

## Алгоритм локальной фильтрации низкочастотных шумов с большой амплитудой в данных электроэнцефалограмм

Е. Г. Климова<sup>1,\*</sup>, С. Б. Медведев<sup>1</sup>, А. Н. Савостьянов<sup>2,3</sup>

<sup>1</sup>Институт вычислительных технологий СО РАН, Новосибирск, Россия

<sup>2</sup>НИИ физиологии и фундаментальной медицины СО РАН, Новосибирск, Россия

<sup>3</sup>Новосибирский государственный университет, Россия

\*e-mail: klimova@ict.nsc.ru

При исследовании функций мозга человека в качестве одного из инструментов используется измерение сложного колебательного электрического процесса, который регистрируется с помощью электроэнцефалографа. Одной из проблем в работе с данными электроэнцефалограмм (ЭЭГ) является наличие в них “шума”, связанного как с поведением пациента, так и с внешними воздействиями. В работе рассматривается алгоритм фильтрации в данных ЭЭГ низкочастотных шумов с большой амплитудой, вызванных различными источниками. Алгоритм основан на фурье-анализе данных и может применяться локально (в заданном временном интервале). Отличительным свойством предлагаемого алгоритма является отсутствие искажения фазы волны (с точностью до вычислительной погрешности метода), что особенно важно для анализа сетевых взаимодействий в головном мозге.

*Ключевые слова:* электроэнцефалограмма, фильтрация шумов, спектральный анализ.

### Введение

Мониторинг состояния здоровья человека включает в себя обработку данных измерений характеристик функционирования различных систем организма. При исследовании функций мозга человека в качестве одного из инструментов используется измерение сложного колебательного электрического процесса, который регистрируется с помощью электроэнцефалографа [1]. Одной из проблем в работе с данными электроэнцефалограмм (ЭЭГ) является наличие в этих данных “шума”, связанного как с “нежелательным” поведением пациента (например, моргание, зевота), так и с внешними воздействиями.

Современная обработка данных ЭЭГ включает целый комплекс процедур. В монографии [2] рассмотрены подходы к фильтрации шумов в данных биомедицинских сигналов, в том числе и в данных ЭЭГ. При этом указано, что применение любого частотного фильтра требует знания спектральных характеристик шума. Для выделения артефактов применяются процедуры фильтрации, реализованные в виде пакетов программ,

как общих математических, таких как MatLab, так и специализированных, предназначенных для работы с данными ЭЭГ, например EEGLAB (<http://sccn.ucsd.edu/eeglab>). При обработке данных ЭЭГ широко применяется метод главных компонент [3], который основан на теории информации и позволяет успешно выделять артефакты в данных ЭЭГ. При этом наряду с достоинствами он обладает рядом недостатков, в частности искажает фазу волны.

Большое количество работ посвящено фильтрам, основанным на статистических свойствах шума, — фильтры Винера [2] и Калмана [4]. Их использование требует оценки статистических характеристик шума, а также знания оператора перехода между шагами по времени (фильтр Калмана), что является отдельной проблемой. Описание используемых процедур математической обработки данных ЭЭГ можно найти, например, в работе [5], где содержатся также ссылки на статьи по обработке данных ЭЭГ.

В настоящей работе исследовалась одна из разновидностей помех, зарегистрированных в данных ЭЭГ, а именно низкочастотный шум с высокой амплитудой. Анализ данных ЭЭГ показал, что в ряде случаев низкочастотный аддитивный шум присутствует не на всем заданном временном интервале, а сильно локализован во времени. Например, низкочастотные возмущения возникают в результате моргания пациента. Разработан метод, позволяющий диагностировать наличие низкочастотного шума автоматически и выделять характерные значения частот. Предложенный алгоритм основан на фурье-анализе данных и может применяться локально (в заданном временном интервале), его отличительным свойством является отсутствие искажения фазы волны (с точностью до вычислительной погрешности метода).

Для многих задач обработки ЭЭГ нарушение фазовых отношений несущественно. В задаче локализации источника активности, задаче анализа частотно-временной динамики, задаче оценки амплитуд вызванных потенциалов не учитываются фазовые отношения, и поэтому существующие методы удаления артефактов всех устраивают. Однако фазовые отношения существенны при построении “сетей мозга” (brain oscillatory networks). Такой подход впервые предложен в работе [6]. Сама эта работа была выполнена без применения ЭЭГ, но на ее основе возникло новое направление в исследовании функциональных сетей. Предлагаемый алгоритм особенно важен именно для анализа сетевых взаимодействий в головном мозге.

## 1. Данные ЭЭГ

Для тестирования предложенного алгоритма использовались данные ЭЭГ, которые регистрировались на базе НИИ физиологии и фундаментальной медицины при помощи 128-канального энцефалографа фирмы Brain Products (Germany) с полосой пропускания 0.1–100 Гц, частота оцифровки сигнала 1000 Гц. Электроды (126 ЭЭГ + VEOG) располагались согласно международной схеме 10-10% с референтом Cz и заземляющим электродом AFz. В ходе эксперимента электроэнцефалограммы регистрировались в экспериментальных условиях парадигмы стоп-сигнал [7], где испытуемый либо нажимал на одну из двух кнопок после появления целевого сигнала (условие Go), либо должен был подавить уже начатое движение, если после целевого сигнала следовал запрещающий сигнал (условие Stop). Каждому испытуемому предлагалось 130 заданий, из которых 30 относились к условию Stop и 100 — к условию Go.

В численных экспериментах рассматривалась выборка из 85 данных ЭЭГ, где содержатся 4000 измерений для 119 датчиков на временном интервале 4 с. Перед проведением

спектрального анализа и фильтрации из всех данных вычиталось среднее по выборке значение.

