

УДК 519.688

К.т.н. Г.Г. Знайко, А.А. Голенко (ПАО «ИНЭУМ им. И.С. Брука)

G. Znaiko, A. Golenko

МУЛЬТИМОДАЛЬНЫЙ АНАЛИЗ БИМЕДИЦИНСКОЙ ИНФОРМАЦИИ

MULTIMODAL ANALYSIS OF BIOMEDICAL INFORMATION

Рассматривается алгоритм эмпирической модовой декомпозиции, посредством которого сигнал, содержащий биомедицинскую информацию, может быть разложен на функции, получившие название эмпирических мод. Описываются результаты использования программной реализации алгоритма. Вычисление модовых функций демонстрируется на электроэнцефалограммах. Для сравнения с традиционными методами сигнальной обработки выполнялся спектральный анализ тех же входных сигналов.

The article considers the empirical mode decomposition algorithm by which a signal containing biomedical information can be decomposed into a function, called empirical modes. Describes the results of using the software implementation of the algorithm. Calculation mode functions demonstrated on electroencephalograms. In comparison with conventional methods of signal processing, spectral analysis was carried out the same input signals.

Ключевые слова: мультимодальный анализ, эмпирическая модовая декомпозиция, обработка электроэнцефалограмм.

Keywords: multi-modal analysis, empirical mode decomposition, electroencephalograms processing.

Введение

Надежная и неинвазивная оценка биомедицинской информации на сегодняшний день является актуальной задачей в современной медицинской диагностике. Часто эта оценка сводится к анализу сигнала, полученного в ходе того или иного медицинского

исследования. В основе этого анализа лежит, как правило, преобразование Фурье, при использовании которого предполагается, что сигнал состоит из наложенных синусоидальных колебаний постоянной амплитуды и периода в заданном частотном диапазоне. Такое предположение неизбежно ограничивает надежность и достоверность метода, потому что сигналы, полученные в клинических условиях, часто нестационарны и модулированы нелинейно взаимодействующими процессами в различных временных масштабах.

Для преодоления проблем, связанных с нестационарностью и нелинейностью сигналов, перспективным представляется метод, получивший название мультимодального анализа. В этом методе используется алгоритм эмпирической модовой декомпозиции сигнала, посредством которого происходит разложение сигналов на функции, называемые внутренними, или эмпирическими модами. Метод представляет собой адаптивную итерационную вычислительную процедуру разложения исходных данных непрерывных или дискретных сигналов на эмпирические моды, или внутренние колебания. Каждая мода представляет собой амплитудно-частотную модуляцию в узкой полосе частот, которая может быть связана с конкретным физиологическим процессом. В отличие от гармонического анализа, где модель дискретного или непрерывного сигнала задаётся заранее, эмпирические моды вычисляются в ходе процесса, что позволяет анализировать локальные явления. Поэтому данный метод может быть эффективно использован при обработке нестационарных временных рядов или процессов.

В статье описаны результаты использования программной реализации алгоритма эмпирической модовой декомпозиции. Вычисление модовых функций демонстрируется на электроэнцефалограммах. Для сравнения с традиционными методами сигнальной обработки выполнялся спектральный анализ тех же входных сигналов.

Основываясь на полученных результатах, в дальнейшем предполагается определить, совокупность каких именно методов и алгоритмов, имеющих в основе как эмпирическую

модовую декомпозицию, так и преобразование Фурье, может дать наиболее качественный и точный анализ биологической информации.

1. Преобразование Гильберта

Под преобразованием Гильберта некоторой действительной функции $x(t)$, $-\infty < t < \infty$, будем понимать действительную функцию, определенную как:

$$\tilde{x}(t) = \text{ГН}[x(t)] = x(t) \cdot (1/\pi t),$$

$$\tilde{x}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau,$$

где функция $1/(t - \tau)$ – ядро преобразования Гильберта, а функции $x(t)$ и $\tilde{x}(t)$ – сопряженные по Гильберту [1]. При вычислении в особой точке $t - \tau \Rightarrow$ используется главное значение интеграла по Коши.

Физическим аналогом преобразования Гильберта может служить фазовращатель, который изменяет на $\pi/2$ фазу всех частотных составляющих входного сигнала $x(t)$.

Полученный сигнал $\tilde{x}(t)$ является, таким образом, ортогональным сигналу $x(t)$ и может быть записан в составе комплексного аналитического сигнала $z(t)$: $z(t) = x(t) + j\tilde{x}(t)$.

Подобная запись позволяет определять мгновенные значения фазы и амплитуды сигнала однозначным образом.

Пусть анализируемый сигнал представлен так:

$$x(t) = u(t) \cdot \cos(\omega_0 t + \varphi(t)),$$

где ω_0 – несущая частота, $u(t)$ – огибающая, $\varphi(t)$ – фаза.

Данный сигнал может быть записан в следующей форме:

$$x(t) = a(t) \cdot \cos(\omega_0 t) + b(t) \cdot \sin(\omega_0 t),$$

где $a(t) = u(t) \cdot \cos(\varphi(t))$, $b(t) = u(t) \cdot \sin(\varphi(t))$.

Тогда $u(t) = \sqrt{a^2(t) + b^2(t)}$, $\text{tg}(\varphi(t)) = b(t)/a(t)$.

Применив преобразование Гильберта, получим сопряженный сигнал $\tilde{x}(t)$:

$$\tilde{x}(t) = a(t) \cdot \sin(\omega_0 t) - b(t) \cdot \cos(\omega_0 t),$$

$$z(t) = x(t) + j \cdot \tilde{x}(t).$$

Тогда

$$|z(t)|^2 = x^2(t) + \tilde{x}^2(t) = a^2(t)[\cos^2(\omega_0 t) + \sin^2(\omega_0 t)] + b^2(t)[\cos^2(\omega_0 t) + \sin^2(\omega_0 t)] = u^2(t),$$

откуда могут быть получены выражения для огибающей, мгновенной фазы и мгновенной частоты:

$$u(t) = \sqrt{x^2(t) + \tilde{x}^2(t)},$$

$$\varphi(t) = \arctg[\tilde{x}(t)/x(t)],$$

$$\omega(t) = d\varphi(t)/dt = [\tilde{x}'(t) - x'(t)\tilde{x}(t)]/[\tilde{x}^2(t) + x^2(t)].$$

Для сигнала, содержащего несколько частотных компонент, мгновенная амплитуда будет отображать огибающую интерференции всех этих компонент с максимальными значениями в точках максимумов суммы модулей первых производных компонент и минимальными значениями в точках минимумов этой суммы, что обеспечивает симметричность верхней и нижней огибающих относительно оси времени. Физически значимые мгновенные частоты могут быть получены только для сигналов, компоненты которых локально симметричны относительно нулевого уровня. Тогда произвольный сигнал может быть представлен в виде суммы некоторых колебательных симметричных функций и некоторого остатка (тренда).

