

Адаптивный классификатор многомерных нестационарных сигналов на основе анализа динамических паттернов

08, август 2010

авторы: Трофимов А. Г., Скругин В. И.

УДК 519.25+004.9

atrofimov@list.ru

Национальный исследовательский ядерный университет "МИФИ"

Введение

Задачи анализа и классификации многомерных динамических данных получили широкое распространение в различных областях науки и техники, среди которых медицинская диагностика, распознавание речи, идентификация динамических систем, эконометрика и другие. В ряде случаев исследователю, работающему в этих областях, приходится иметь дело с принципиально нестационарными динамическими данными. Существующие методы цифровой обработки сигналов в основном ориентированы на анализ и обработку одномерных временных рядов, в то время как многомерным динамическим данным уделяется существенно меньшее внимание. Методы анализа и классификации, ориентированные на одномерные сигналы, зачастую оказываются неэффективными для многомерных данных или вообще не могут быть обобщены на многомерный случай. С другой стороны, нестационарность временных рядов накладывает до-

полнительные ограничения на возможность применения известных методов цифровой обработки сигналов. Как правило, методы обработки одномерных нестационарных сигналов сводятся к использованию скользящих оконных преобразований (в частности, вейвлет-преобразования) или к сегментации с последующим анализом выделенных квазистационарных участков. [1-3]

Построение классификатора временных рядов предполагает выполнение следующих этапов:

1) Предобработка сигналов. На этом шаге осуществляется фильтрация сигналов от шумов, устранение артефактов или другие операции, позволяющие на ранних этапах отсеять неинформативные данные, используя априорно известную информацию о природе сигнала.

2) Выделение характерных признаков сигнала. Целью этого этапа является переход от исходного пространства динамических данных к пространству многомерных статических данных. Как только такой переход сделан, задача классификации сигналов сводится к известной задаче классификации многомерных данных.

3) Классификация в пространстве характерных признаков сигналов. На этом этапе могут использоваться классические алгоритмы классификации многомерных статических данных, например метод дискриминантных функций, нейросетевой аппарат и другие [4,5].

В настоящей работе вопросы предобработки сигналов рассматриваться не будут. При работе с конкретными временными рядами исследователь, обычно, знает, какие алгоритмы предобработки нужно применить к этим сигналам, чтобы увеличить их информативность. Как правило, это какие-либо стандартные методы обработки сигналов или известные специальные методы для рядов рассматриваемой природы. Наибольший интерес в данной работе представляет проблема выделения характерных признаков сигнала для классификации.

Динамическая система, продуцирующая многомерный временной ряд, может находиться в различных функциональных состояниях или работать в различных режимах. Работа системы в том или ином функциональном состоянии находит отражение в наблюдаемых сигналах в виде паттернов. Как правило, моментам смены функционального состояния системы соответствуют

моменты резкого изменения динамических характеристик наблюдаемого временного ряда, т.е. моменты смены паттернов.

В данной работе предлагается методика расчёта характерных признаков многомерных нестационарных сигналов, основанная на выделении функциональных состояний динамической системы, продуцирующей сигнал. Анализ функциональных состояний и переходов между ними, а также соответствующих им паттернов, может послужить основой при классификации сигналов, а также позволить выявить скрытые закономерности, присутствующие в сигнале, и использовать их для лучшего понимания природы наблюдаемой динамической системы.

В работе рассматривается практическое применение предлагаемой методики для анализа и классификации многомерных нестационарных сигналов электроэнцефалограмм (ЭЭГ). Классификатор сигналов ЭЭГ является основой телекоммуникационных систем нового поколения – интерфейсов "мозг-компьютер" (ИМК). ИМК – системы, позволяющие осуществлять управление внешними техническими устройствами непосредственно с помощью электрической активности нейронов мозга, минуя мышцы [6]. Такие системы могут использоваться в медицинских целях – для организации канала коммутации с парализованными пациентами, а также в быту и промышленности – для предоставления нового канала взаимодействия с внешним миром. Современные разновидности этих устройств, основанные на анализе сигналов ЭЭГ, не поддерживают необходимых для промышленного внедрения скоростей передачи информации от мозга к электронному устройству, что в высокой степени обусловлено недостаточной эффективностью существующих алгоритмов анализа сигналов ЭЭГ, отражающих электрическую активность мозга.

Статья состоит из 4 разделов. В первом разделе приводится математическая постановка задачи классификации многомерных нестационарных сигналов. В разделе 2 вводится понятие динамического паттерна и излагаются теоретические основы предлагаемого алгоритма выделения характерных признаков сигнала для классификации на основе анализа динамических паттернов.

В разделе 3 описывается математическая модель нейросетевого классификатора в пространстве выделенных характерных признаков. Раздел 4 посвящён результатам применения предложенной методики для классификации сигналов ЭЭГ, соответствующих различным типам мыслительной активности человека, и обсуждению полученных результатов.