## 2. Методика выделения низкочастотной составляющей измерений

На рис. 1 представлен фрагмент ЭЭГ с выраженным низкочастотным шумом, полученный с помощью пакета программ EEGLAB, шкала сверху означает номер элемента выборки данных ЭЭГ, шкала снизу — время, слева — наименования датчиков. Как видно из рисунка, низкочастотный шум обладает высокой амплитудой, на порядок большей, чем амплитуда основных колебаний.

Рассмотрим подход к разделению высоко- и низкочастотных колебаний, широко применяемый в теории турбулентности [8, 9]. Пусть данные ЭЭГ в момент времени  $t$  представляют собой сумму

$$eeg(t) = eeg_0(t) + L(t), \quad (1)$$

где  $eeg_0(t)$  — “истинные” данные (без шума), а  $L(t)$  — низкочастотный шум.

Как показано в [8], операция осреднения по времени  $\frac{1}{T} \int_{-T/2}^{T/2} eeg(t+\tau) d\tau$  представляет собой фильтр высоких частот. Если представить  $eeg(t)$  в виде вещественного ряда Фурье на заданном временном интервале  $[-l, l]$

$$eeg(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \cos \frac{n\pi t}{l} + b_n \sin \frac{n\pi t}{l} \right),$$

то для функций  $\sin \omega t$  и  $\cos \omega t$  операция осреднения влечет за собой умножение на множитель  $u_\omega = \frac{2}{\omega T} \sin \frac{\omega T}{2}$ . Пусть в представлении (1) “характерной частотой”  $L(t)$  является  $\omega_1$ , а  $eeg_0(t)$  —  $\omega_2$ . Требуется подобрать такой интервал осреднения, чтобы

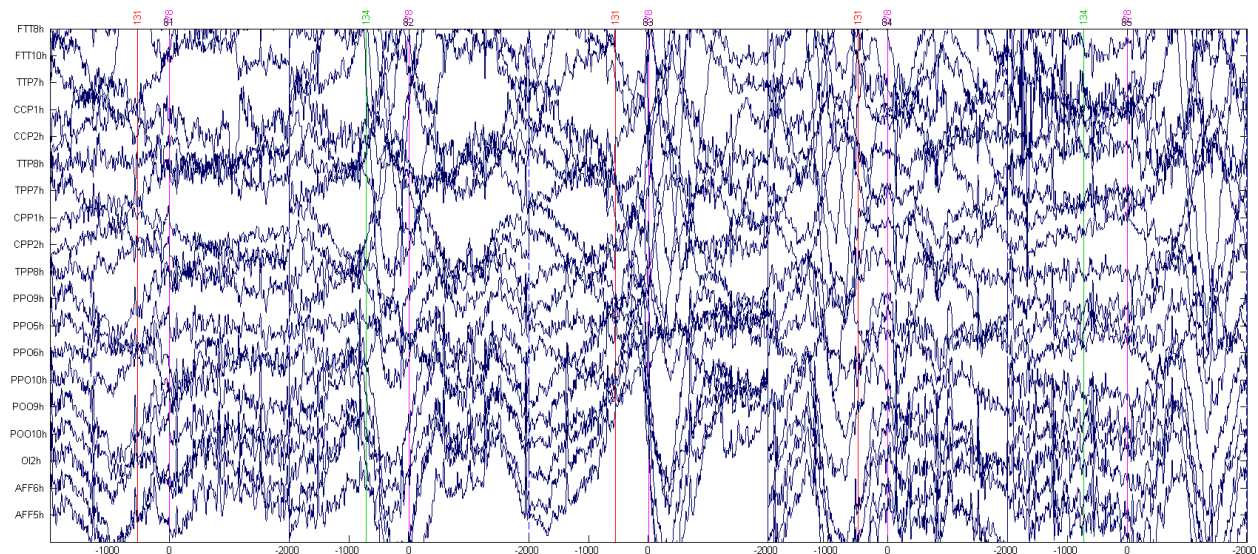


Рис. 1. Данные ЭЭГ, содержащие низкочастотный шум: элементы выборки 81–84

множитель  $u_\omega$  при  $eeg_0(t)$  был минимальным, а при  $L(t)$  — максимальным. Как показано в [9], такой оптимальный период осреднения может быть выбран из условия

$$\omega_1 T = \frac{1}{\omega_2 T} = \varepsilon \ll 1.$$

В этом случае он равен  $T = 1/\sqrt{\omega_1 \omega_2}$ .

Таким образом, в результате применения операции осреднения в момент времени  $t$  мы получаем оценку низкочастотной составляющей  $L(t)$ , поскольку высокочастотная составляющая удаляется при фильтрации, и  $eeg_0(t) = eeg(t) - L(t)$ . Из вышесказанного очевидно, что данный метод корректирует только амплитуду сигнала, сохраняя фазу неизменной.

Так как измерения ЭЭГ представляют собой дискретный временной ряд длины  $N$ , вместо осреднения с помощью интеграла берется аппроксимация интеграла квадратурными формулами [10].

В дискретном случае в заданный момент времени  $t_k = \Delta t k$  измерение может быть представлено в виде конечного дискретного ряда Фурье

$$eeg(t_k) = \frac{a_0}{2} + \sum_{n=1}^p \left( a_n \cos \frac{2\pi nk}{N} + b_n \sin \frac{2\pi nk}{N} \right), \quad (2)$$

где  $p = N/2$ .

