

Исследование проблемы множественного сравнения при анализе всплескообразной электрической активности мышц у пациентов с болезнью Паркинсона

О.С. Сушкова¹, А.А. Морозов¹, А.В. Габова², А.В. Карабанов³

¹Институт радиотехники и электроники им. В.А. Котельникова РАН, Моховая 11, корп. 7, Москва, Россия, 125009

²Институт высшей нервной деятельности и нейрофизиологии РАН, Бултерова 5А, Москва, Россия, 117485

³ФГБНУ «Научный центр неврологии», Волоколамское шоссе 80, Москва, Россия, 125367

Аннотация. Разработан метод анализа всплескообразной электрической активности мышц, основанный на вейвлет-анализе и ROC-анализе, который позволяет изучать частотно-временные особенности сигналов электромиограмм (ЭМГ) и акселерометра (АКС) «треморных» конечностей пациентов с болезнью Паркинсона (БП). Идея метода заключается в поиске локальных максимумов («всплесков») на вейвлет-спектрограмме и вычислении различных характеристик, описывающих эти максимумы: ведущая частота всплеска, длительность всплеска в периодах, ширина полосы частот всплеска, число всплесков в секунду. Анализируется степень отличия группы пациентов от контрольной группы испытуемых в пространстве этих параметров. Для этого используется ROC-анализ. Исследуется функциональная зависимость AUC (площади под ROC-кривой) от значений границ диапазонов рассматриваемых параметров. Применение этого метода позволило выявить новые нейрофизиологические закономерности в диапазоне 3–7 Гц в сигналах ЭМГ и АКС. В этих частотных диапазонах наблюдаются отличия между группой пациентов с БП и контрольной группой испытуемых. Разработанный метод предусматривает перебор большого количества диапазонов выбранных характеристик, поэтому при статистическом оценивании различных гипотез возникает проблема множественного сравнения. Необходимо найти компромисс между степенью детализации изучаемых характеристик и величиной поправки Бонферрони. В работе рассказывается о процедуре проверки статистических гипотез на данных пациентов с БП на ранней стадии.

1. Введение

Исследование проблемы множественного сравнения [1–5] в настоящее время является актуальной задачей биомедицины. Для решения проблемы множественного сравнения разработано большое количество методов, основанных на анализе групповой вероятности ошибки первого рода (familywise error rate, FWE), средней доли ложных отклонений гипотез (false discovery rate, FDR), применении теории случайных полей (RFT) [5–7], метода перестановок (permutation method) [3] и др. К сожалению, подавляющее большинство указанных методов не рассматривает связь проблемы множественного сравнения с проблемой многомасштабного анализа биомедицинских данных. Между тем эти две проблемы очень тесно связаны между собой, как с точки зрения математического аппарата, так и с точки зрения практического применения методов статистического анализа данных. Проблема многомасштабного анализа заключается в том, что биомедицинские данные могут содержать закономерности, проявляющиеся на различных, неизвестных заранее масштабах изображений или сигналов. Для выявления таких закономерностей необходимо исследование данных в некотором пространстве шкал (масштабов). Вместе с тем, разные масштабы изображений и сигналов соответствуют разному количеству множественных сравнений и, следовательно, предполагают разный уровень статистической поправки на количество множественных сравнений. Более того, исследование данных на множестве шкал само по себе является случаем множественного сравнения и, следовательно, может потребовать применения дополнительной поправки на множественное сравнение. Эта проблема, в частности, была рассмотрена в рамках теории случайных полей, и был предложен подход к анализу данных на основе пространства шкал (scale space approach) [8], однако в данной статье в качестве поправки на количество множественных сравнений мы будем использовать стандартную поправку Бонферрони [1], чтобы сконцентрироваться на проблеме многомасштабного анализа.

Ранее мы разработали метод анализа всплескообразной электрической активности коры головного мозга, основанный на анализе вейвлетов и ROC-анализе [9,10]. Идея этого метода анализа заключается в том, что мы рассматриваем сигнал электроэнцефалограммы (ЭЭГ) как набор всплесков [11]. В отличие от работ, посвящённых обнаружению электрической активности одного или двух конкретных типов, таких как альфа-веретёна [12] и сонные веретёна [13–18], мы анализируем любые виды всплескообразной электрической активности в коре мозга в широком диапазоне частот. Кроме того, мы рассматриваем всплеск как типичный компонент ЭЭГ, а не как особый вид сигналов ЭЭГ [19]. Рассмотренный в данной статье метод анализа основан на статистическом анализе вейвлет-спектрограмм, новом методе визуализации результатов статистического анализа и новом алгоритме обнаружения всплесков.

