

УДК 004.9

# АВТОМАТИЗИРОВАННОЕ РАСПОЗНАВАНИЕ ЭЛЕКТРОЭНЦЕФАЛОГРАММ С ИСПОЛЬЗОВАНИЕМ МЕТОДА АНАЛИЗА ФРАКТАЛЬНОЙ ДИНАМИКИ

**Р. И. Полонников,**

доктор техн. наук, профессор

**Е. Л. Вассерман,**

канд. мед. наук, старший научный сотрудник

**Н. К. Карташев,**

научный сотрудник

Санкт-Петербургский институт информатики и автоматизации РАН

Приводятся результаты применения признаков, рассчитанных методом анализа фрактальной динамики, для обучаемой классификации электроэнцефалограмм. Путем соотнесения с результатами, полученными на случайных выборках, показана значимость положительных результатов распознавания.

**Ключевые слова** — цифровая электроэнцефалография, метод анализа фрактальной динамики, распознавание функциональных состояний, случайный поиск с адаптацией.

## Введение

Электроэнцефалография является одним из основных методов изучения и диагностики функций центральной нервной системы. Электроэнцефалограмма (ЭЭГ) — это запись суммарной электрической активности мозга. Она представляет собой многоканальный шумоподобный нестационарный сигнал. Анализируя ЭЭГ, врач-электроэнцефалографист делает заключение о функциональных и морфологических особенностях головного мозга обследуемого, о характере и выраженности имеющейся у него церебральной патологии. Эта работа плохо поддается автоматизации из-за сложности объекта исследования и порождаемого им сигнала, обилия трудноустраняемых артефактов, большого удельного веса слабоформализуемых эмпирических знаний и использования при принятии решения контекстной информации. Поэтому автоматизация анализа и интерпретации ЭЭГ до сих пор остается не решенной в полном объеме актуальной теоретической и практической задачей. Программные средства, созданные в результате ее решения, могут быть востребованы в телемедицине, медицине катастроф, военной медицине, а также при мониторинге ЭЭГ и скрининговых исследованиях.

Задача распознавания ЭЭГ в этом случае может быть сформулирована следующим образом. Требуется построить автоматизированную систему, спо-

собную на основе анализа ЭЭГ испытуемого высказать предположение о его принадлежности к одному из классов. Ключевым элементом подобной системы может стать метод, позволяющий сопоставить электроэнцефалограмме вектор числовых признаков, используя который можно было бы проводить обучаемую классификацию статистическими и другими методами. Авторами был предложен такой метод, получивший название «метод анализа фрактальной динамики» (АФД) [1, 2]. Он предполагает разбиение многоканальной ЭЭГ на 30 односекундных одноканальных фрагментов с последующим вычислением ряда ее параметров, которые изменяются во времени и пространстве (т. е. по поверхности головы) и доступны для вторичной статистической обработки. В целом с помощью метода можно получить набор информативных параметров ЭЭГ, которые характеризуют ее динамику как мультифрактала [3].

В данной работе обсуждаются результаты эксперимента по обучаемой классификации файлов ЭЭГ. Классификация проводилась с использованием наиболее информативных подсистем признаков из общего набора признаков, найденных методом АФД.

## Материал и методы

В нашем распоряжении была база из 109 ЭЭГ, зарегистрированных по 12 каналам с верхним гра-

ничным фильтром 30 Гц, постоянной времени 0,3 с и частотой дискретизации 200 Гц. ЭЭГ были разделены врачом-экспертом на несколько классов. Наиболее представительными оказались классы «все больные» (62 чел.) и «все здоровые» (47 чел.). Далее описываются эксперименты по распознаванию именно этих классов.

Все ЭЭГ были подвергнуты обработке методом АФД с использованием реализации алгоритмов метода, приведенной в работе [1]. В результате каждому файлу оказалась сопоставлен один или несколько векторов из  $n = 28$  информативных признаков (действительных чисел); число векторов соответствует количеству полных 30-секундных безартефактных фрагментов фоновой ЭЭГ, которые удалось выделить в файле.

Поскольку прямая физиологическая трактовка большинства АФД-признаков затруднительна, заранее невозможно выбрать те признаки, использование которых дает положительный эффект при распознавании объектов из выбранных классов, и исключить те, которые лишь ухудшают его результаты. Кроме того, число объектов в выборке было небольшим. Возникла необходимость решить хорошо известную задачу выбора наиболее информативной подсистемы (подмножества) из  $m$  признаков ( $m < n$ ). Чтобы избежать полного перебора, был использован алгоритм случайного поиска с адаптацией (СПА). Алгоритм СПА реализует направленный случайный поиск наиболее информативной подсистемы признаков. В отличие от метода Монте-Карло, в процессе поиска производится «поощрение» и «наказание» отдельных признаков в зависимости от результатов испытаний групп подсистем. Алгоритм СПА был реализован на языке MATLAB в полном соответствии с работой Г. С. Лбова [4].