### 3. Фурье-анализ данных ЭЭГ

Для оценки оптимального интервала осреднения требуется знание характерных частот низко- и высокочастотной составляющих  $\omega_1$  и  $\omega_2$ . Для оценки этих частот проведен анализ с помощью разложения в ряд Фурье (2) данных ЭЭГ без шума и с низкочастотным шумом. Рассмотрены четыре случая ЭЭГ без выраженного шума и четыре случая с низкочастотным шумом (этот фрагмент данных приведен на рис. 1).

На рис. 2 представлены коэффициенты разложения в ряд Фурье  $E_n = \sqrt{a_n^2 + b_n^2}$  для первой группы, а на рис. 3 — для второй. Одновременно с этим по аналогии с исследованиями, приведенными в [9], рассматривалась зависимость  $E_n \omega_n$  от  $\ln \omega_n$ , где  $\omega_n$  — частота. На рисунке в этих переменных площадь, ограниченная кривой, дает среднее значение в интервале частот в силу формулы

$$\int E(\omega) d\omega = \int \omega E(\omega) d(\ln \omega).$$

Это позволяет по максимальному значению функции  $E_n \omega_n$  оценить значение частоты, вокруг которого сосредоточены основные колебания. На рис. 4 и 5 представлены графики этой зависимости для двух выбранных групп измерений.

Для рассматриваемых данных согласно формуле (2) частота  $\omega_n = \frac{2\pi n}{T_0}$ , где  $T_0 = 4$  с. Графики на рис. 2 и 3 дают возможность оценить изменение амплитуды в случае данных с шумом, в то время как графики на рис. 4 и 5 позволяют выделить характерную частоту низкочастотной составляющей. Сравнивая графики на рис. 4 и 5, легко видеть существенные различия в поведении функции  $E_n \omega_n$  в низких частотах. Для данных без

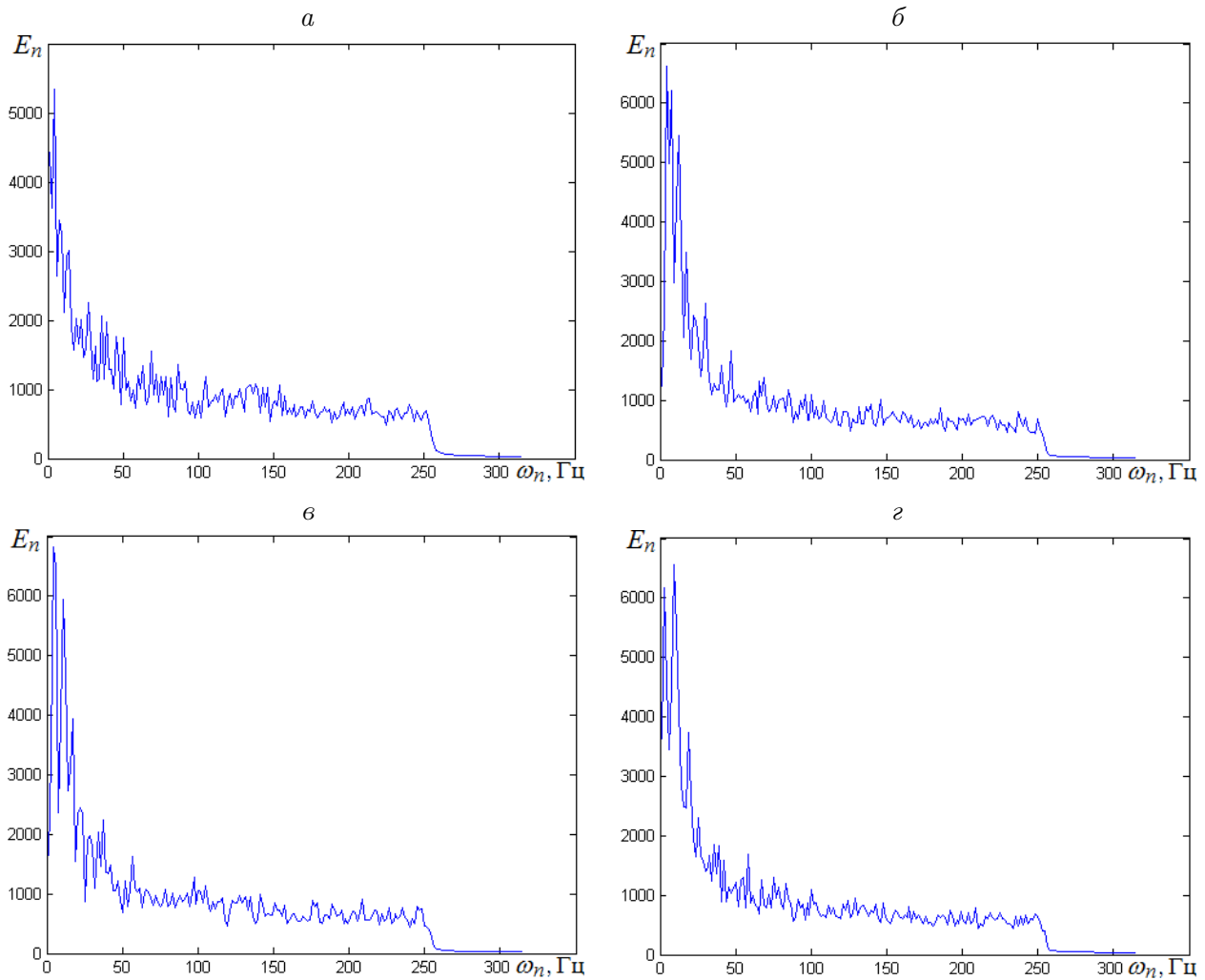


Рис. 2. Коэффициенты разложения в ряд Фурье  $E_n$  в зависимости от частоты  $\omega_n$  для элементов выборки 1–4 (а–г соответственно)

низкочастотного шума (см. рис. 4) кривые в “среднем” достаточно плавные, но в присутствии низкочастотного шума имеются явные отклонения в “среднем” от монотонного поведения незашумленных сигналов.