2. Базовый метод эмпирической модовой декомпозиции. Преобразование Гильберта–Хуанга

Под преобразованием Гильберта–Хуанга (Hilbert–Huang transform, ННТ) понимается эмпирический метод разложения нелинейного и нестационарного сигнала на эмпирические моды. При этом базис, в котором происходит разложение, выводится адаптивно непосредственно из самого сигнала. Данное преобразование было предложено Норденом Хуангом в 1995 году с целью изучения поверхностных волн тайфунов. В 1998 году Хуанг вместе с коллективом авторов обобщил преобразование на произвольные *временные* ряды [2].

В основе преобразования Гильберта–Хуанга лежит метод эмпирической модовой декомпозиции, посредством которого исходный сигнал раскладывается на набор функций, получивших название «эмпирических мод», или «внутренних модовых функций» (intrinsic

mode functions, IMF). Извлеченные моды с четко определенными мгновенными частотами (точными оценками колебательной информации) эмпирически связаны с конкретными физическими аспектами исследуемого явления. При этом полагается, что любое колебание, входящее в состав исходного сигнала, представляет собой отдельную внутреннюю модовую функцию, имеющую экстремумы, пересечения нуля и накладывающуюся на другие модовые функции, существующие в этот же момент времени.

Каждая такая функция должна обладать следующими свойствами:

- число экстремумов функции и число пересечений нуля не должны отличаться более, чем на единицу;
- в каждой точке функции среднее значение огибающих, определенных локальными максимумами и локальными минимумами, должно быть нулевым.

Первое свойство означает, что локальные максимумы и минимумы будут всегда положительными и отрицательными соответственно, а второе, что мгновенные частоты не будут иметь нежелательных флуктуаций, вызванных асимметричностью сигнала. В отличие от косинусных гармоник модовые функции могут иметь изменяющиеся значения амплитуды и частоты.

Если произвольный сигнал содержит произвольную последовательность локальных максимумов и минимумов (но не менее двух), то он может быть представлен как совокупность внутренних модовых функций и остаточный тренд. Если вместо максимумов в сигнале присутствуют точки перегиба, в которых точка кривой переходит с одной стороны касательной на другую, то перед выполнением модовой декомпозиции производится дифференцирование сигнала.

Для произвольного сигнала $y(t)$ модовая декомпозиция будет заключаться в последовательном вычислении внутренних эмпирических функций $c_j(t)$ и остатков $r(t) = r_{j-1}(t) - c_j(t)$, где $j = 1, 2, 3, \dots, n$ при $r_0 = y(t)$. Результатом декомпозиции является совокупность модовых функций и конечного остатка:

$$x(t) = \sum_{j=1}^n c_j(t) + r_n(t), \text{ где } n - \text{ число внутренних модовых функций.}$$

Рассмотрим процесс декомпозиции модельного сигнала $y(t)$, представленного на рис. 1. Для пояснения алгоритма используются рисунки из [3].

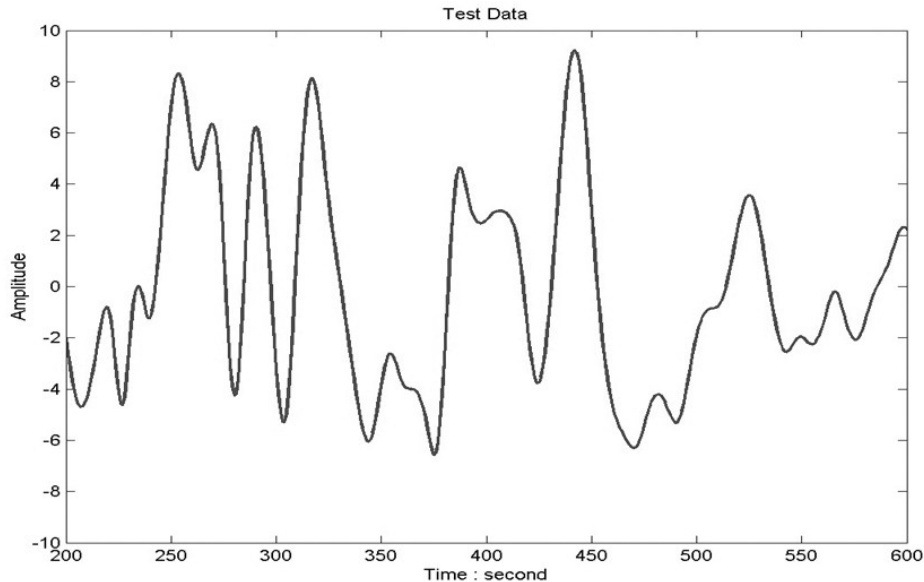


Рисунок 1. Тестовый сигнал $y(t)$

Алгоритм эмпирической декомпозиции включает следующие шаги:

1. Определяется положение всех локальных максимумов и минимумов входного сигнала и значение $y(t)$ в этих точках.

2. С использованием значений локальных экстремумов вычисляются посредством сплайновой интерполяции верхняя и нижняя огибающие сигнала $Xup(t)$ и $Xlow(t)$ (рис. 2, штриховые линии).