1. Постановка задачи

В связи с тем, что обработка данных осуществляется преимущественно в рамках классических ЭВМ, будем считать, что рассматриваемые сигналы либо имеют дискретную природу, либо дискретизованы. В этом случае значение каждого многомерного сигнала в момент времени t является вектором вида:

$$x(t) = (x_1(t), x_2(t), \dots, x_L(t)) \in X, \quad t = \overline{1, T}, \quad (1.1)$$

где $t = \overline{1, T}$ – номер такта дискретного времени, T – число дискретных отсчетов, L – размерность сигнала, $x_l(t)$ – значение l -ого компонента сигнала $\{x(t), t = \overline{1, T}\}$, в момент времени t , $l = \overline{1, L}$, $X \in R^L$ – множество значений временного ряда.

Пусть в результате P наблюдений получено множество многомерных временных рядов в общем случае различной длины, каждый из которых отнесён к одному из K классов. Таким образом, каждый класс представлен конечным числом сигналов. Обозначим P_k – число наблюдений, относящихся к k -му классу, $k = \overline{1, K}$. Таким образом, общее число рассматриваемых временных рядов $P = \sum_{k=1}^K P_k$.

Значение временного ряда, относящегося к k -му классу, полученное в результате p -го наблюдения в момент времени t , будем обозначать

$$x^{k,p}(t) = (x_1^{k,p}(t), x_2^{k,p}(t), \dots, x_L^{k,p}(t)) \in X, \quad t = \overline{1, T_p^k}, \quad p = \overline{1, P_k}, \quad k = \overline{1, K} \quad (1.2)$$

где T_p^k – число дискретных отсчетов p -го сигнала k -го класса, $x_l^{k,p}(t)$ – значение l -ого компонента сигнала $\{x^{k,p}(t), t = \overline{1, T_p^k}\}$, в момент времени t , $l = \overline{1, L}$.

Классификатор сигналов запишем как оператор, сопоставляющий сигналу $\{x(t), t = \overline{1, T}\}$ номер класса k , к которому он относится:

$$C[\{x(t), t = \overline{1, T}\}] = k, \quad k \in \{1, 2, \dots, K\}. \quad (1.3)$$

Задача построения классификатора временных рядов, таким образом, сводится к задаче нахождения оператора C на основе наблюдений вида (1.2), для которых результат действия этого оператора, т.е. номер класса, известен. В случае классификации статических данных построение оператора C , как правило, сводится к аппроксимации дискриминантных функций классов или границ между классами либо в пространстве самих данных, либо в некотором новом пространстве, рассчитанном в результате предобработки исходных данных [4]. Использование такого подхода к временным рядам, т.е. рассмотрение временной последовательности как вектора признаков, не представляется возможным в силу ряда причин, среди которых высокая размерность этого пространства (длина сигналов может достигать нескольких тысяч), а также различная длина наблюдаемых сигналов. Практически все методы классификации временных рядов предполагают сведение задачи классификации динамических данных к задаче классификации статических данных. Для такого сведения необходимо осуществить переход от временного ряда $\{x(t), t = \overline{1, T}\}$ к некоторому вектору признаков v размерности M , характеризующему этот временной ряд. Обозначая оператор этого перехода через Ψ , запишем

$$\Psi[\{x(t), t = \overline{1, T}\}] = z, \quad z \in R^M. \quad (1.4)$$

Основная проблема состоит в том, что не существует единого алгоритма такого перехода и в каждом конкретном случае исследователь сталкивается с необходимостью поиска оператора Ψ

и пространства признаков для эффективной классификации. Особенно остро эта проблема встает для случая нестационарных временных рядов.

Учитывая (1.3) и (1.4), задача классификации временных рядов, таким образом, состоит в поиске операторов Ψ и c :

$$\begin{cases} \Psi\left[\{x(t), t = \overline{1, T}\}\right] = z, & z \in R^M, \\ c[z] = k, & k \in \{1, 2, \dots, K\}. \end{cases} \quad (1.5)$$

Отметим, что задача нахождения этих операторов на основе результатов наблюдений является неоднозначной. Неточность в оценке оператора Ψ может быть компенсирована за счёт увеличения сложности оператора c , что, однако, может привести к ухудшению обобщающих способностей классификатора.

Качество построенного классификатора напрямую связано с его обобщающими способностями, которые могут быть оценены по результатам классификации данных тестовой выборки, не использовавшихся при нахождении операторов Ψ и c . Также при построении адаптивного классификатора нередко применяется валидационная выборка, назначение которой состоит в контроле процесса настройки классификатора по обучающим примерам, в частности, валидационная выборка может участвовать в критерии останова итеративной процедуры обучения. В связи с вышесказанным, перед построением классификатора разобьём множество наблюдаемых временных рядов на обучающую, тестовую и валидационную выборки в заданной пропорции и перенумеруем сигналы внутри этих выборок так, что каждый из них может быть представлен в форме (1.2).

2. Алгоритм выделения характерных признаков многомерного сигнала на основе анализа динамических паттернов

2.1. Выделение моментных характерных признаков сигнала

Оператор Ψ в модели (1.5) ставит в соответствие временному ряду $\{x(t), t = \overline{1, T}\}$ некоторый вектор характерных признаков z , не зависящий от времени t . Это означает, что все интересные особенности динамики значений временного ряда на рассматриваемом промежутке времени должны быть интегрированы в конечное множество статических показателей. Эти показатели могут быть рассчитаны как по исходным значениям временного ряда, так и по некоторым производным показателям, возможно, лучше отражающим эти особенности.

