

ИМПБ РАН 50 лет. Лаборатория обработки данных

Махортых С.А.

ИМПБ РАН – филиал ИПМ им. М.В. Келдыша РАН

makh@impb.ru

Разработаны основные принципы обобщенного спектрально-аналитического метода (ОСАМ). Этот подход реализован при решении задач анализа биомедицинских и биоинформационных данных [1, 2]. Разработаны методы диагностики, распознавания, классификации и прогнозирования с использованием настраиваемых (адаптивных) спектральных признаков. Разработаны основные принципы использования сферических функций при анализе данных на сфере (например, в задачах картографии, анализе магнитоэнцефалограмм). Разработана математическая библиотека преобразований данных в спектральном представлении по системам функций непрерывного и дискретного аргументов. На основе этого подхода предложены способы решения и реализовано соответствующее программно-алгоритмическое обеспечение в ряде актуальных задач биомедицины, биоинформатики, экологии. Ниже кратко представлены основные результаты.

Разработаны параллельные алгоритмы ОСАМ для решения задач анализа электрофизиологических данных деятельности ЦНС. На основе предложенных моделей выполнен анализ источников и их связей для ЭЭГ и МЭГ с высоким разрешением (более 200 отведений). Для различных типов связей сгенерированы данные МЭГ/ЭЭГ и найдено оптимальное сочетание методов для анализа источников, а также сравнение ЭЭГ и МЭГ методов регистрации [3, 4]. Пример с тремя связанными источниками и результаты оценки связности с моделированными данными для разных способов оценки приводится на рис. 1. Найдены зоны мозга с повышенной биомагнитной активностью для клинических случаев, связанных с болезнью Паркинсона и синдромом Tinnitus.

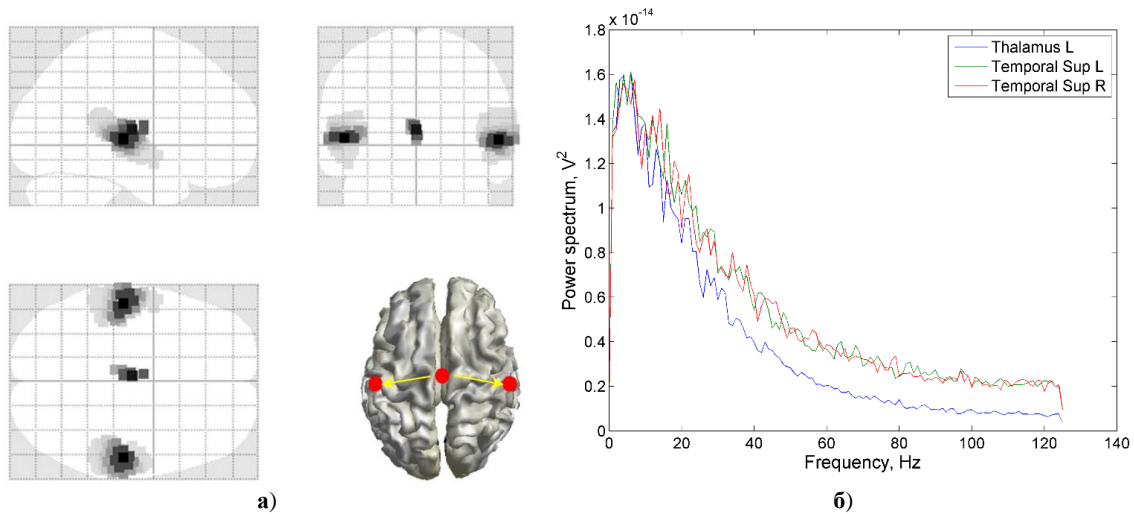


Рис. 1. Сгенерированная активность с тремя источниками: таламус и височные доли а), спектр для различных источников б).

Разработан спектрально-аналитический подход к выявлению размытых протяженных повторов в геномных последовательностях [5, 6]. Метод основан на разномасштабном интегральном оценивании сходства нуклеотидных последовательностей в пространстве коэффициентов разложения фрагментов кривых GC- и GA-содержания по классическим ортогональным базисам. Метод одинаково хорошо работает на разных масштабах данных. Он позволяет выявлять следы сегментных дубликаций и мегасателлитные участки в геноме, районы синтении при сравнении пары геномов. Найдены условия оптимальной аппроксимации, обеспечивающие автоматическое распознавание повторов различных видов (прямых и инвертированных, а также тандемных) на спектральной матрице сходства (рис. 2). Реализованы эффективные параллельные алгоритмы анализа геномов [7].

Разработан информационно-вычислительный сервис [8, 9] для исследования структурно-функциональных элементов белковых молекул, который включает в себя базу данных спиральных пар с различными геометрическими характеристиками и соответствующими аминокислотными

последовательностями и определенный базовый набор инструментов визуализации и анализа характеристик структурных мотивов для установления закономерностей их структурной организации. Реализован сервис, позволяющий проводить численные эксперименты молекулярной динамики спиральных пар. Сервис доступен по адресу <http://protodb.org>.

Создана база данных, содержащая сайты щепления человека, а также ортологичные сайты, найденные в других организмах [10]. Разработан ряд критериев консервативности. Разработанная модель (рис. 3) выявления апоптотических сайтов щепления успешно опробована на воспалительных каспазах [11].