В качестве частоты  $\omega_1$  может быть взято значение, при котором достигается максимум  $E_n \omega_n$  в области низких частот  $\omega \in [1, 6]$  (в логарифмическом масштабе  $\ln \omega \in [0, 1.79]$ ). При проведении процедуры фильтрации для  $\omega_2$  может быть принято характерное значение  $\omega_2 = 35$  (оно соответствует среднему значению частоты для высокочастотной части спектра).

#### 4. Алгоритм локальной фильтрации низкочастотного шума

Данные ЭЭГ представляют собой массив  $D(nn, nt, nx)$ , где  $nn$  — размер выборки;  $nx$  — количество датчиков;  $nt$  — число временных шагов. Алгоритм локальной фильтрации условно можно разбить на три шага.

*Шаг 1.* Производится спектральный анализ — разложение в ряд Фурье массива значений временного ряда для элементов выборки по каждому датчику и вычисляется

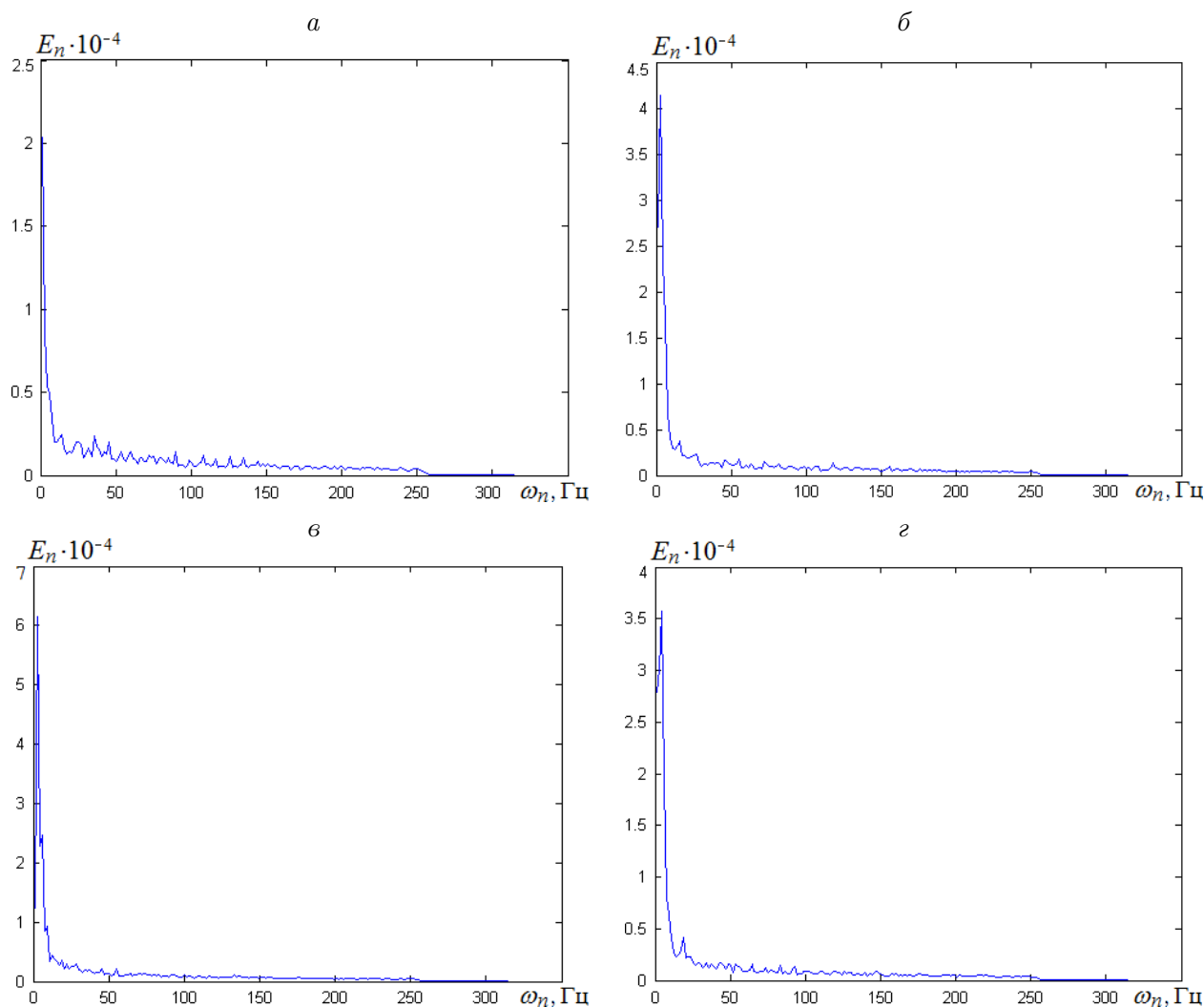


Рис. 3. Коэффициенты разложения в ряд Фурье  $E_n$  в зависимости от частоты  $\omega_n$  для элементов выборки 81–84 (а–г соответственно)

среднее значение по всем датчикам массива коэффициентов Фурье  $E_n$ . Затем вычисляется максимум функции  $E_n \omega_n$  в диапазоне  $\omega \in [1, 6]$ . Значение  $\omega$ , при котором достигается максимум, будет соответствовать “характерной” частоте в заданном диапазоне  $\omega$ .