Введём оператор φ , определённый на множестве рассматриваемых временных рядов вида (1.1) и возвращающий в каждый момент времени $t, t = \overline{1, T}$, вектор, характеризующий рассматриваемый временной ряд в этот момент:

$$\varphi\left[\{x(\tau), \tau = \overline{1, T}\}, t\right] = v(t), \quad v(t) \in V, \quad t = \overline{1, T}. \quad (1.6)$$

Вектор $v(t)$, характеризующий временной ряд $\{x(t), t = \overline{1, T}\}$ в момент времени $t, t = \overline{1, T}$, будем называть вектором *моментных характерных признаков (МХП)*. Важно подчеркнуть, что значение вектора $v(t)$ в каждый момент времени t формируется в общем случае на основе значений сигнала $\{x(t), t = \overline{1, T}\}$ на всём рассматриваемом интервале времени.

Последовательность векторов $\{v(t), t = \overline{1, T}\}$ сама является многомерным временным рядом и определяет фазовую траекторию сигнала $\{x(t), t = \overline{1, T}\}$ в пространстве V моментных характерных признаков.

Цель перехода от пространства X значений исходного временного ряда к пространству МХП V состоит во введении в рассмотрение дополнительных параметров исходного временного ряда, характеризующих его динамические свойства в каждый момент времени. Выбор конкретного типа оператора φ осуществляется с учетом специфики анализируемых данных. Так, для сигналов гармонической природы в качестве элементов вектора МХП $v(t)$ могут быть выбраны зна-

чения спектральной плотности мощности, рассчитанные по результатам оконного преобразования Фурье с центром окна в момент времени t ; в задачах выделения специфических графоэлементов элементами вектора $v(t)$ могут являться коэффициенты непрерывного вейвлет-преобразования в заданном частотном диапазоне. В простейшем случае $v(t) = x(t)$, $t = \overline{1, T}$.

Отметим, что размерность вектора $v(t)$ может быть выше, чем размерность значений исходного временного ряда $x(t)$.

2.2. Понятие статического и динамического паттерна временного ряда

В связи с тем, что элементы вектора МХП $v(t)$ связаны с динамическими особенностями сигнала $x(t)$ в момент времени t , можно предположить, что близость векторов МХП $v(t_1)$ и $v(t_2)$ в пространстве V означает близость динамических свойств сигнала в моменты времени t_1 и t_2 . Так, на квазистационарных участках динамические свойства сигнала с течением времени меняются слабо, что приводит к группированию соответствующих значений МХП в некоторой локальной области или появлению регулярности переходов между несколькими локальными областями. В то же время изменение динамических свойств сигнала приведёт к появлению новых локальных областей в пространстве МХП и/или изменению регулярности в переходах между существующими областями.

Отметим, что обратное утверждение, вообще говоря, неверно. Близость динамических свойств сигнала в моменты времени t_1 и t_2 может не означать близость векторов $v(t_1)$ и $v(t_2)$ в пространстве МХП. Вектор МХП стационарного сигнала с неизменными динамическими свойствами может иметь в разные моменты времени существенно различные значения МХП.

Пусть пространство V некоторым образом разбито на Q областей S_1, \dots, S_Q таких, что

$$\begin{cases} \bigcup_{q=1}^Q S_q = V, \\ S_{q_1} \cap S_{q_2} = \emptyset, \quad q_1 = \overline{1, Q}, \quad q_2 = \overline{1, Q}, \quad q_1 \neq q_2; \end{cases}$$

Эти области в пространстве МХП V далее будем называть *статическими паттернами (СП)*.

В каждый момент времени t , $t = \overline{1, T}$, вектор МХП $v(t)$ будет принадлежать одной из областей S_1, \dots, S_Q . Будем говорить, что статический паттерн S_q проявляется в сигнале $x(t)$ в момент времени t , если соответствующий ему вектор МХП $v(t) \in S_q$. Такой паттерн S_q будем называть *активным* в момент времени t . Таким образом, каждый сигнал $\{x(t), t = \overline{1, T}\}$ вида (1.1) может быть представлен как последовательность номеров активных статических паттернов $\{q(t), t = \overline{1, T}\}$.

С течением времени активные статические паттерны для сигнала $\{x(t), t = \overline{1, T}\}$ некоторым образом сменяют друг друга. Долю времени, в течение которого СП S_q был активным, будем называть *индексом активности* p статического паттерна S_q :

$$p(S_q) = \frac{T(S_q)}{T}, \quad q = \overline{1, Q}, \quad (1.7)$$

где $T(S_q)$ – число тактов дискретного времени, в течение которого СП S_q был активным.

Таким образом, для временного ряда $\{x(t), t = \overline{1, T}\}$ может быть составлен вектор $p = (p(S_1), \dots, p(S_Q))$ индексов активностей всех рассматриваемых СП.