Для анализа сигналов электромиограмм (ЭМГ) и акселерометра (АКС) разработанный алгоритм анализа сигналов [9,10] был модифицирован. Был добавлен дополнительный этап сглаживания вейвлет-спектрограмм сигналов. Сглаживание требуется потому, что стандартные быстрые алгоритмы вычисления вейвлетов имеют следующую проблему: вейвлет-спектрограммы (при обработке сигналов сложной формы) неизбежно загрязняются артефактами (выбросами и высокочастотными колебаниями). Эти артефакты можно ошибочно распознать как «всплески». Для решения этой проблемы было применено сглаживание вейвлет-спектрограмм гауссовым окном со специально выбранными параметрами. Параметры окна сглаживания (ширина по времени и ширина полосы частот) изменяются адаптивно в зависимости от параметра масштабирования вейвлета. Кроме того, был уменьшен порог, используемый для первоначального выбора всплесков; в настоящее время длительность

всплесков должна быть больше или равна 1 периоду. Длительность всплеска измеряется на высоте, равной квадратному корню из 1/2 от максимума пика.

Применение метода анализа всплескообразной электрической активности позволило выявить новые нейрофизиологические закономерности у пациентов с болезнью Паркинсона (БП) по сравнению со здоровыми испытуемыми в диапазоне 3–7 Гц в сигналах ЭМГ и АКС. Было обнаружено, что всплески в диапазоне 3–7 Гц имеют следующие характеристики: $50 \text{ мкВ}^2/\text{Гц}$ и выше, 1–5 периодов, ширина полосы частот 1–5 Гц [20]. Отличия между группами было обнаружено как на здоровых руках, так и на «треморных» руках пациентов (с тремором на левой руке и с тремором на правой руке). Причём отличия на здоровых левых руках пациентов с тремором на правой руке более выражены на ЭМГ, а отличия на здоровых правых руках у пациентов с тремором на левой руке более выражены на сигналах акселерометра.

2. Подготовка эксперимента

Данные нелеченых (то есть, ранее не принимавших специальные лекарства) пациентов с БП на ранних стадиях сравнивались с данными здоровых добровольцев. Отметим, что группа пациентов с БП включала как пациентов с тремором левой руки (15 человек), так и пациентов с тремором правой руки (18 человек), всего 33 пациента. Число здоровых добровольцев составляло 18 человек. Все пациенты и здоровые добровольцы были правшами. Статистически значимых различий между возрастными пациентами и здоровых добровольцев обнаружено не было. ЭМГ-электроды располагались на внешних сторонах рук, на мышцах-разгибателях, акселерометры были установлены на тыльных сторонах ладоней. Сигналы ЭМГ и акселерометра записывались в специальном положении испытуемого – испытуемый сидел в кресле, руки лежали на подлокотниках кресла, ладони были выпрямлены и повернуты ребром к полу, ноги были поставлены на пяточки. Глаза во время всех записей были закрыты. Для записи ЭМГ использовался 41-канальный многофункциональный комплекс для проведения нейрофизиологических исследований Нейрон-Спектр-5 (фирма Нейрософт). Для записи АКС использовался акселерометр, разработанный Е. М. Тиманиным в ИПФ РАН [21]. Частота дискретизации ЭМГ составляла 500 Гц, частота дискретизации акселерометра 1378.125 Гц. Для ЭМГ и акселерометра использовались фильтр высоких частот с частотой отсечки 0.5 Гц и режекторный фильтр с частотой 50 Гц, кроме того, для ЭМГ использовался фильтр Баттерворта от 60 до 240 Гц. К сигналам ЭМГ после фильтрации применялось преобразование Гильберта для выделения огибающей сигнала. Продолжительность каждой записи составляла около двух минут. Записи анализировались как есть, без выбора отдельных областей в сигнале.