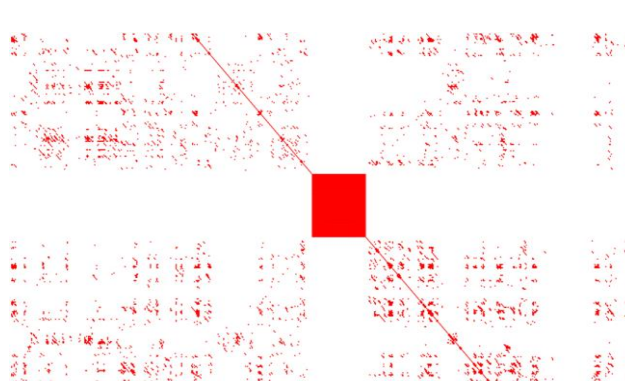


Рис. 2. Фрагмент точечной матрицы, содержащий тандемный повтор.

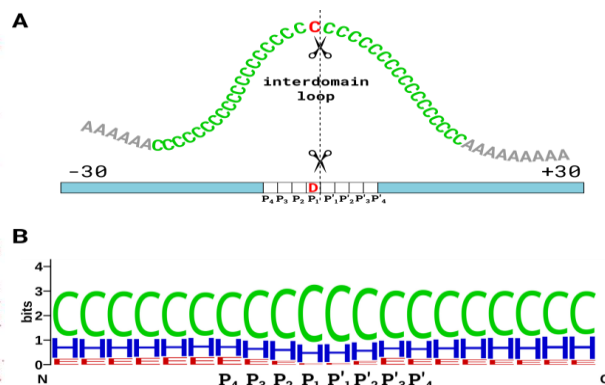


Рис. 3. Модель сайта щепления.

Разработаны эффективные специально-аналитические методы классификации и диагностики данных в области кардиологии [12] и экологии [13].

Список литературы

1. Makhortykh S.A. Generalized spectral-analytical method for biomedical data processing. *Mathematica Montisnigri*. 2016. V. 36. P. 104–113.
2. Махортых С.А., Куликова Л.И., Панкратов А.Н., Тетуев Р.К. Generalized spectral-analytical method and its applications in problems of image analysis and pattern recognition. *Pattern Recognition and Image Analysis*. 2019. V. 29. No. 3. P. 621–638.
3. Makhortykh S., Lyzhko E. Sources localization for brain biomagnetic activity. *Review of Applied Physics (RAP)*. 2014. V. 3. P. 25–28.
4. Kumar A., Lyzhko E., Hamid L., Srivastav A., Stephani U., Japaridze N. Differentiating ictal/subclinical spikes and waves in childhood absence epilepsy by spectral and network analyses: A pilot study. *Clinical Neurophysiology*. 2021. T. 132. № 9. С. 2222–2231. doi: [10.1016/j.clinph.2021.06.011](https://doi.org/10.1016/j.clinph.2021.06.011)
5. Pyatkov M.I., Pankratov A.N. SBARS: fast creation of dotplots for DNA sequences on different scales using GA-,GC-content. *Bioinformatics*. 2014. V. 30. № 12. P. 1765–1766.
6. Pankratov A.N., Tetuev R.K., Pyatkov M.I., Toigildin V.P., Popova N.N. Spectral analytical method of recognition of inexact repeats in character sequences. *Proceedings of the Institute for System Programming*. 2015. V. 27. No. 6. P. 335–344.
7. Тетуев Р.К., Пятков М.И., Панкратов А.Н. Параллельный алгоритм глобального выравнивания протяжённых аминокислотных и нуклеотидных последовательностей. *Мат. биол. и биоинф.* 2017. Т. 12. № 1. С. 137–150.
8. Руднев В.Р., Панкратов А.Н., Куликова Л.И., Дедус Ф.Ф., Тихонов Д.А., Ефимов А.В. Распознавание и анализ устойчивости структурных мотивов типа α - α -уголок в глобулярных белках. *Мат. биол. и биоинф.* 2013. Т. 8. № 2. С. 398–406. doi: [10.17537/2013.8.398](https://doi.org/10.17537/2013.8.398)
9. Rudnev V.R., Kulikova L.I., Kaysheva A.L., Efimov A.V., Tikhonov D.A. Use of the Molecular Dynamics Method to Investigate the Stability of α - α -Corner Structural Motifs in Proteins. *Symmetry*. 2021. V. 13. P. 1193. doi: [10.3390/sym13071193](https://doi.org/10.3390/sym13071193)
10. Gubina N.E., Leboeuf D., Piatkov K.I., Pyatkov M.I. Novel apoptotic mediators identified by conservation of vertebrate caspase targets. *Biomolecules*. 2020. V. 10. No. 4. P. 612.
11. Leboeuf D., Pyatkov M., Zatsepin T.S., Piatkov K. The Arg/N-Degron Pathway—A Potential Running Back in Fine-Tuning the Inflammatory Response? *Biomolecules*. 2020. V. 10. No. 6. P. 903.
12. Москаленко А.В., Тетуев Р.К., Махортых С.А. Математические основы и перспективы кардиофизики. *The Scientific Heritage*. 2020. № 56. С. 13–32. doi: [10.24412/9215-0365-2020-56-2-13-32](https://doi.org/10.24412/9215-0365-2020-56-2-13-32)
13. Makhortykh S. Underground waveguide and the ecology problems of vibration forecast. In: *18th International Multidisciplinary Scientific GeoConference SGEM 2018* (30 June – 9 July, 2018, Albena, Bulgaria): Conference proceedings. Vol. 18. Geology. Applied and Environmental Geophysics. 2018. P. 1035–1042.