Для распознавания использовались методы дискриминантного анализа: линейный классификатор, квадратичный классификатор, классификатор с использованием расстояния Махаланобиса и классический байесовский классификатор. Первые три представлены М-функцией classify, взятой из Statistics Toolbox пакета MATLAB (версия 6.5, SP 1). Метод Байеса был реализован в виде отдельной М-функции.

### Результаты

Из векторов информативных признаков, полученных при анализе файлов ЭЭГ, было составлено четыре обучающие и четыре контрольные выборки одинаковой величины в соответствии с вариантом разбиения на классы «все больные» (1-й класс) и «все здоровые» (2-й класс). Выборки набирались случайным образом так, чтобы количества объектов в классах совпадали. Попадания векторов признаков одного и того же файла ЭЭГ одновременно на обучение и на контроль не допускалось. Таким образом, количество объектов в каждой выборке равнялось 46 (по 23 объекта 1-го и 2-го класса).

Для каждой из четырех пар «обучающая выборка — контрольная выборка» было проделано следующее. Зафиксированы параметры алгоритма СПА и минимизируемая алгоритмом СПА функция — количество ошибок, допущенных примененным методом распознавания на обучающей выборке. Параметр  $m$  (размер подбираемой подсистемы признаков) варьировался от 2 до 9. Также перебирались 4 вышеуказанных метода распознавания. Для каждого сочетания  $m$  и метода распознавания производилось 7 запусков алгоритма СПА. Таким образом, всего было выполнено 896 запусков и составлена таблица испытаний, в которую для каждого запуска вносились следующие результаты: найденная подсистема признаков, значение критериальной функции в момент останова алгоритма, проценты правильных распознаваний (полученные применением каждого базового метода распознавания к обучающей и контрольной выборкам при использовании найденной подсистемы признаков), количество произведенных групп испытаний и машинное время, затраченное на вычисления. В таблице приведены некоторые характеристики, полученные путем усреднения результатов всех запусков алгоритма СПА.

Также была подсчитана частота вхождения отдельных информативных признаков в найденные алгоритмом СПА подсистемы как отношение количества подсистем, в которые входит признак, к общему числу проведенных запусков алгоритма.

Заключительный эксперимент проведен для проверки значимости полученных результатов.

### Интегральные результаты распознавания

Показатель	Разбиение на классы «все больные» — «все здоровые»	Случайное разбиение на классы
1	2	3
Средний процент правильных распознаваний, сделанных на контрольной выборке	68 %	51 %
Среднеквадратическое отклонение процента правильных распознаваний	6,2 %	6,3 %
Среднее количество испытаний за один запуск алгоритма СПА	3356	3570
Среднее машинное время, затрачиваемое на один запуск алгоритма СПА (характеристики ЭВМ: P3 667 МГц, ОЗУ 256 Мб)	32,7 с	34,4 с
Процент запусков алгоритма СПА, на которых не происходило останова до выполнения максимального числа испытаний (15 тыс.)	4 %	5 %

Было построено четыре пары обучающих и контрольных выборок из тех же объектов, которые входили в выборки, обсуждавшиеся выше. Объекты в них также были разделены поровну на два класса, но номер класса для каждого объекта определялся не экспертом, а генератором случайных чисел. Для этих «бесмысленных» выборок проделано то же количество запусков алгоритма СПА в тех же условиях, что и для основных выборок. Интегральные показатели для соответствующей таблицы испытаний приведены в столбце 3.

Если на контрольной выборке распознавалось в среднем 68% объектов для варианта разбиения на классы «все больные» — «все здоровые», то аналогичная процедура распознавания, примененная к случайной выборке из тех же объектов, привела лишь к 51% правильных распознаваний на контроле при среднеквадратических отклонениях в обоих случаях в 6%. Поэтому можно утверждать, что полученный для неслучайного разбиения на классы результат представляет собой не вычислительный артефакт, а отражение закономерности, присутствующей в исходных данных и выявленной использованными методами.

### Заключение

Таким образом, применение метода анализа фрактальной динамики для обработки ЭЭГ в сочетании с методом случайного поиска с адаптацией

для сокращения числа информативных признаков дает положительные результаты при распознавании указанных экспертом классов испытуемых. Значимость полученных результатов подтверждается путем соотнесения их с результатами распознавания случайных выборок. Мы также полагаем, что анализ частот встречаемости отдельных информативных признаков в «успешных» подсистемах позволит наметить направление поиска новых признаков и лучше понять физиологический смысл уже имеющихся.

### Литература

1. Вассерман Е. Л., Карташев Н. К., Полонников Р. И. Фрактальная динамика электрической активности мозга. СПб.: Наука, 2004. 208 с.
2. Polonnikov R. I., Wasserman E. L., Kartashev N. K. Regular developmental changes in EEG multifractal characteristics // International journal of neuroscience. 2003. Vol. 113. N 11. P. 1615–1639.
3. Шрёдер М. Фракталы, хаос, степенные законы. Миниатюры из бесконечного рая: Пер. с англ. / НИЦ «Регулярная и хаотическая динамика». Ижевск, 2001. 528 с.
4. Лбов Г. С. Методы обработки разнотипных экспериментальных данных. Новосибирск: Наука, 1981. 160 с.