3. Методы

Для обнаружения статистически значимых отличий в изучаемых характеристиках всплесков электрической активности мышц у групп пациентов и контрольных испытуемых используется непараметрический тест Манна-Уитни. На Рис. 1 и Рис. 2 представлены диаграммы значений AUC (далее, AUC-диаграммы) с обнаруженными статистически значимыми областями в диапазоне частот от 3 до 7 Гц (без поправки на множественное сравнение). Видно, что статистически значимые области есть практически на всех изображениях, за исключением тех изображений, на которых сравниваются количества всплесков в ЭМГ-сигналах на правых руках контрольных испытуемых и на правых (здоровых) руках пациентов с тремором на левой руке (Рис. 1).

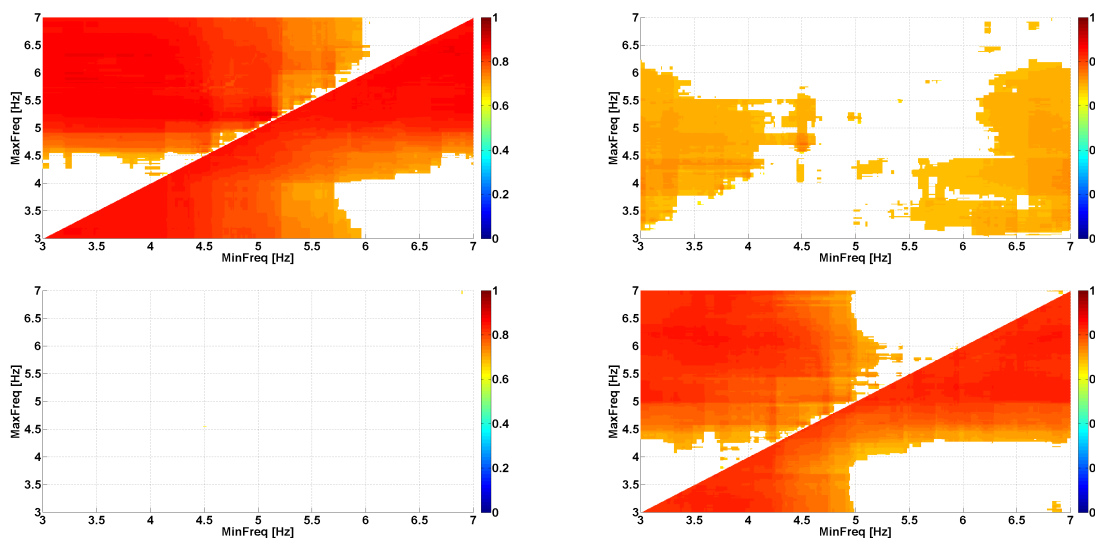


Рисунок 1. Диаграммы значений AUC на основе сигналов ЭМГ с обнаруженными статистически значимыми областями без поправки на множественное сравнение. Первый столбец – пациенты с тремором на левой руке, второй столбец – пациенты с тремором на правой руке. Первая строка – левые руки. Вторая строка – правые руки. Ось абсцисс – нижняя граница диапазона частот, ось ординат – верхняя граница диапазона частот.

Разработанный метод предусматривает перебор большого количества диапазонов выбранных характеристик, поэтому при проверке статистических гипотез возникает проблема множественного сравнения. Самое простое решение – использование поправки Бонферрони. Альфа-уровень остаётся постоянным, 0.05. Варьируется количество K ячеек на диаграмме. Будем называть разрешением R метода число колонок (или строк) в диаграмме. Возникает вопрос – какое разрешение необходимо выбрать? Большое разрешение позволяет обнаруживать на диаграмме области небольшого размера, однако поправка Бонферрони при этом может оказаться слишком большой и, следовательно, статистический тест не обнаружит никаких статистически значимых отличий. Необходимо добиться того, чтобы были видны интересные детали, но при этом поправка на множественное сравнение не была слишком большой, то есть, необходимо найти компромисс между степенью детализации изучаемых характеристик и величиной поправки Бонферрони.

Идея нашего подхода заключается в многомасштабном анализе, мы изменяем разрешение на диаграммах для поиска такой поправки Бонферрони, при которой не пропадают статистически значимые эффекты. Интересно найти наибольшее разрешение, где появляется статистически значимая область, прошедшая поправку Бонферрони.

Рассмотрим AUC-диаграммы в следующем подпространстве значений параметров всплеск: диапазон частот от 3 до 7 Гц, амплитуда от 50 мкВ2/Гц и выше, число периодов от 1 до 5 периодов, ширина полосы частот всплеск от 1 до 5 Гц. В данном примере исследуются сигналы акселерометра, пациенты с тремором левой «треморной» руки.