3. Точка за точкой определяется среднее значение верхней и нижней огибающей (рис. 2, штрихпунктирная линия):

$$m_1(t) = (Xup(t) + Xlow(t))/2.$$

4. Разность между исходным сигналом $y(t)$ и средним значением огибающих $m_1(t)$ принимается за первое приближение к первой внутренней модовой функции (так называемый процесс *отсеивания*):

$$h_1(t) = y(t) - m_1(t).$$

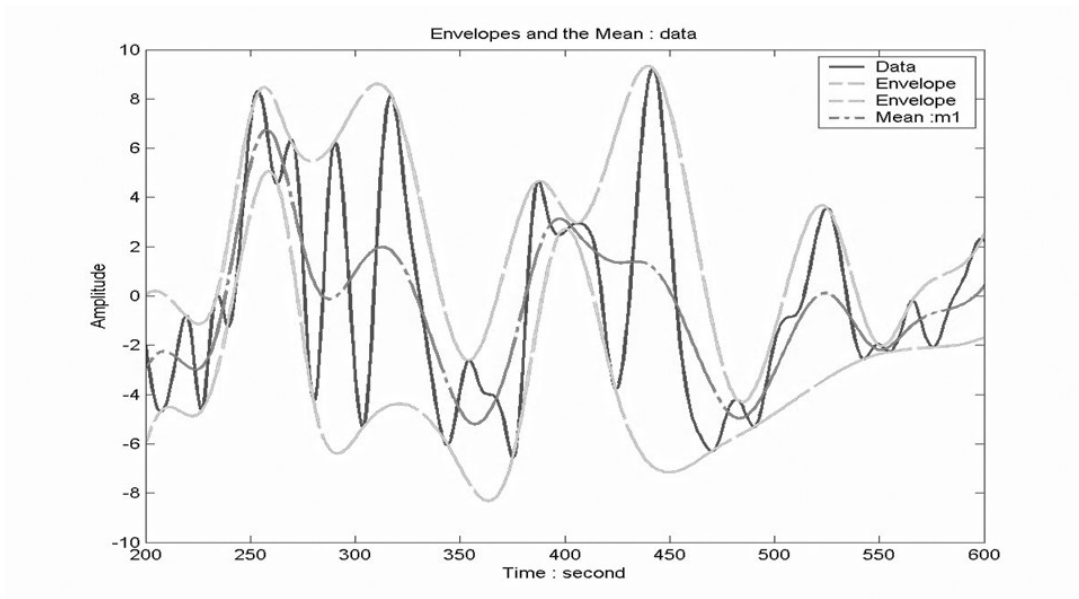


Рисунок 2. Верхняя и нижняя огибающие исходного сигнала $y(t)$ (штриховые линии) и среднее значение огибающих $m_1(t)$ (штрихпунктирная линия)

5. Для полученной функции $h_1(t)$ выполняется следующая проверка:

– если $h_1(t)$ отвечает двум определенным выше свойствам внутренней модовой функции, то модовая функция считается извлеченной, $c_1 = h_1(t)$, $y(t)$ заменяется разностью $r(t) = y(t) - h_1(t)$;

– если $h_1(t)$ не отвечает двум определенным выше свойствам внутренней модовой функции, то $y(t)$ заменяется на $h_1(t)$, повторяются шаги 1–5, и вычисляется второе приближение к первой внутренней модовой функции (рис. 3):

$$h_2(t) = h_1(t) - m_2(t).$$

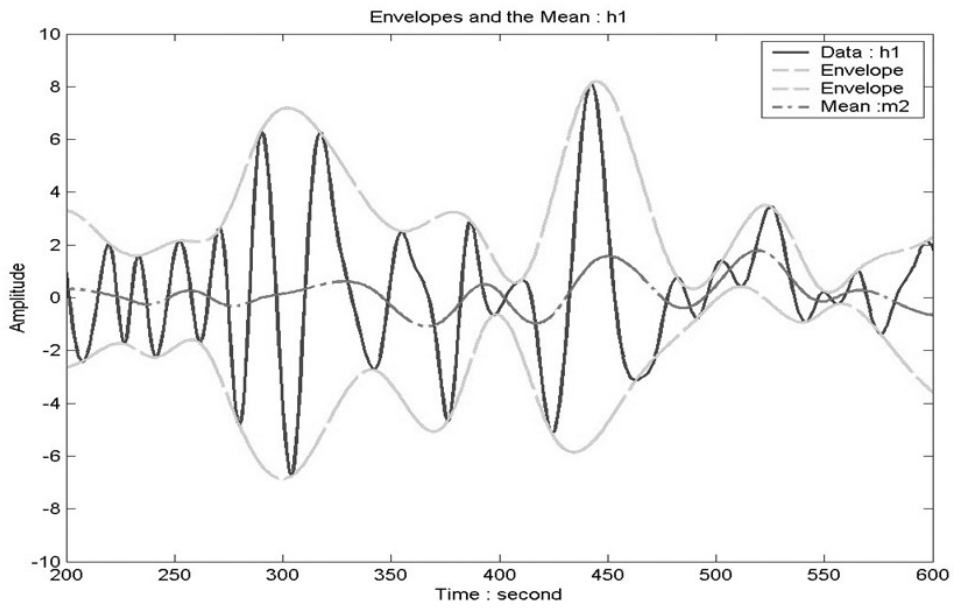


Рисунок 3. Огибающие функции $h_1(t)$ и их среднее значение $m_2(t)$

Последующие шаги выполняются аналогично. На рис. 4 и 5 приведены результаты дальнейшего выполнения алгоритма.