*Шаг 2.* Производится оценка периода осреднения по формуле  $T = 1/\sqrt{\omega_1 \omega_2}$ . При этом характерное значение частоты  $\omega_2$  может быть взято как среднее по выборке в высокочастотной части спектра. В выполненных нами экспериментах  $\omega_2 = 35$ . Отметим, что возможна модификация алгоритма с автоматизацией выбора параметра  $\omega_2$ . Процедура оценки интервала осреднения может проводиться по средним значениям всех датчиков либо отдельно для каждого датчика, в этом случае существенно возрастает время вычислений.

*Шаг 3.* Для временного ряда при заданных номерах выборки и номерах датчика осуществляется процедура фильтрации данных. В каждый момент времени в результате осреднения в интервале  $(t - T/2, t + T/2)$ , где  $T$  — период осреднения, вычисленный на втором шаге процедуры, имеется информация о низко- и высокочастотной составляющих. Рассматривается соотношение между осредненным значением и исходным и принимается решение о преобразовании этого исходного значения. Решение

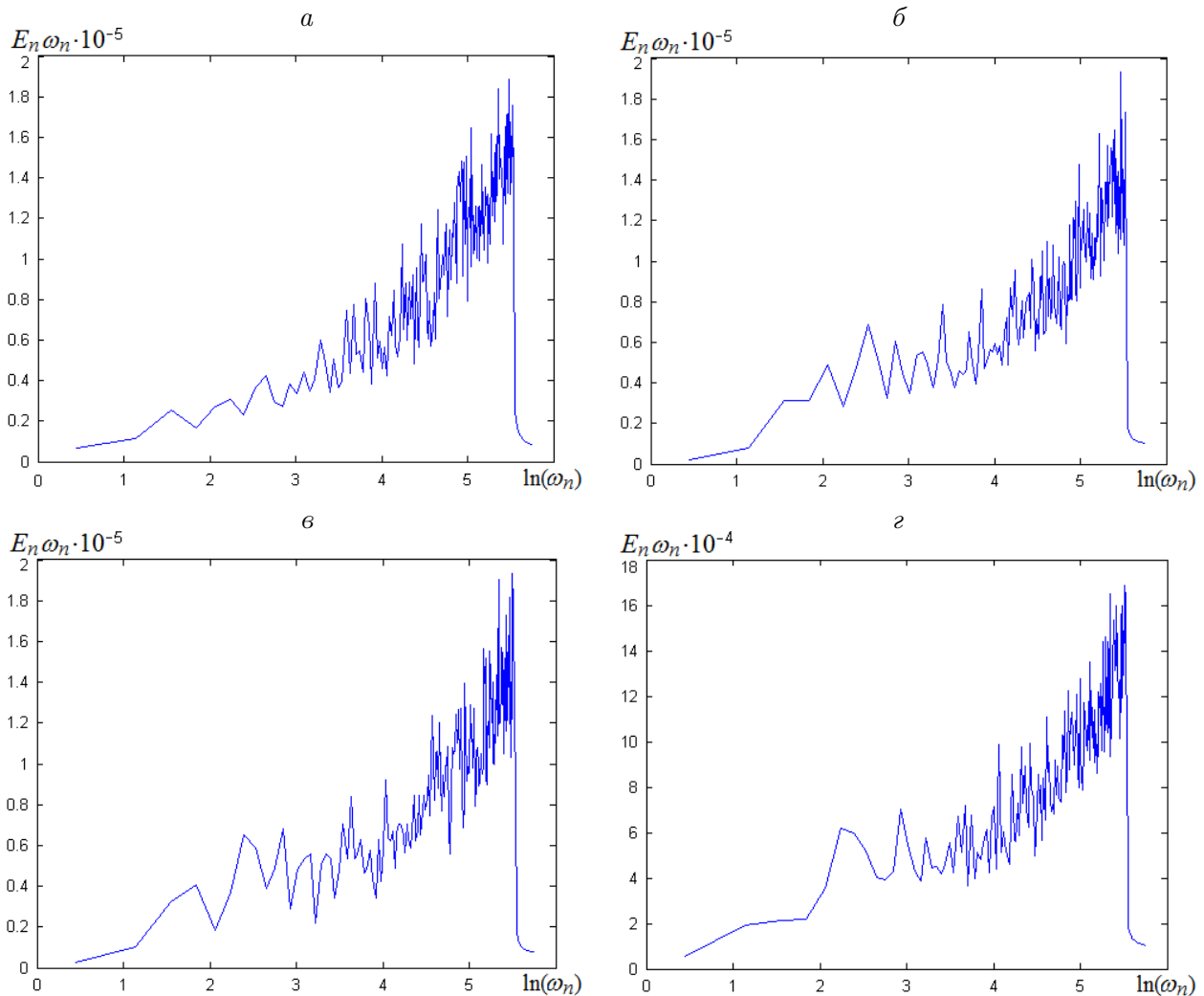


Рис. 4. Зависимость  $E_n \omega_n$  от  $\ln \omega_n$  для 1–4-го элементов выборки (а–г соответственно)

о проведении удаления низкочастотного шума может быть принято на основе критерия  $\tilde{u}_\omega = \left| \overline{eeg(t)} \right| / |eeg(t)| \geq u_c$ , где  $u_c$  — заданная константа. Критерий характеризует долю низкочастотной составляющей в данных. В расчетах эмпирическим путем взято значение параметра  $u_c = 0.4$ , что соответствует некоторым характерным масштабам возмущений. В случае выполнения заданного критерия производится коррекция значения  $eeg(t)$ .

Как видно из данных ЭЭГ, низкочастотный шум может присутствовать не на всем временном промежутке, а локально. Предложенная процедура может работать как поточно, так и в заданном временном промежутке, и таким образом мы можем избежать неоправданного искажения “незашумленных” участков ЭЭГ.