Первый раз статистически значимая область появляется на разрешение 108 (то есть, число рассматриваемых ячеек 108×108). На Рис. 3 приведено разрешение 69 с хорошо выраженной статистически значимой областью (примерно в диапазоне от 4 до 6 Гц). Обратите внимание, что поправка на множественное сравнение сделана только на количество ячеек в AUC-диаграмме.

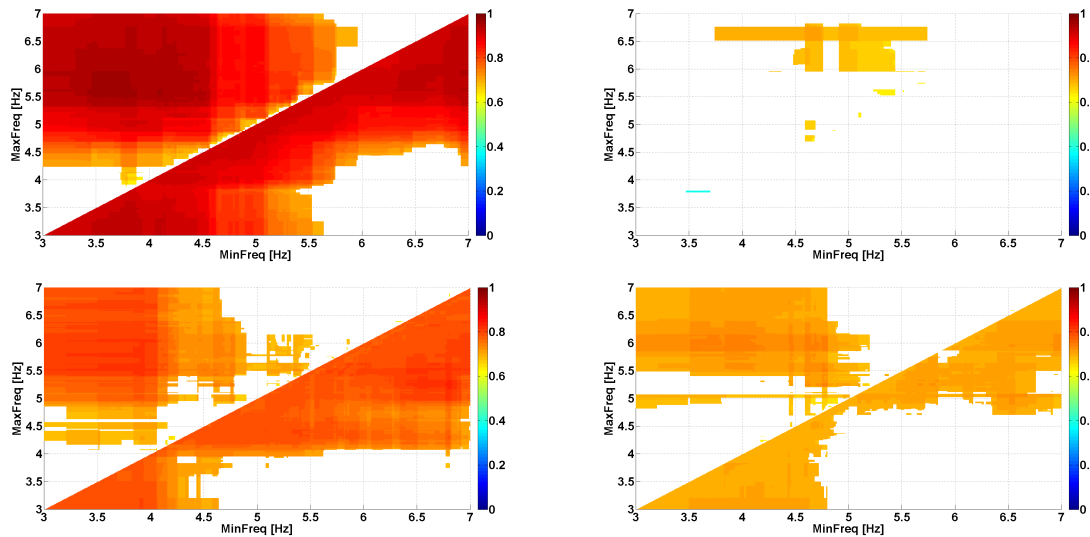


Рисунок 2. Диаграммы значений AUC на основе сигналов АКС с обнаруженными статистически значимыми областями без поправки на множественное сравнение. Первый столбец – пациенты с тремором на левой руке, второй столбец – пациенты с тремором на правой руке. Первая строка – левые руки. Вторая строка – правые руки. Ось абсцисс – нижняя граница диапазона частот, ось ординат – верхняя граница диапазона частот.

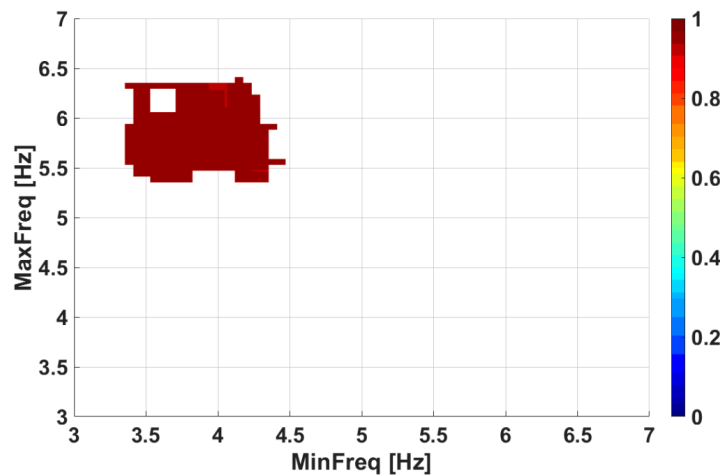


Рисунок 3. Разрешение 69 с хорошо выраженной статистически значимой областью у пациентов с тремором на левой «треморной» руке.

Дополнительная поправка на количество рассматриваемых шкал не производится, поэтому форма полученных областей статистически значимых отличий является случайной. В данной работе форма областей статистически значимых отличий не исследуется, нам важен лишь факт наличия статистически значимых отличий в исследуемых частотных областях.