Классификация временного ряда может быть проведена по результатам анализа индексов активностей СП, присутствовавших в сигнале. Так, среди множества СП могут быть найдены паттерны, специфичные для временных рядов одного класса и неспецифичные для временных рядов всех остальных классов. Эта специфичность паттерна может выражаться в существенном превышении индекса его активности для сигналов одного из классов над другими.

Однако, как показывают эксперименты и как показано в работах [7,8], в ряде случаев классификация сигналов лишь на основе индексов активности СП не оказывается эффективной. Связано это с тем, что сигналы различных классов могут иметь одни и те же СП с примерно

одинаковыми индексами активности. Различие же между классами сигналов проявляется в последовательности смены этих статических паттернов.

Обобщим понятие индекса активности (1.7) на цепочку статических паттернов $\Sigma_{q_1, \dots, q_r} = (S_{q_1}, \dots, S_{q_r})$ длины r , которую далее будем называть *динамическим паттерном (ДП) r -го порядка*:

$$p(\Sigma_{q_1, \dots, q_r}) = p(S_{q_1}, \dots, S_{q_r}) = \frac{T(S_{q_1}, \dots, S_{q_r})}{T - r + 1}, \quad q_i \in \{1, \dots, Q\}, \quad i = \overline{1, r}, \quad (1.8)$$

где $T(S_{q_1}, \dots, S_{q_r})$ – количество вхождений цепочки (q_1, \dots, q_r) в последовательность номеров СП $\{q(t), t = \overline{1, T}\}$, соответствующую рассматриваемому сигналу $\{x(t), t = \overline{1, T}\}$. Индекс активности динамического паттерна фактически представляет собой относительную частоту встречаемости соответствующей последовательности переходов между статическими паттернами на рассматриваемом интервале времени.

С ростом порядка r число возможных динамических паттернов комбинаторно растёт, в то время как индексы активностей динамических паттернов, наблюдающихся в сигнале, стремятся к нулю – с увеличением длины цепочки вероятность её возникновения уменьшается. Для ДП порядка $r = 2$ индексы активности (1.8) образуют матрицу P размерности $Q * Q$. Визуально ДП 2-го порядка могут быть представлены в виде ориентированного графа с взвешенными ребрами, каждая вершина которого соответствует статическому паттерну, а веса рёбер равны индексам соответствующих ДП 2-го порядка.

2.3. Алгоритм формирования вектора характерных признаков сигнала

Индексы активности динамических паттернов являются интегральными характеристиками временного ряда на рассматриваемом интервале времени. Однако использование индексов активности всех ДП (при $r = 2$ – всех элементов матрицы P) в качестве характерных признаков сигнала для последующей классификации не представляется рациональным в силу нескольких

причин. Во-первых, некоторые ДП являются общими для сигналов всех классов, т.е. отражают некоторые общие свойства динамической системы, продуцирующей сигналы, и поэтому не представляют интереса с точки зрения классификации. Присутствие других ДП (с невысокими индексами активности) может являться индивидуальной особенностью конкретного сигнала или быть обусловлено действием различного рода случайных факторов и не может являться характеристикой класса сигналов. В связи с этим такие ДП также целесообразно исключить из дальнейшего рассмотрения. Во-вторых, общее число ДП r -го порядка может быть довольно большим, особенно при высоких порядках r , что усложняет процесс их классификации.

Введем понятие специфичности динамического паттерна Σ по отношению к паре классов (k_1, k_2) , $k_1 \in \{1, \dots, K\}$, $k_2 \in \{1, \dots, K\}$, $k_1 \neq k_2$. Под специфичностью ДП Σ будем понимать степень успешности классификации значений индексов активности этого ДП на обучающих сигналах, относящихся к классам k_1 и k_2 , с помощью некоторого классификатора. В качестве *показателя специфичности* динамического паттерна Σ по отношению к паре классов (k_1, k_2) , таким образом, выберем точность работы классификатора:

$$I_{k_1, k_2}(\Sigma) = \frac{1}{2} \left(\frac{n_{k_1}}{N_{k_1}} + \frac{n_{k_2}}{N_{k_2}} \right), \quad k_1 \in \{1, \dots, K\}, \quad k_2 \in \{1, \dots, K\}, \quad k_1 \neq k_2, \quad (1.9)$$

где n_{k_1} и n_{k_2} – число правильно классифицированных примеров классов k_1 и k_2 соответственно по значению индекса активности ДП Σ , а N_{k_1} и N_{k_2} – общее число примеров обучающей выборки соответствующих классов. Возможные значения показателя $I_{k_1, k_2}(\Sigma)$ лежат в отрезке $[0; 1]$. Значение $I_{k_1, k_2}(\Sigma) = 1$ означает разделимость классов k_1 и k_2 с помощью выбранного классификатора.

Наиболее простым классификатором индексов активности p динамического паттерна Σ на два класса k_1 и k_2 является линейный классификатор:

$$c[p(\Sigma)] = \begin{cases} k_1, & \text{если } p(\Sigma) \leq p_0, \\ k_2, & \text{в ином случае,} \end{cases} \quad \text{или} \quad c[p(\Sigma)] = \begin{cases} k_2, & \text{если } p(\Sigma) \leq p_0, \\ k_1, & \text{в ином случае,} \end{cases} \quad (1.10)$$

где значение разделяющего порога p_0 выбирается таким образом, чтобы обеспечить максимальное значение критерия (1.9).