## 5. Численные эксперименты

С применением предложенного алгоритма выполнялась фильтрация данных ЭЭГ, их первоначальный вид приведен на рис. 1. Для вычисления коэффициентов разложения в ряд Фурье использовалась процедура быстрого преобразования Фурье из библиотеки

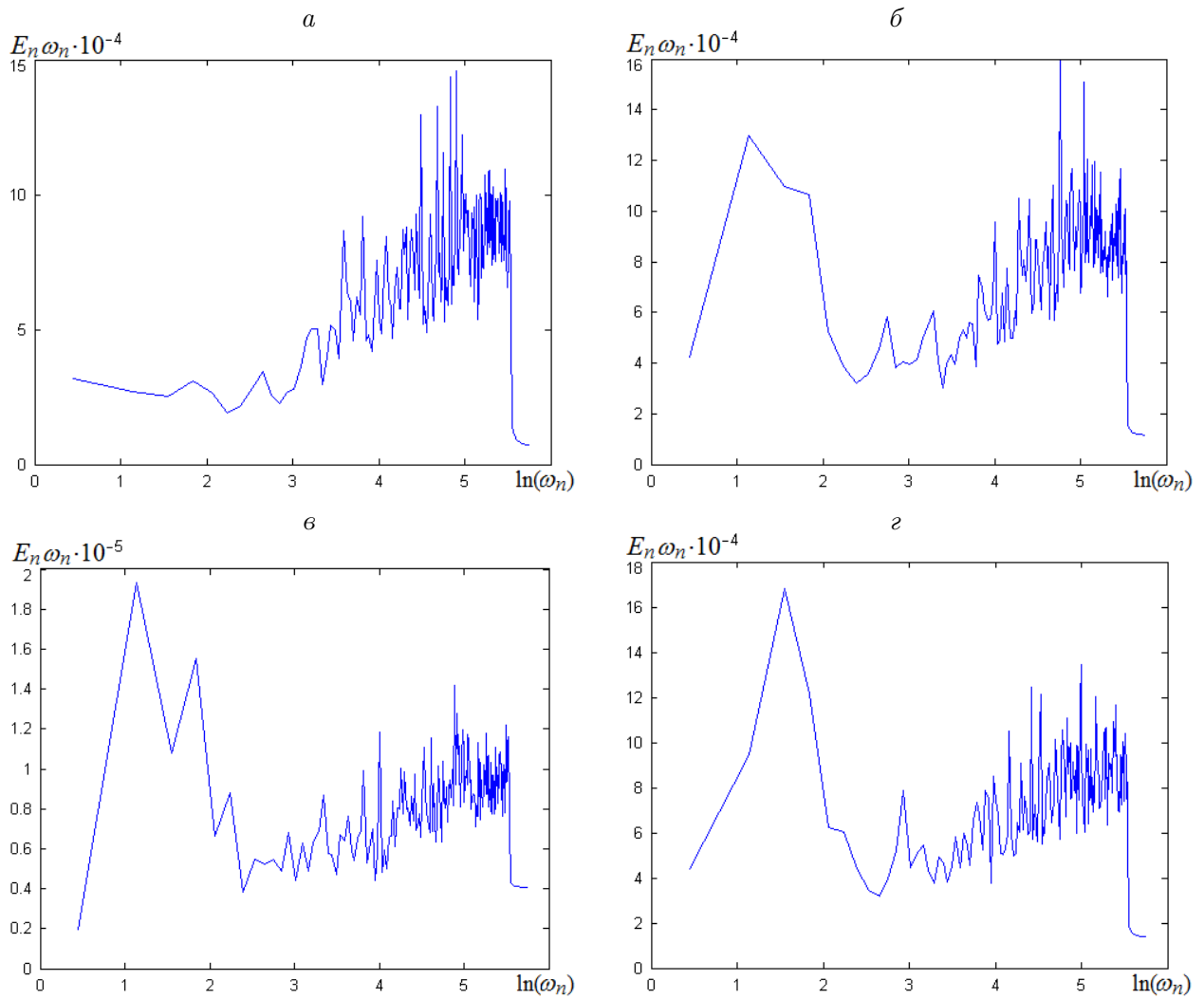


Рис. 5. Зависимость  $E_n \omega_n$  от  $\ln \omega_n$  для элементов выборки 81–84 (*a–г* соответственно)