Полученные зависимости статистически значимых областей от величины разрешения можно представить в виде изоповерхностей. На Рис. 4 приведена изоповерхность $p\text{-value} \leq 0.05$. Пациенты с тремором левой «треморной» руки, акселерометр, частоты от 3 до 7 Гц.

На рассмотренных рисунках видно, что статистически значимые области имеют формы «груш», то есть, от верха до низа они постепенно увеличиваются, а затем снова уменьшаются (примерно на разрешении 3–4). Уменьшение статистически значимых областей происходит

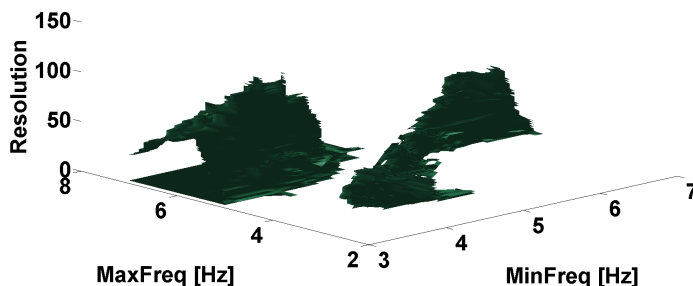


Рисунок 4. Изоповерхность $p\text{-value} \leq 0.05$. Пациенты с тремором левой «треморной» руки, акселерометр, частоты от 3 до 7 Гц.

потому, что на малом количестве рассматриваемых ячеек, в одни и те же ячейки попадают области с закономерностями противоположной направленности, а также области, в которых какие-либо закономерности не проявляются.

Рассмотрим виды сверху изоповерхностей, полученных на основе сигналов ЭМГ (Рис. 5) и АКС (Рис. 6). Статистически значимые области, прошедшие поправку Бонферрони, остались на ЭМГ только на «треморных» руках. Они наблюдаются как у пациентов с тремором на левой руке, так и у пациентов с тремором на правой руке. На акселерометре статистически значимые области остались только у пациентов с тремором на левой руке и только на «треморной» (левой) руке.

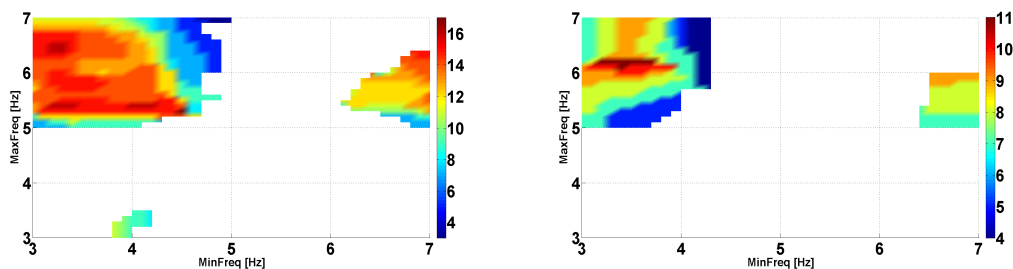


Рисунок 5. Виды сверху на изоповерхности, полученные на основе ЭМГ. Первый столбец – пациенты с тремором на левой руке, «треморная» (левая) рука, второй столбец – пациенты с тремором на правой руке, «треморная» (правая) рука. Ось абсцисс – нижняя граница диапазона частот, ось ординат – верхняя граница диапазона частот.

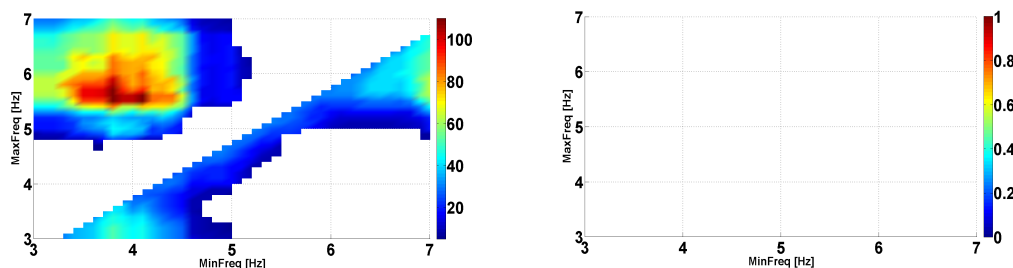


Рисунок 6. Виды сверху на изоповерхности, полученные на основе АКС. Первый столбец – пациенты с тремором на левой руке, «треморная» (левая) рука, второй столбец – пациенты с тремором на правой руке, «треморная» (правая) рука. Ось абсцисс – нижняя граница диапазона частот, ось ординат – верхняя граница диапазона частот.