Отсортируем динамические паттерны, встречающиеся в обучающих сигналах классов k_1 и k_2 , по значению показателя специфичности (1.9) в порядке убывания. Обозначим $\Sigma_{k_1, k_2}^{(i)}$ – ДП, имеющий i -ый номер в таком отсортированном ряду, а $\Sigma_{k_1, k_2}^* = \Sigma_{k_1, k_2}^{(1)}$ – наиболее специфичный динамический паттерн. Предполагая, что индекс ДП с максимальной специфичностью в наибольшей степени характеризует тип классифицируемого сигнала, будем использовать его в качестве характерного признака $z_{k_1 k_2}$ сигнала $\{x(t), t = \overline{1, T}\}$ для его классификации на два класса: k_1 и k_2 :

$$z_{k_1 k_2} = p(\Sigma_{k_1, k_2}^*). \quad (1.11)$$

Отметим, что для большей устойчивости классификации размерность вектора $z_{k_1 k_2}$ может быть увеличена:

$$z_{k_1 k_2} = (p(\Sigma_{k_1, k_2}^{(1)}), \dots, p(\Sigma_{k_1, k_2}^{(\lambda)})), \quad (1.12)$$

где λ – число используемых для классификации наиболее специфичных ДП. Далее будет рассматриваться вектор характерных признаков $z_{k_1 k_2}$ вида (1.11).

Задача линейной классификации сигналов на $K > 2$ классов может быть сведена к построению ряда линейных классификаторов на два класса вида (1.10), каждый из которых осуществляет классификацию в пространстве соответствующих признаков $z_{k_1 k_2}$, $k_1 = \overline{1, K}$, $k_2 = \overline{1, K}$, $k_1 \neq k_2$. Таким образом, вектор признаков z для классификации сигнала на K классов будет составлен из всех векторов $z_{k_1 k_2}$, $k_1 = \overline{1, K-1}$, $k_2 = \overline{k_1+1, K}$:

$$z = (z_{12}, \dots, z_{K-1,K}). \quad (1.13)$$

Размерность вектора z не превышает $K*(K-1) / 2$. Именно этот вектор предполагается использовать в качестве вектора характерных признаков сигнала $\{x(t), t = \overline{1, T}\}$ для его классификации.

2.4. Алгоритм оптимизации специфичности динамических паттернов

Для некоторых классов k_1 и k_2 , $k_1 = \overline{1, K}$, $k_2 = \overline{1, K}$, $k_1 \neq k_2$, выделенные наиболее специфичные динамические паттерны r -го порядка, образующие вектор $z_{k_1 k_2}$, могут иметь значения показателя специфичности (1.9), не позволяющие проводить классификацию с помощью выбранного классификатора с требуемой точностью. Причиной этому может быть неудачное разбиение пространства МХП V на статические паттерны. Индексы активности, а, следовательно, и специфичность ДП r -го порядка $\Sigma_{q_1, \dots, q_r} = (S_{q_1}, \dots, S_{q_r})$ зависят от формирующих его статических паттернов S_{q_1}, \dots, S_{q_r} , т.е. областей в пространстве МХП V . Изменение вида этих областей будет отражаться на значении показателя специфичности $I_{k_1 k_2}(\Sigma_{q_1, \dots, q_r})$, $k_1 = \overline{1, K}$, $k_2 = \overline{1, K}$, $k_1 \neq k_2$.

Поставим задачу разбиения пространства V на статические паттерны S_1, \dots, S_Q таким образом, чтобы максимизировать точность работы выбранного классификатора в пространстве признаков z , определяемом выражением (1.13). В связи с тем, что точность классификации на K классов связана с точностью попарной классификации каждой пары классов, введём следующий показатель точности:

$$I = \prod_{\substack{k_1=1, \\ k_2=1, \\ k_1 \neq k_2}}^K I_{k_1 k_2}(\Sigma_{k_1, k_2}^*). \quad (1.14)$$

Значение показателя I , равное 1, будет означать разделимость всех пар классов в пространстве характерных признаков z с помощью выбранного классификатора.

Задача поиска оптимальных статических паттернов состоит в разбиении пространства признаков МПХ V на Q областей S_1, \dots, S_Q таким образом, чтобы максимизировать значение точности классификации (1.14):

$$\begin{cases} I \rightarrow \max_{(S_1, \dots, S_Q)}, \\ \bigcup_{q=1}^Q S_q = V, \\ S_{q_1} \cap S_{q_2} = \emptyset, \quad q_1 = \overline{1, Q}, \quad q_2 = \overline{1, Q}, \quad q_1 \neq q_2; \end{cases} \quad (1.15)$$

Задачу (1.15) можно свести к задаче разбиения множества векторов МХП сигналов обучающей выборки, каждый из которых представляет собой точку в пространстве V , на Q групп. В связи с тем, что число точек в пространстве V может достигать десятков или даже сотен тысяч, перед решением оптимизационной задачи (1.15) сократим её размерность. Для этого проведём предварительную кластеризацию данных в пространстве V на N кластеров ($N \gg Q$), используя стандартные методы кластеризации, например, метод k -средних или самообучающиеся карты Кохонена [5,9].