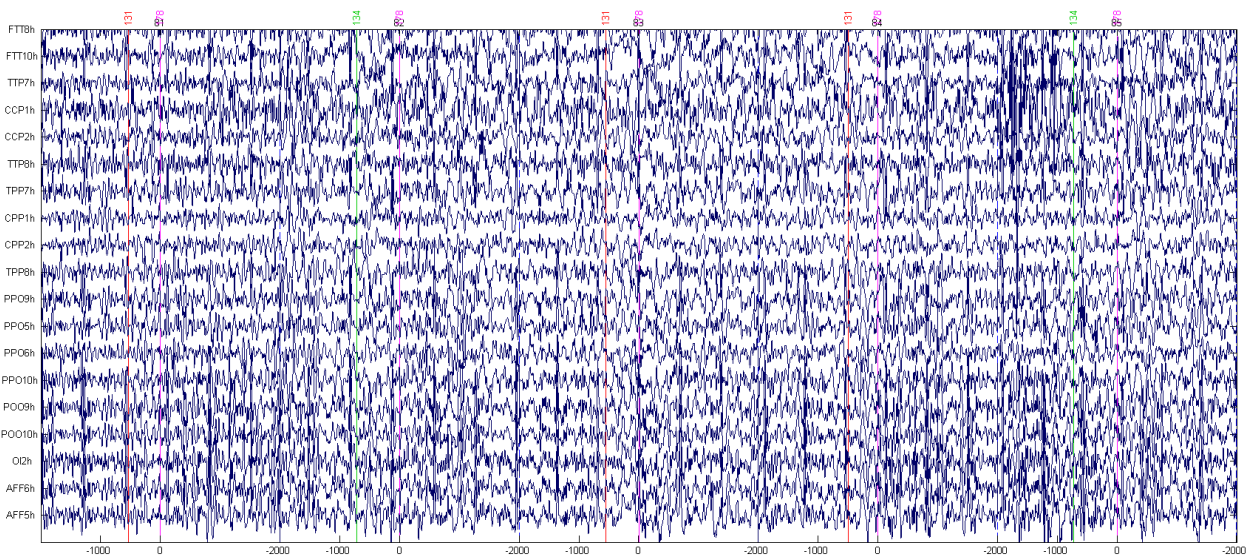


Рис. 6. Данные ЭЭГ после применения процедуры фильтрации для элементов выборки 81–84



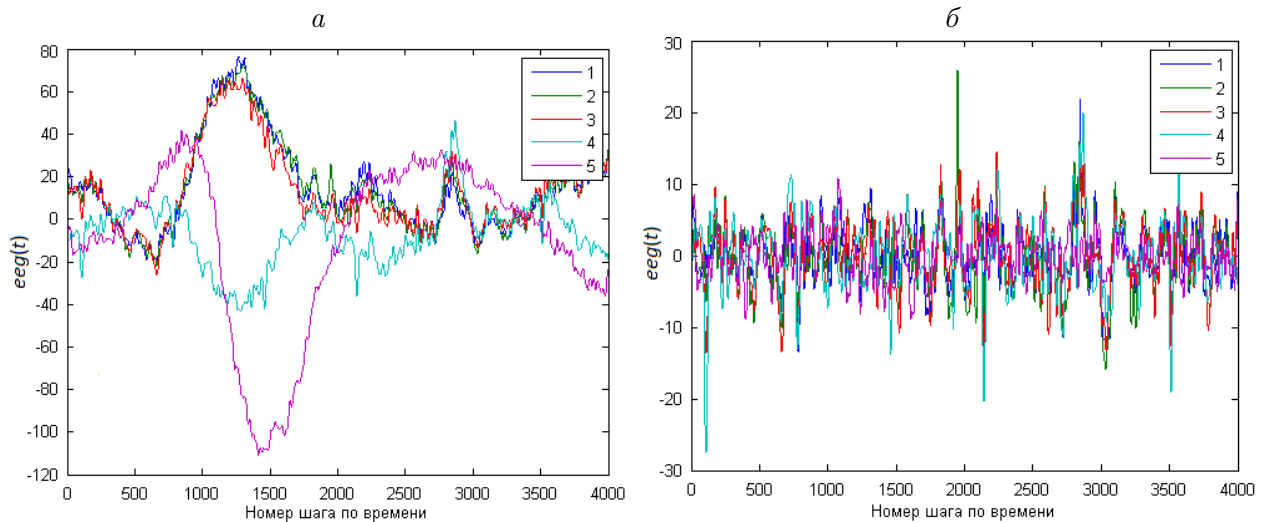


Рис. 7. Данные ЭЭГ  $eeg(t)$  для пяти датчиков: *a* — до фильтрации, *б* — после фильтрации

NETLIB (<http://www.netlib.org>). На рис. 6 представлены результаты фильтрации для элементов выборки с низкочастотным шумом, на рис. 7 — данные пяти датчиков до и после фильтрации. Из рисунков видно, что процедура фильтрации убирает низкочастотный шум с высокой амплитудой.

На практике важным свойством преобразования ЭЭГ является то, что оно не искажает фазу, поэтому при разработке методики фильтрации отдавалось предпочтение методам, обладающим этим свойством. Для проверки влияния данного алгоритма на фазу волны рассмотрена тестовая задача, когда процедура фильтрации применялась к сумме синусов с различными значениями частот. Коэффициент корреляции восстановленной после фильтрации высокочастотной и исходной волн (без шума) составил 0.9.

## Заключение

Рассмотрен алгоритм фильтрации низкочастотных шумов от различных источников в данных ЭЭГ. Алгоритм основан на фурье-анализе данных и может применяться локально (в заданном временном интервале). Алгоритм является аналогом метода скользящего среднего [11], однако имеет ряд принципиальных отличий: включает процедуру оценки интервала осреднения исходя из спектрального анализа данных, а также критерий, согласно которому принимается решение об удалении низкочастотной составляющей данных (шума) в заданной точке. Фактически процедура фильтрации применяется в подобластях исходной временной области. В дальнейшем планируется оптимизация выбора этих подобластей.

Следует отметить, что обычные фильтры убирают медленные частоты на всем протяжении записи ЭЭГ, поэтому есть риск вместе с артефактами удалить и полезный медленно-волновой сигнал. Предлагаемый алгоритм удаляет медленно-волновой шум локально в местах всплесков такой активности, что соответствует именно глазным движениям. В этом случае шум убирается, а полезный сигнал остается и фазовые отношения не нарушаются.