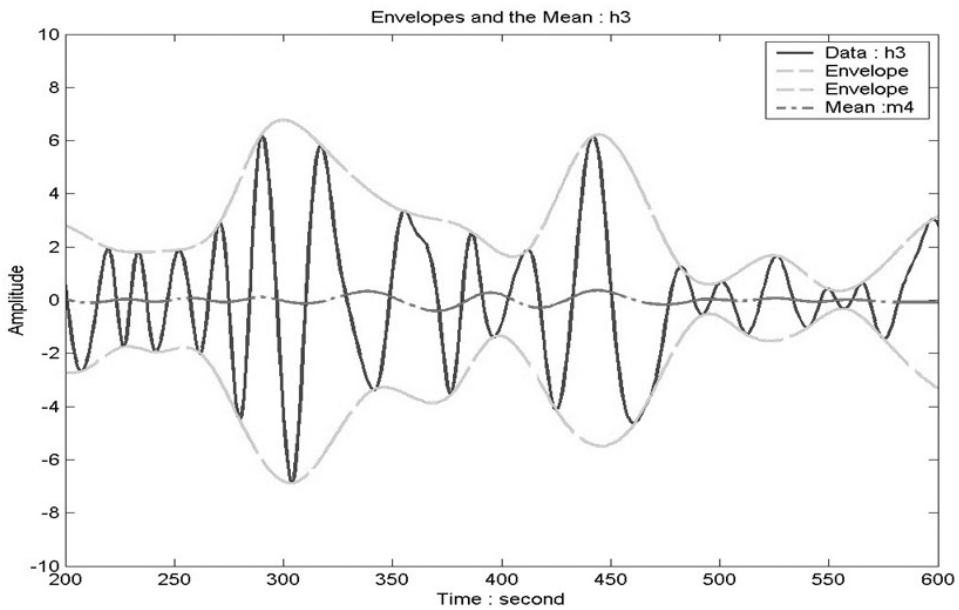


Рисунок 4. Огибающие функции $h_3(t)$ и их среднее значение $m_4(t)$

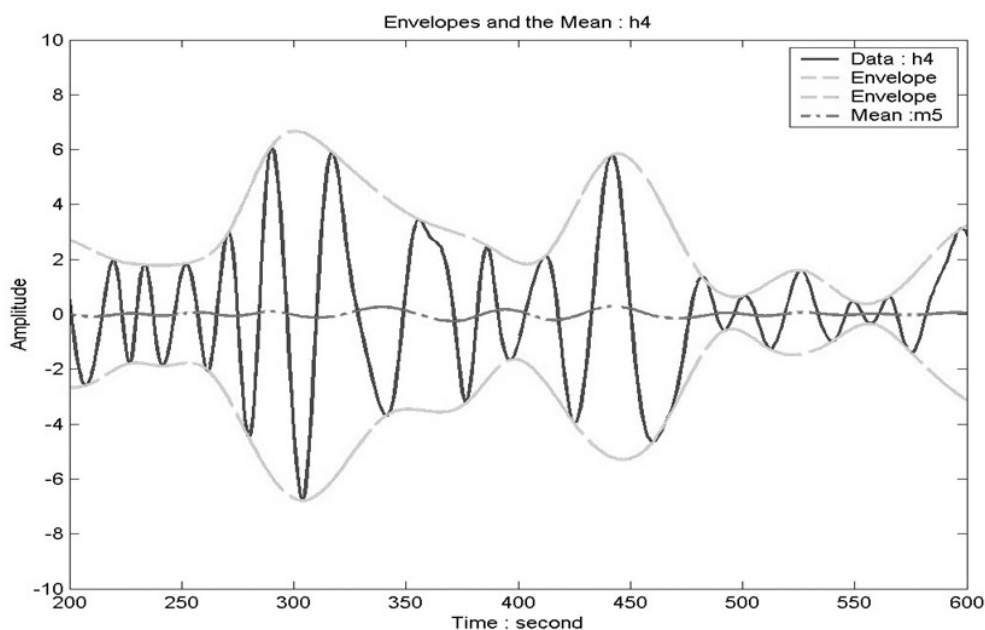


Рисунок 5. Огибающие функции $h_4(t)$ и их среднее значение $m_5(t)$

По мере увеличения итераций, среднее значение огибающих стремится к нулю, а текущая функция h_i – к неизменной форме. Существуют различные критерии останова алгоритма, о которых пойдет речь ниже.

Последнее значение $h_i(t)$ считается наиболее высокочастотной внутренней модовой функцией c_1 , которая входит в данный сигнал и при вычитании которой в сигнале останутся компоненты с более низкими частотами. Полученный остаток декомпозируется аналогичным образом, в результате чего выделяется вторая внутренняя модовая функция c_2 , и так далее. В результате декомпозиции сигнал $y(t)$ может быть записан следующим образом:

$$y(t) = \sum_{i=0}^N c_i(t) + r_N(t),$$

где N – число внутренних мод; $r_N(t)$ обозначает окончательный остаток, который может быть интерпретирован как постоянная составляющая сигнала; $c_j(t)$ – внутренние моды, ортогональные друг другу.

Критерии останова

Рассмотрим два необходимых вида критериев останова для приведенного выше алгоритма: критерии останова *отсеивания*, когда для очередной функции h_i принимается решение о том, является ли она внутренней модовой функцией, и критерии останова

декомпозиции, когда определяется, содержит ли полученный после вычитания очередной внутренней модовой функции остаток сигнала какую-либо полезную информацию.

Критерии останова процесса отсеивания. Исторически первым использовался метод SD (1998 год), в котором в качестве критерия служила квадратичная разность между двумя последовательными операциями приближения:

$$\delta = \sum_k \left[\frac{|h_{i-1}(k) - h_i(k)|^2}{h_{i-1}^2(k)} \right].$$

В 2003 году Quek с коллегами предложил более эффективный критерий останова переменной меры [4]:

$$\delta = \sum_k |h_{i-1}(k) - h_i(k)|^2 / \sum_k h_{i-1}^2(k).$$

Нередко применяются критерии, выводимые опытным путем и зависящие от характера обрабатываемых сигналов. Часто используется и двойной критерий, например: выполнять, пока не достигнута δ , но не более n итераций; n , как правило, выбирается равным 6–8 повторениям. В случае слишком строгого критерия останова вероятно появление внутренних модовых функций, не несущих полезной информации, а при недостаточно строгом – появление отрицательных мгновенных частот в модовых функциях.

Критерии останова процесса декомпозиции. К таким критериям можно отнести следующие утверждения:

– после вычитания очередной модовой функции остаток сигнала $r_k(t)$ имеет менее двух экстремумов;

– после вычитания очередной модовой функции остаток сигнала $r_k(t)$ по своим значениям не представляет больше интереса для анализа.