Обозначим s_1, \dots, s_N – кластеры, полученные в результате кластеризации данных в пространстве V , которые далее будем называть *микросостояниями*. Таким образом, задача (1.15) сводится к задаче разбиения N кластеров на Q групп (задаче формирования Q статических паттернов из N микросостояний) таким образом, чтобы максимизировать критерий (1.14). Для решения этой комбинаторной задачи могут быть применены различные эвристические методы. В данной работе для её решения предложено использовать генетические алгоритмы (ГА) [10].

На рис. 1 представлена последовательность операций, выполняемых при формировании вектора характерных признаков z для временного ряда $\{x(t), t = \overline{1, T}\}$. Эта последовательность фактически определяет алгоритм построения оператора Ψ в модели (1.5).

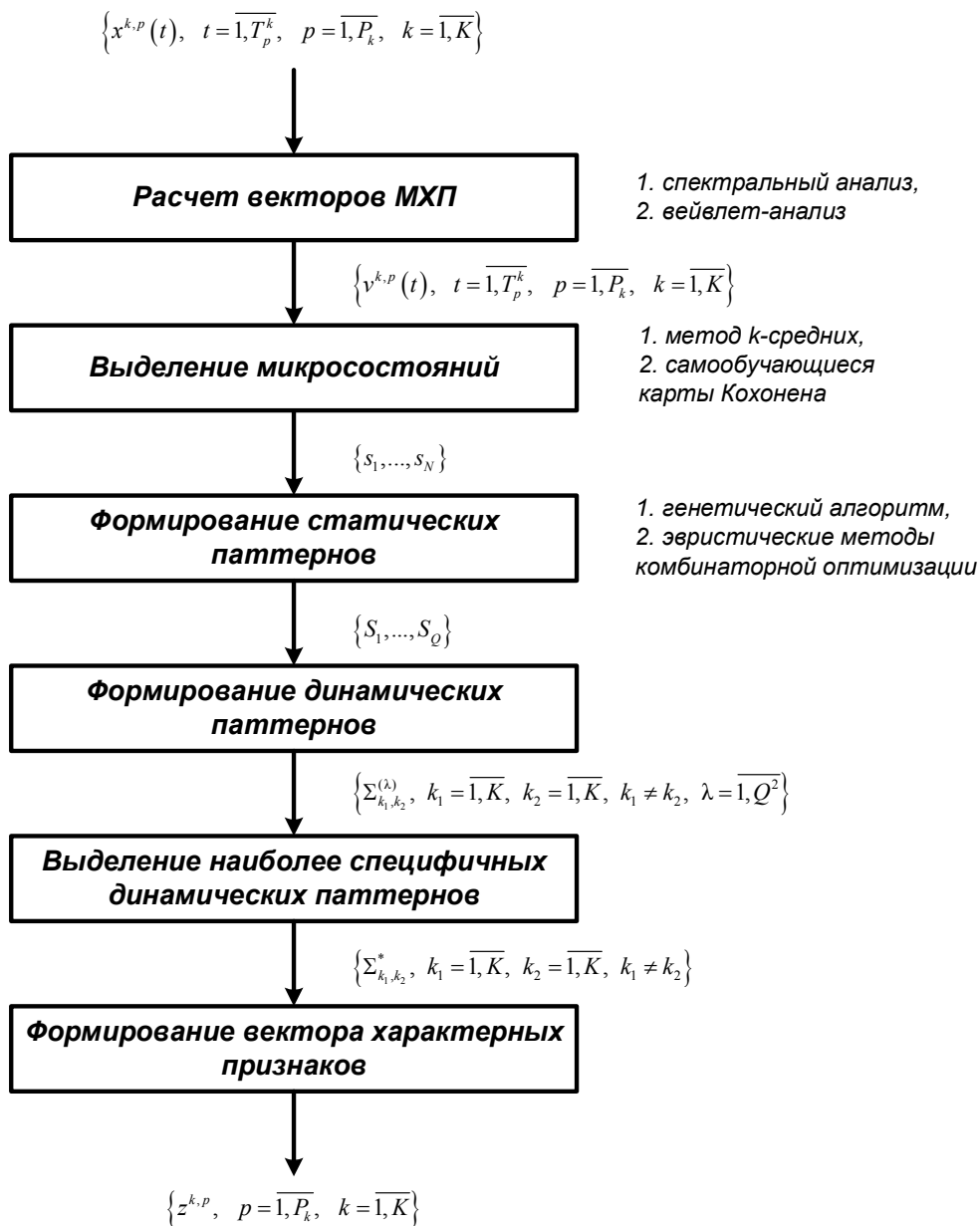


Рис. 1. Схема формирования вектора характерных признаков $z^{k,p}$ для многомерного временного ряда $\{x^{k,p}(t), t = \overline{1, T_p^k}\}$ из обучающей выборки, $p = \overline{1, P_k}$, $k = \overline{1, K}$. Справа указаны методы, которые могут быть использованы при выполнении соответствующей операции.