4. Результаты

Разработан новый метод разведочного анализа данных. Этот метод включает вычисление диаграмм значений AUC и непараметрическую проверку статистических гипотез для обнаружения статистически значимых отличий в характеристиках «всплесков» электрической активности мышц. Проведён детальный анализ данных пациентов с болезнью Паркинсона и контрольных испытуемых. Доказано, что выявленные отличия между группами пациентов с БП и здоровых испытуемых являются статистически значимыми. Найдены области с наименьшими p-value, прошедшие поправку Бонферрони. Найдены статистически значимые отличия в данных ЭМГ – на «треморных» руках. Статистически значимые отличия в данных ЭМГ выявлены как у пациентов с тремором на левой руке, так и у пациентов с тремором на правой руке; на основе данных акселерометра – только у пациентов с тремором на левой руке и только на «треморной» руке, однако закономерности выражены более сильно, чем в ЭМГ-данных. Было показано, что анализ пациентов с болезнью Паркинсона с помощью ЭМГ и акселерометра даёт разные результаты. ЭМГ и акселерометр могут взаимно дополнять друг друга. Обнаруженные закономерности в данных ЭМГ и акселерометра могут быть полезны для диагностики ранних стадий болезни Паркинсона.

5. Литература

- [1] Shaffer, J.P. Multiple hypothesis testing: A review // *Annual Review of Psychology*. — 1995. — Vol. 46(1). — P. 561-584.
- [2] Pigeot, I. Basic concepts of multiple tests – a survey // *Statistical papers*. — 2000. — Vol. 41(1). — P. 3-36.
- [3] Nichols, T.E. Multiple testing corrections, nonparametric methods, and random field theory // *Neuroimage*. — 2012. — Vol. 62(2). — P. 811-815.
- [4] Austin, S.R. Multiple hypothesis testing: A review / S.R. Austin, I. Dialsingh, N. Altman // *J. Indian Soc. Of Agricultural Stat.* — 2014. — Vol. 68. — P. 303-314.
- [5] Petersson, K.M. Statistical limitations in functional neuroimaging II. signal detection and statistical inference / K.M. Petersson, T.E. Nichols, J.-B. Poline, A.P. Holmes // *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. — 1999. — Vol. 354(1387). — P. 1261-1281.
- [6] Brett, M. Introduction to Random Field Theory / M. Brett, W. Penny, S. Kiebel // *Human Brain Function / Academic Press*, 2003.
- [7] Rohani, F. Nonparametric Random Fields with Applications in Functional Imaging: Ph.D. thesis / McGill University, 2009.
- [8] Worsley, K.J. Scale space searches for a periodic signal in fMRI data with spatially varying hemodynamic response / K.J. Worsley, M. Wolforth, A.C. Evans // *Proceedings of BrainMap*. — 1997. — Vol. 96.
- [9] Sushkova, O.S. A method of analysis of EEG wave trains in early stages of Parkinson's disease / O.S. Sushkova, A.A. Morozov, A.V. Gabova // *International Conference on Bioinformatics and Systems Biology*, 2016. — P. 1-4.
- [10] Sushkova, O.S. Data mining in EEG wave trains in early stages of Parkinson's disease / O.S. Sushkova, A.A. Morozov, A.V. Gabova // *Advances in Soft Computing. Lecture Notes in Computer Science*. — 2017. — Vol. 10062. — P. 403-412.

