получать более детальную информацию об исследуемых объектах на основе улучшения спектральных оценок анализируемых сигналов и анализа динамики их поведения во времени.

Литература

1. Биоэлектрическая активность мозга у больных с невротическими и неврозоподобными расстройствами (по данным спектрального анализа) Шульц Е.В., Бабулин И.Н., Караваева Т.А., Карвасарский Б.Д., Слезин В.Б. Обзорение психиатрии и медицинской психологии им. В.М. Бехтерева. 2010. № 3. С. 26-31.
2. Взаимосвязь быстрых изменений индивидуального поддиапазона волн электроэнцефалограммы и кардиоритма во время сна Васильев Е.Н., Урываев Ю.В. Физиология человека. 2006. Т. 32. № 4. С. 18-23.
3. Зенков Л.Р. Клиническая электроэнцефалография (с элементами эпилептологии). – М.: МЕДПресс-информ, 2011.– 356 с.
4. Анализ мышечных биопотенциалов методом спектрографии Дегтярев Ю.Г., Осипов А.Н., Ковалев В.В., Кульчицкий В.А. Военная медицина. 2013, – № 3 (28). – С. 94-97.
5. Состояние вегетативной нервной регуляции у больных атопическим дерматитом Аксенова О.И., Марченко В.Н., Монахов К.Н. Вестник современной клинической медицины. 2014. Т. 7. № 4. – С. 15-17.
6. Анализ электроэнцефалограмм на основе исследования изменяющейся во времени структуры локальных максимумов матрицы вейвлет-коэффициентов Туровский Я.А., Кургалин С.Д., Максимов А.В., Семёнов А.Г. Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2012. № 2. – С. 69-73.
7. Анализ энцефалограмм на основе исследования цепочек локальных максимумов скейлограмм Туровский Я.А., Кургалин С.Д., Семёнов А.Г. Цифровая обработка сигналов. 2013. № 2. – С. 20-23.
8. Динамика цепочек локальные максимумов спектров электроэнцефалограмм человека Туровский Я.А., Кургалин С.Д., Семёнов А.Г. Биофизика. 2014. Т. 59. № 1. – С. 185-190.
9. Фактор времени при реализации непрерывного вейвлет-преобразования для анализа сигналов ЭЭГ Туровский Я.А., Кургалин С.Д., Вахтин А.А., Максимов А.В. Информационные технологии в проектировании и производстве. 2012. № 2. С. 61-66.
10. Исследование динамики максимумов локальных вейвлет-спектров вызванных зрительных потенциалов головного мозга Туровский Я.А., Кургалин С.Д., Семенов А.Г. Информационные технологии. 2013. № 10. С. 46-50.
11. Вегетативная регуляция сердечно-сосудистой системы плодов и новорожденных детей, перенесших хроническую внутриутробную гипоксию Туровский Я.А. автореферат диссертации на соискание ученой степени кандидата медицинских наук / Воронежская государственная медицинская академия им. Н.Н. Бурденко. Воронеж, 2005
12. <http://www.ncbi.nlm.nih.gov/nuccore/EF032883.1?report=genbank&log>
13. http://www.atlas-yakutia.ru/weather/climate_russia-III.html
14. <http://www.finam.ru/>

GENERALIZATION OF THE METHOD OF CHAINS OF LOCAL EXTREMA FOR ANALYSIS OF SIGNALS OF DIFFERENT ORIGINS ABSTRACT

Turovsky Y.A., Kurgalin S.D., Vahtin A.A., Borzunov S.V., Belobrodsky V.A.

The method of construction of chains of local extrema for the matrix of squared coefficients of the wavelet transformation, developed and applied previously to analyze EEG signals («the method of chains of local extrema»), is summarized in this paper for the research of all one-dimensional signals or data, regardless of their nature. An essential feature of the development of the method is the formulation of a general rule for adding an existing extremum of the local spectrum to the already formed chain by identifying values of the parameters u and v , which determine the characteristics of the formed Chain of Local Maxima or Chain of Local Minima. Selection of parameters must ensure the maximization of the ratio of extrema numbers (only minimums or only maximums) at the time of the existence of the chain. This approach avoids an excessive union of extrema into the chains with large gaps in the time domain. The proposed universal method has been tested on data of different nature: on the recoded RNA genome sequences of hepatitis C virus, on meteorological data - the average monthly temperatures, on information about stock exchange quotations. It may be concluded that the value of $k / \Delta b$ is sufficiently sensitive criterion for selecting values of u and v necessary for the correct construction of CLM and CLMin signals of different nature.