Кроме того, поскольку суммирование всех полученных внутренних модовых функций должно давать в результате исходный сигнал, можно в качестве критерия останова декомпозиции использовать относительную среднеквадратичную погрешность

реконструкции сигнала. Текущий остаток $r_k(t)$ при этом не учитывается.

Итерация, в ходе которой очередная выделенная внутренняя модовая функция увеличивает погрешность реконструкции, также может служить сигналом к завершению декомпозиции. Полученный в результате декомпозиции остаток сигнала должен быть максимально «выпрямленным», т.е. содержать минимальное количество локальных максимумов и минимумов.

3. Обработка электроэнцефалограмм посредством эмпирической модовой декомпозиции

В качестве исходных электроэнцефалограмм использовались сигналы, полученные разработчиками системы VCI2000, предназначенной для проведения исследований в области интерфейсов «мозг–компьютер» [5].

Авторами была выполнена запись 64-канальной электроэнцефалограммы 109 добровольцев, выполнявших последовательности моторных и умственных задач. Использовалась международная система 10–10 (исключены электроды Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9 и P10), как показано на рис. 6.

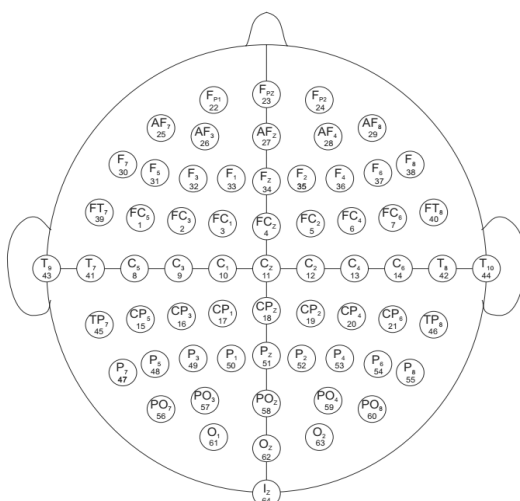


Рисунок 6. Расположение электродов

Каждый испытуемый проходил пять прогонов: два одномоментных базовых прогона (один с закрытыми глазами, один с открытыми), и три двухминутных прогона, в каждом из

которых выполнялось четыре задачи:

1) цель появляется возле левого или правого края монитора, испытуемый сжимает и разжимает соответствующий кулак. Когда цель исчезает, испытуемый расслабляется;

2) цель появляется возле левого или правого края монитора, испытуемый воображает сжатие соответствующего кулака. Когда цель исчезает, испытуемый расслабляется;

3) цель появляется возле верхнего или нижнего края монитора, испытуемый сжимает либо оба кулака (если цель вверху), либо обе ступни (если цель внизу). Когда цель исчезает, испытуемый расслабляется;

4) цель появляется возле верхнего или нижнего края монитора, испытуемый воображает сжатие кулаков (если цель вверху), либо ступней (если цель внизу). Когда цель исчезает, испытуемый расслабляется.

На рис. 7 представлен фрагмент электроэнцефалограммы испытуемого №1, записанный в базовом прогоне с закрытыми глазами. На рис. 7а представлен спектр этого сигнала.