- [11] Sushkova, O.S. Data mining in EEG wave trains in early stages of Parkinson's disease / O.S. Sushkova, A.A. Morozov, V. Gabova, A.V. Karabanov // Proceedings of the 12th Russian-German Conference on Biomedical Engineering, 2016. – P. 80-84.
- [12] Lawhern, V. Detecting alpha spindle events in EEG time series using adaptive autoregressive models / V. Lawhern, S. Kerick, K.A. Robbins // BMC Neuroscience. – 2013. – Vol. 14. – P. 101.
- [13] Parekh, A. Sleep spindle detection using time-frequency sparsity / A. Parekh, I. Selesnick, D. Rapoport, I. Ayappa // IEEE Signal Processing in Medicine and Biology Symposium. – Philadelphia, PA, 2014. – P. 1-6.
- [14] O'Reilly, C. Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools / C. O'Reilly, T. Nielsen // Frontiers in Human Neuroscience. – 2015. – Vol. 9. – P. 353. DOI: 10.3389/fnhum.2015.00353.
- [15] Huupponen, E. Determination of dominant simulated spindle frequency with different methods / E. Huupponen, W.D. Clercq, G. Gómez-Herrero // Journal of Neuroscience Methods. – 2006. – Vol. 156. – P. 275-283.
- [16] Nonclercq, A. Sleep spindle detection through amplitude-frequency normal modelling / A. Nonclercq, C. Urbain, D. Verheulpen // Journal of Neuroscience Methods. – 2013. – Vol. 214. – P. 192-203.
- [17] Jaleel, A. Improved spindle detection through intuitive pre-processing of electroencephalogram / A. Jaleel, B. Ahmed, R. Tafreshi // Journal of Neuroscience Methods. – 2014. – Vol. 233. – P. 1-12.
- [18] Camilleri, T.A. Automatic detection of spindles and K-complexes in sleep EEG using switching multiple models / T.A. Camilleri, K.P. Camilleri, S.G. Fabri // Biomedical Signal Processing and Control. – 2014. – Vol. 10. – P. 117-127.
- [19] Жирмунская, Е.А. Клиническая электроэнцефалография (цифры, гистограммы, иллюстрации). — Москва: Межотраслевой научно-исследовательский инженерно-технологический центр «Скан», 1993.
- [20] Sushkova, O.S. Investigation of surface EMG and acceleration signals of limbs' tremor in Parkinson's disease patients using the method of electrical activity analysis based on wave trains / O.S. Sushkova, A.A. Morozov, V. Gabova, A.V. Karabanov // Advances in Artificial Intelligence: 16th Ibero-American Conference on AI. – Cham: Springer International Publishing, 2018. – P. 253-264.
- [21] Тиманин, Е.М. Устройство для комплексного исследования различных видов тремора человека / Е.М. Тиманин, А.В. Густов, Е.В. Ерёмин // Патент № 2483676 Российской Федерации. 10.06.2013.

Благодарности

Авторы благодарны академику С.Н. Иллариошкину за помощь в проведении работы. Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-37-20021 с привлечением средств государственного задания по теме «Фундаментальные основы радиоэлектронных методов для проблем биомедицины», стипендии Президента РФ молодым учёным и аспирантам № СП-5247.2018.4 и поддержано Российской академией наук.

Investigation of the multiple comparisons problem in the wave train electrical activity analysis of the muscles in Parkinson's disease patients

O.S. Sushkova¹, A.A. Morozov¹, A.V. Gabova², A.V. Karabanov³

¹Kotel'nikov Institute of Radio Engineering and Electronics of RAS, Mokhovaya 11-7, Moscow, Russia, 125009

²Institute of Higher Nervous Activity and Neurophysiology of RAS, Butlerova 5A, Moscow, Russia, 117485

³FSBI "Research Center of Neurology", Volokolamskoe shosse 80, Moscow, Russia, 125367

Abstract. We have developed a new method for analyzing wave train electrical activity of the muscles based on the wavelet analysis and ROC analysis that enables to study the time-frequency features of electromyograms (EMG) and acceleration (ACC) signals of limbs' tremor in patients with Parkinson's disease (PD). The idea of the method is to find local maxima (named wave trains) in the wavelet spectrogram and to calculate various characteristics describing these maxima: the leading frequency, the duration in periods, the bandwidth, the number of wave trains per second. The degree of difference of the group of patients from the control group of subjects in the space of these parameters is analyzed. ROC analysis is used for this purpose. The functional dependence of AUC (the area under the ROC curve) on the values of the bounds of the ranges of the parameters under consideration is investigated. This method allows revealing new neurophysiological regularities in the range of 3–7 Hz in the EMG and ACC signals. The differences between the group of patients with PD and the control group of subjects are observed in these frequency ranges. The method involves investigation of a large number of ranges of selected characteristics; therefore a multiple comparisons problem appears during the statistical check of the statistical hypotheses. It is necessary to find a compromise between the degree of detail of the studied characteristics and the magnitude of the Bonferroni correction. The paper describes the statistical check of statistical hypotheses on the data of patients with early Parkinson's disease.