3. Классификация в пространстве характерных признаков

После выбора пространства характерных признаков z , следующим шагом алгоритма классификации является построение собственно классификатора данных $c[z]$ в этом пространстве. В связи с тем, что при формировании вектора z уже был выбран и использован классификатор для разделения пар классов по значениям индексов ДП, на данном этапе для разделения K классов тип классификатора целесообразно сохранить. Так, при использовании линейного классификатора (1.10) в расчётах показателя (1.9) классификатор $c[z]$ также целесообразно выбирать линейным.

В настоящей работе в качестве классификатора $c[z]$ предложено использовать многослойный персептрон [5]. Число входов персептрона определяется размерностью вектора z , рассчитываемого по формуле (1.13), число выходов определяется числом классов. Обучающая выборка, таким образом, представляет собой совокупность векторов

$$\{(z^{k,p}; \sigma^{k,p}), \quad p = \overline{1, P_k}, \quad k = \overline{1, K}\},$$

где $z^{k,p}$ – вектор характерных признаков для сигнала $\{x^{k,p}(t), t = \overline{1, T_p^k}\}$, $\sigma^{k,p} = (\sigma_1^{k,p}, \dots, \sigma_K^{k,p})$ – биполярный вектор желаемых выходов персептрона, определяющий принадлежность входного вектора k -му классу:

$$\sigma_i^{k,p} = \begin{cases} 1, & \text{если } i = k, \\ -1, & \text{иначе.} \end{cases}$$

Математическая модель многослойного персептрона имеет вид [5]:

$$\begin{cases} y = s^R, \\ s^i = f_i(h^i), \\ h^i = \tilde{W}^i \tilde{s}^{i-1}, \\ \tilde{s}^i = (1; s^i), \\ s^0 = z, \\ i = \overline{1, R}, \end{cases}$$

где R – число слоёв многослойного персептрона, z – вектор входов, y – вектор выходов, s^i – вектор выходов i -го слоя нейронов, h^i – вектор потенциалов i -го слоя нейронов, \tilde{W}^i – расширенная матрица синаптических коэффициентов m -го слоя нейронов, f_i – активационная характеристика i -го слоя нейронов.

В результате работы классификатора подаваемый на его вход пример будет отнесён к одному из K классов либо будет принято решение о невозможности классификации. Построение классификатора проводится на основе данных обучающей выборки, оценка качества его работы – на основе данных тестовой выборки. Для оценки качества классификации будем использовать показатель

$$I = \frac{1}{K} \sum_{k=1}^K \frac{n_k}{N_k}, \quad (1.16)$$

где n_k – число верно распознанных примеров k -ого класса при предъявлении N_k сигналов этого класса. Значение показателя (1.16) равное 1 означает правильную классификацию всех предъявленных на вход классификатора примеров.

4. Результаты экспериментальных исследований

4.1. Описание исходных данных

Экспериментальные исследования предложенных алгоритмов классификации многомерных нестационарных временных рядов проводились на сигналах электроэнцефалограмм (ЭЭГ), полученных в лаборатории высшей нервной деятельности человека Института высшей нервной деятельности и нейрофизиологии Российской Академии Наук. Каждая ЭЭГ соответствует процессу решения испытуемым задачи одного из $K = 4$ типов, связанных с различными типами мышления: планиметрическое мышление (тип ‘G’), пространственное мышление (тип ‘C’), составление предложений (тип ‘S’) и исключение лишнего слова из перечня (тип ‘W’).

Испытуемый решил примерно по 70 задач каждого типа, потратив на каждую от четырех секунд до нескольких десятков секунд. Все сигналы ЭЭГ были оцифрованы с частотой дискретизации $\omega_0 = 250$ Гц и очищены от артефактов. Запись ЭЭГ велась по 22 каналам, из которых информативными являются $L = 19$. Все сигналы были разбиты на обучающую и тестовую выборки, объёмы которых соответственно равны $P^{обуч} = 48+55+42+53 = 198$, $P^{тест} = 21+24+19+23 = 87$ ($P = P_1 + P_2 + P_3 + P_4$).

4.2. Расчет векторов моментных характерных признаков

В связи с тем, что электрическая активность мозга имеет гармоническую природу [11], для формирования моментных характерных признаков многомерного сигнала ЭЭГ предложено использовать непрерывное вейвлет-преобразование (НВП) с материнской вейвлет-функцией Морле, имеющей аналитический вид [12]:

$$\psi(t) = e^{-\frac{t^2}{2}} e^{i2\pi t}.$$

Расчёт вейвлет-коэффициентов сигнала $\{x(t), t = \overline{1, T}\}$, производился по формуле

$$W(a, b) = \frac{1}{\sqrt{a}} \sum_{t=0}^{T-1} x(t) \int_{t\Delta T}^{(t+1)\Delta T} \psi^* \left(\frac{\tau - b}{a} \right) d\tau,$$

где ΔT – интервал дискретности, a, b – параметры масштаба и временного сдвига соответственно, $\psi(t)$ – материнская вейвлет-функция, оператор $*$ означает комплексное сопряжение.