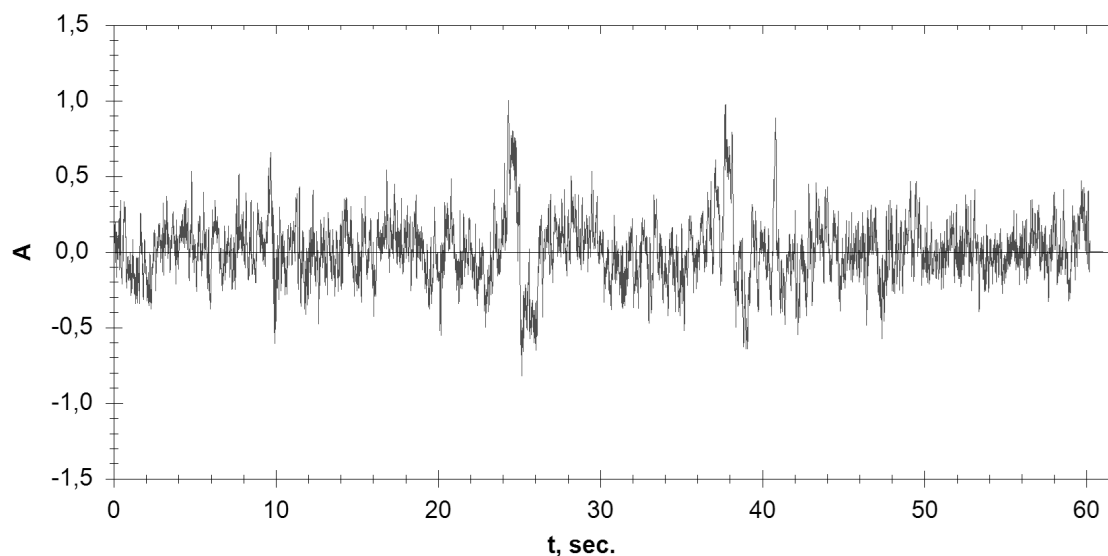


Рисунок 7. Электроэнцефалограмма испытуемого №1. Состояние с закрытыми глазами

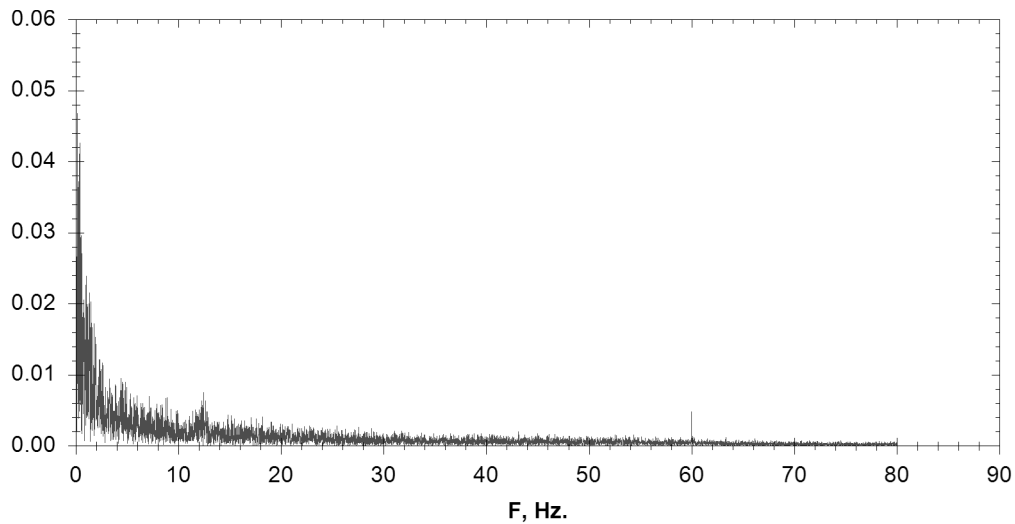


Рисунок 7а. Спектр исходной электроэнцефалограммы

Использование модовой декомпозиции позволяет анализировать и обрабатывать отдельные участки полного спектра, например, исключать частоту сетевого напряжения или биологические процессы, не несущие полезной диагностической информации. Информативные функции, напротив, можно выделять и обрабатывать отдельно с целью повышения локализации и снижения вычислительной сложности.

На рис. 8 представлены внутренние моды, полученные в ходе выполнения эмпирической модовой декомпозиции.

На рис. 8а представлен спектр моды Mode 0. Хорошо видна частота сетевого напряжения – 60 Гц в Северной Америке. Поскольку измерения выполнялись с закрытыми глазами в состоянии покоя, частоты, соответствующие обработке визуальной информации, активности, концентрации или тревожному мышлению на спектре явно не выделены, хотя можно заметить небольшой подъем в районе 10–20 Гц, соответствующих бета-волнам.