**Благодарности.** Коллекция данных ЭЭГ собрана при поддержке Российского научного фонда (грант № 14-15-00202).

## Список литературы / References

- [1] **Марютина Т.М., Ермолаев О.Ю.** Введение в психофизиологию. М.: Флинта, 2004. 400 с.  
**Marjutina, T.M., Ermolaev, O.Yu.** Introduction to psychophysiology. Moscow: Flinta, 2004. 400 p. (In Russ.)
- [2] **Рангайан Р.М.** Анализ биомедицинских сигналов. Практический подход. М.: Физматлит, 2010. 439 с.  
**Rangayyan, R.M.** Biomedical signal analysis. A case-study approach. Wiley, 2002. 552 p.
- [3] **Lee, Te-Won, Girolami, M., Bell, A.J., Sejnowski, T.J.** A unifying information-theoretic framework for independent component analysis // *Computers and Mathematics with Applications*. 2000. Vol. 39(11). P. 1–21.
- [4] **Morbidi, F., Garulli, A., Prattichizzo, D., Rizzo, C., Manganotti, P., Rossi, S.** Off-line removal of TMS-induced artifacts on human electroencephalography by Kalman filter // *J. of Neuroscience Methods*. 2007. Vol. 162, iss. 1-2. P. 293–302.
- [5] **Knyazev, G.G., Savostyanov, A.N., Bocharov, A.V., Tamozhnikov, S.S., Sapri-gyn, A.E.** Task-positive and task-negative networks and their relation to depression: EEG beamformer analysis // *Behavioural Brain Research*. 2016. Vol. 36, No. 1. P. 160–169. DOI: 10.1016/j.bbr.2016.03.033.
- [6] **Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S.** Functional connectivity in the motor cortex of resting human brain using echo-planar MRI // *Magnetic Resonance in Medicine*. 1995. Vol. 34(4). P. 537–541.
- [7] **Savostyanov, A.N., Tsai, A.C., Liou, M., Levin, E.A., Lee, J.D., Yurganov, A.V., Knyazev, G.G.** EEG correlates of trait anxiety in the stop-signal paradigm // *Neuroscience Letters*. 2009. Vol. 449(2). P. 112–116.
- [8] **Яглом А.М.** Корреляционная теория стационарных случайных функций. Л.: Гидрометеоиздат, 1981. 279 с.  
**Yaglom, A.M.** Correlation theory of stationary and related random functions. Supplementary Notes and References. New York: Springer, 1987. 258 p.
- [9] **Динамическая метеорология / Под ред. Д.Л. Лайхтмана.** Л.: Гидрометеоиздат, 1976. 606 с.  
*Dynamic meteorology / Ed. by D.L. Laykhtman.* Leningrad: Gidrometeoizdat, 1976. 606 p. (In Russ.)
- [10] **Калиткин Н.Н.** Численные методы. М.: Наука, 1978. 512 с.  
**Kalitkin, N.N.** Numerical methods. Moscow: Nauka, 1978. 512 p. (In Russ.)
- [11] **Справочник по прикладной статистике. В 2-х т. Т. 1: Пер. с англ. / Под ред. Э. Ллойда, У. Ледермана, Ю.Н. Тюрина.** М.: Финансы и статистика, 1989. 510 с.  
*Handbook on applied statistics. In two vol. V. 1 / Ed. by E. Lloyd, U. Lederman, Yu.N. Tyurin.* Moscow: Finansy i Statistika, 1989. 510 p. (In Russ.)

*Поступила в редакцию 22 июня 2016 г.,  
с доработки — 29 сентября 2016 г.*

## Algorithm for local filtering of low frequency and large amplitude noise in electroencephalograms data

KLIMOVA, EKATERINA G.<sup>1,\*</sup>, MEDVEDEV, SERGEY B.<sup>1</sup>, SAVOSTYANOV, ALEXANDER N.<sup>2,3</sup>

<sup>1</sup>Institute of Computational Technologies SB RAS, Novosibirsk, 630090, Russia

<sup>2</sup>Scientific Research Institute of Physiology and Fundamental Medicine, Novosibirsk, 630117, Russia

<sup>3</sup>Novosibirsk State University, Novosibirsk, 630090, Russia

\*Corresponding author: Klimova, Ekaterina G., e-mail: klimova@ict.nsc.ru

**Purpose.** One of the problems in operation of electroencephalograms (EEG) data is the existence of “noise” in these data. The noise is connected both to a behavior of the patient, and to external influences. The purpose of this study is creation of an algorithm for filtering of low frequency and large amplitude noise arising from different sources in EEG data. At the same time one of requirements to the algorithm is elimination of the phase distortion in a wave.

**Methods.** The algorithm is based on the Fourier analysis of data and can be applied locally in the set time interval. In the developed algorithm the approach to division of high-frequency and low-frequency fluctuations which is widely applied in the theory of turbulence is used. The algorithm is analog of the method of the sliding average. However it has a number of fundamental differences. The algorithm includes an averaging interval assessment procedure proceeding from the spectral analysis of data, and also criterion according to which the decision on removal of a low-frequency component of data (noise) in the set point is made.

**Results.** Numerical experiments with real data of electroencephalograms are made. EEG registered on the basis of scientific research institute of physiology and fundamental medicine with the help of 128 channel encephalograph of Brain Products, Germany with bandpass range of 0.1–100 Hz, the frequency of digitization of a signal of 1000 Hz. Electrodes (126 EEG + VEOG) settled down according to the international diagram of 10-10 % with the reviewer Cz and a grounding electrode of AFz. Numerical experiments with real data of electroencephalograms (selection from 85 EEG) showed that the algorithm allows to adjust EEG data, cleaning low frequency noise and slightly changing wave phase.

**Conclusions.** In this work the algorithm of a filtration of low-frequency noise in EEG data is proposed. Distinctive property of the offered algorithm is lack of distortion of a phase of a wave (within computational error of the method).

*Keywords:* electroencephalogram, filtration of noise, spectral analysis.

**Acknowledgements.** EEG data has been collected with the assistance of the Russian Scientific Fund, grant No. 14-15-00202.

*Received 22 June 2016,*

*Received in revised form 29 September 2016*