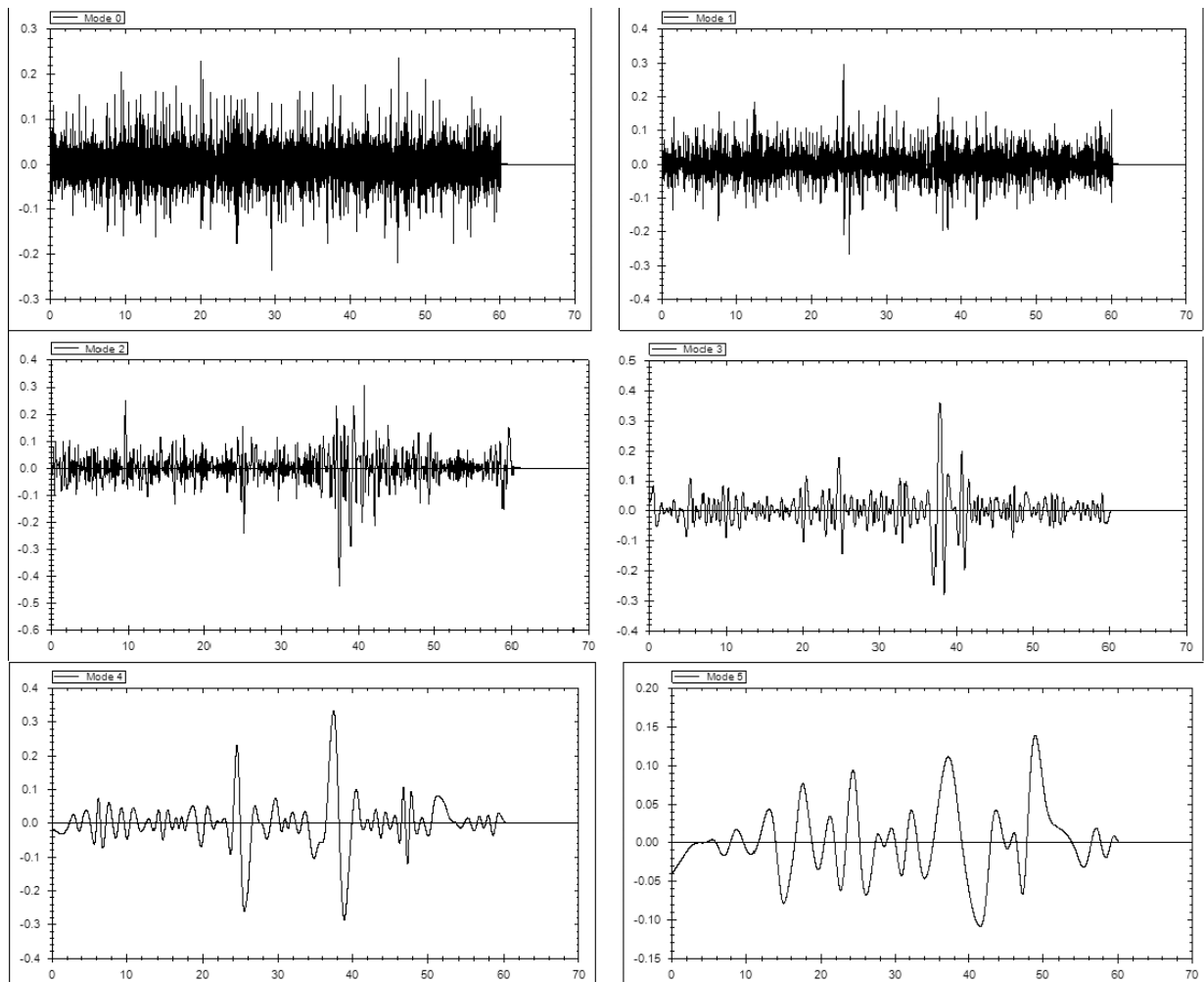


Рисунок 8. Разложение электроэнцефалограммы на внутренние модовые функции

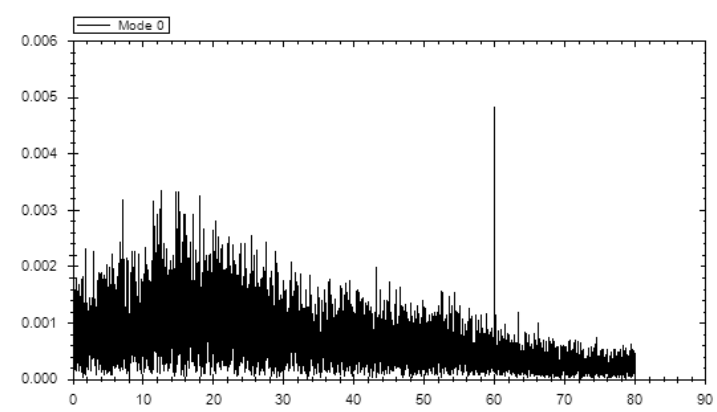


Рисунок 8а. Спектр моды Mode 0

Спектр следующей моды – Mode 1 представлен на рис. 8б. Виден подъем в районе 7–13 Гц – диапазон, соответствующий альфа-волнам, имеющим наибольшую амплитуду в состоянии спокойного бодрствования с закрытыми глазами.

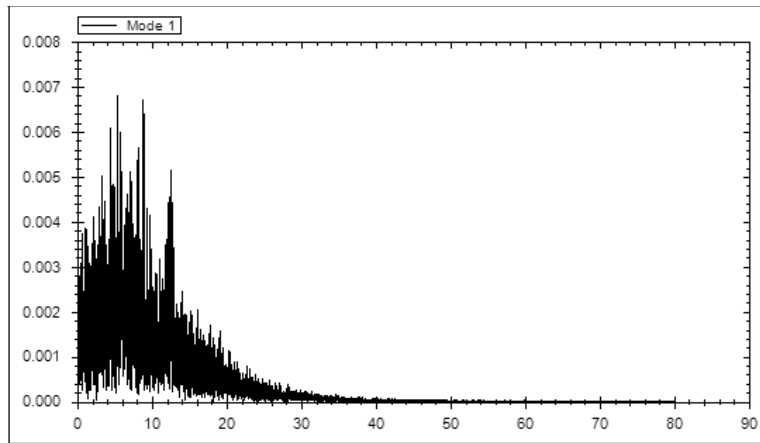


Рисунок 8б. Спектр моды Mode 1

Спектры следующих двух мод примерно соответствуют диапазонам тета- и дельта-волн, последние две моды содержат низкочастотные компоненты, частотой менее 1 Гц.

Следует отметить, что преобразование Фурье и вычисление спектров использовались исключительно для визуального представления полученных результатов в данной статье, в процессе вычисления внутренних мод никакие дополнительные преобразования сигнала не использовались.

Предположим, что полученный сигнал зашумлен в ходе записи или же каким-либо параллельно протекающим биологическим процессом. Рассмотрим два варианта: белый шум (рис. 9а) и модуляция двумя гармоническими сигналами с частотами 50 и 120 Гц (рис. 9б).

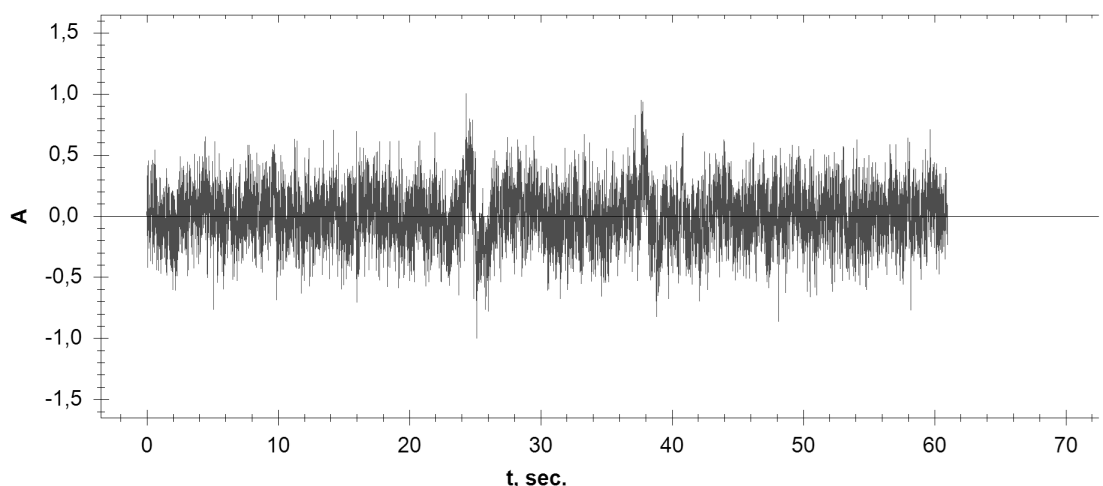


Рисунок 9а. Электроэнцефалограмма с белым шумом

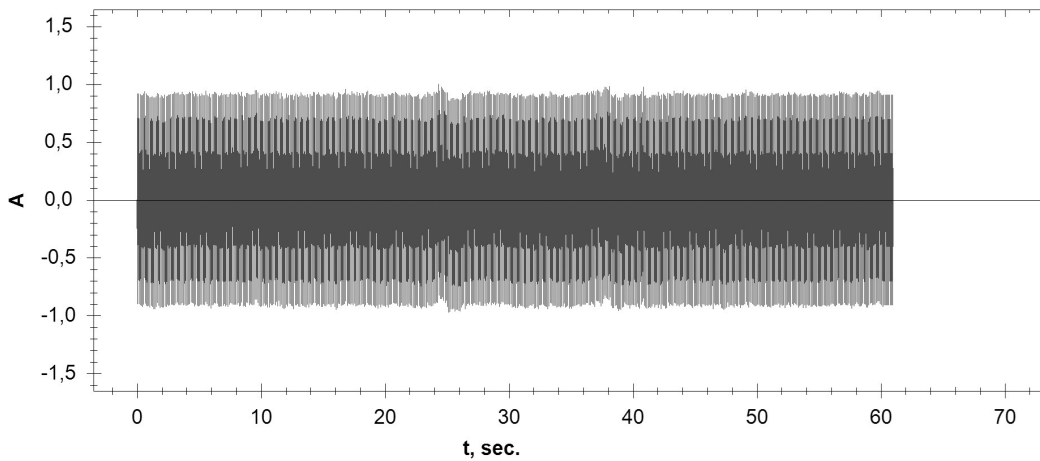


Рисунок 9б. Электроэнцефалограмма, модулированная гармоническими сигналами

После выполнения эмпирической модовой декомпозиции был получен набор мод, некоторые из которых представлены на рис. 10а и 10б.

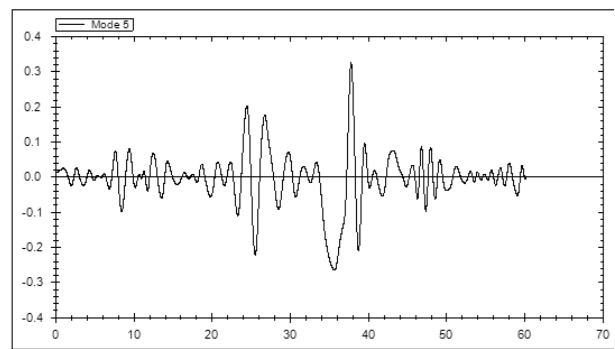
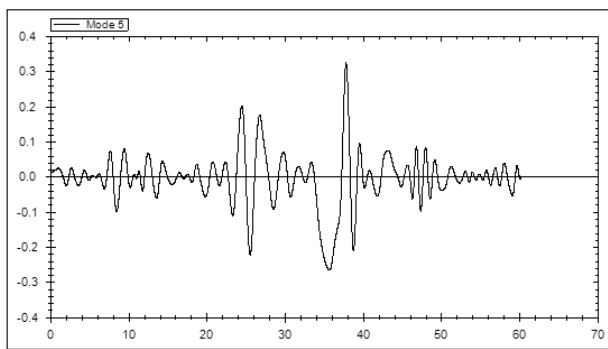


Рисунок 10а. Пятая мода, выделенная из оригинальной электроэнцефалограммы и из электроэнцефалограммы с белым шумом

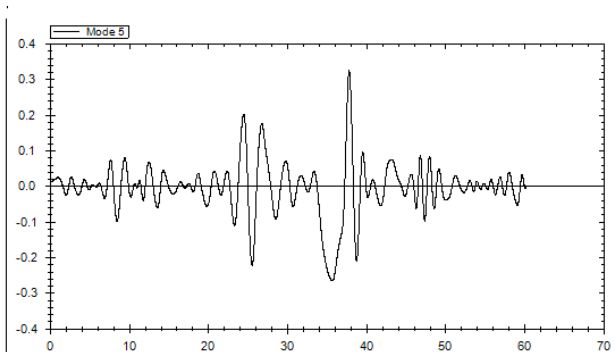
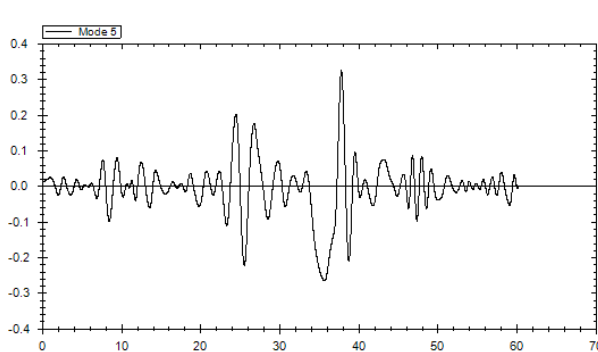


Рисунок 10б. Пятая мода, выделенная из оригинальной электроэнцефалограммы и из электроэнцефалограммы, модулированной гармоническими сигналами

Приведенные выше результаты также были получены без использования каких-либо фильтров, оконных функций и прочих средств обработки сигналов. Применялся только алгоритм эмпирической модовой декомпозиции.

Заключение

Описанный выше метод эмпирической модовой декомпозиции представляется перспективным для преодоления проблем, связанных с нестационарностью и нелинейностью сигналов, например, в случаях, когда сигнал не может быть адекватно представлен совокупностью гармонических функций. Все приведенные выше результаты были получены без использования дополнительных сигнальных преобразований, таких как фильтрация, использование окон и т.п. Перед выполнением алгоритм позволяет явно задавать число итераций при отсеивании и число полученных модовых функций, что дает возможность регулировать сложность вычислений и затрачиваемые вычислительные ресурсы, а адаптивные критерии останова помогают предотвратить избыточные прогоны, не дающие значимого результата. Возможность работы только с функциями, содержащими диагностически ценную информацию, также снижает сложность вычислений и упрощает последующую обработку. После уточнения требований, специфических для биомедицинской информации, и при параллельном использовании традиционных методов анализа сигналов метод эмпирической модовой декомпозиции, как ожидается, должен повысить точность и надежность неинвазивной медицинской диагностики.

Другая область использования метода связана с системами типа интерфейс мозг–компьютер. Применение мультимодального анализа должно облегчить выделение в сигнале паттернов, характерных для тех или иных моторных или умственных операций. Для определения таких паттернов используются, как правило, либо сложные статистические методы, требующие больших вычислительных затрат, либо различные методы сегментации, предполагающие, что исследуемый сигнал уже есть в наличии

целиком. Эмпирическая модовая декомпозиция, в свою очередь, позволяет регулировать сложность вычислений и может работать только с той частью сигнала, которая зарегистрирована на данный момент.

Литература

1. Дж. Бендат, А. Пирсол. Прикладной анализ случайных данных. – М.: Мир, 1989.
2. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, Henry H. Liu. The Royal Society, 1998.
3. An Introduction to Hilbert-Huang Transform: A Plea for Adaptive Data Analysis. Norden E. Huang. Research Center for Adaptive Data Analysis. National Central University, 2008.
4. Quek S., Tua P. and Wang Q. Detecting anomalies in beams and plate based on the Hilbert–Huang transform of real signals. Smart Materials and Structures 12, 2003, pp. 447-460.
5. Schalk G., McFarland D.J., Hinterberger T., Birbaumer N., Wolpaw J.R. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. IEEE Transactions on Biomedical Engineering 51(6): 1034-1043, 2004.