Для представления результатов НВП сигнала l -го канала ЭЭГ в памяти компьютера используется матрица $W^l = \{w_{ij}^l\}$ вейвлет-коэффициентов, рассчитанных в узлах дискретной сетки на плоскости (a, b) : $w_{ij}^l = W^l(a_i, b_j)$, $l = \overline{1, L}$.

На рис. 2 показан пример визуального представления матрицы квадратов вейвлет-коэффициентов (вейвлетограммы), рассчитанной для одного из каналов ЭЭГ. Каждый столбец матрицы отражает распределение энергии сигнала по частотам в фиксированный момент вре-

мени. Каждая строка отражает распределение энергии сигнала на фиксированной частоте по времени.

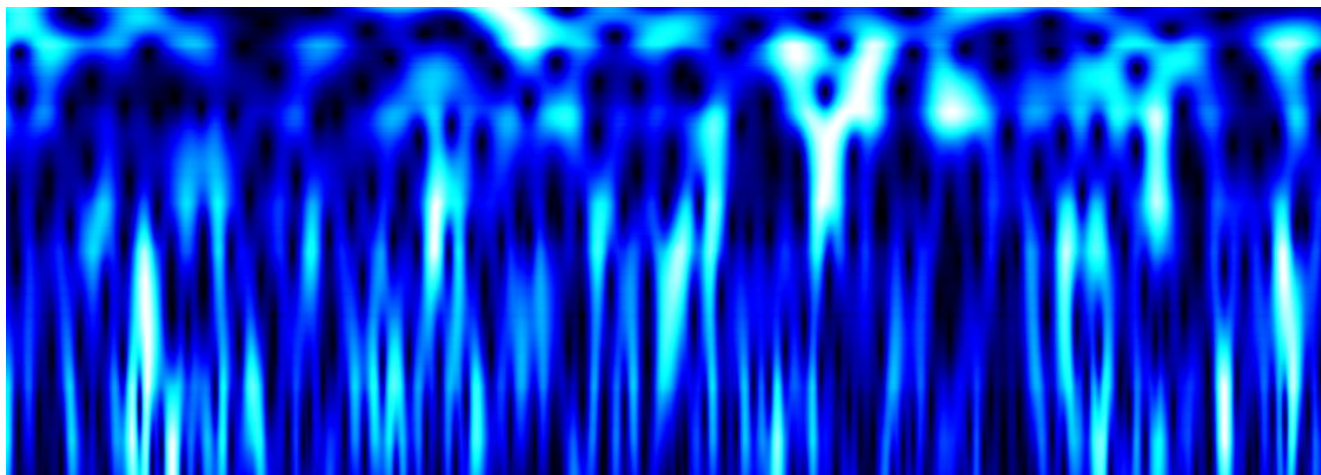


Рис. 2. Вейвлетограмма сигнала ЭЭГ в диапазоне частот от 5 Гц до 30 Гц; по оси абсцисс отложены значения временных сдвигов b , по оси ординат – масштабных сдвигов a , соответствующих частотам.

Поскольку сигналы ЭЭГ являются многомерными, вектор МХП ЭЭГ $v(t)$ составлен из значений энергии сигнала каждого из каналов ЭЭГ в заранее заданных частотных диапазонах $\Omega_1, \dots, \Omega_\mu$ в момент времени t :

$$v(t) = (v_1(t), \dots, v_L(t)), \quad v_l(t) = (v_{l1}(t), \dots, v_{l\mu}(t)), \quad l = \overline{1, L},$$

$$v_{lm}(t) = \sum_{i \in \Omega_m} |w_{it}'|^2, \quad l = \overline{1, L}, \quad m = \overline{1, \mu}.$$

Размерность M вектора МХП $v(t)$ равна $\mu * L$, где μ – число выбранных частотных полос, L – число каналов ($L = 19$). В работе исследовался диапазон частот от 5 до 30 Гц, что соответствует диапазонам альфа- и бета-ритма. В качестве элементов вектора МХП выбраны моментные значения энергии в частотных полосах шириной 0.5 Гц в рассматриваемом диапазоне. Таким образом, размерность вектора МХП $v(t)$ равна $M = 50 * 19 = 475$.

4.3. Выделение микросостояний

Для каждого обучающего сигнала в каждый момент времени были рассчитаны векторы МХП $\{v^{k,p}(t), t = \overline{1, T_p^k}\}$, $p = \overline{1, P_k}$, $k = \overline{1, K}$. Общее число векторов в пространстве МХП оказалось примерно равным 182000. Для сокращения их числа была проведена их кластеризация на $N = 200$ кластеров с использованием самообучающихся карт Кохонена с механизмом утомления *CWTA (Consciousness Winner Takes All)* [5]. В процессе самообучения уровень соседства экспоненциально менялся от $\lambda_{max} = 100$ до $\lambda_{min} = 0,001$, коэффициент обучения линейно менялся от $\mu_{max} = 1$ до $\mu_{min} = 0,001$. Самообучение продолжалось в течение 100 эпох.

В результате процесса кластеризации векторов МХП были сформированы $N = 200$ микросостояний, из которых далее формировались статические паттерны. На рис. 3 показано распределение числа точек в кластере по кластерам после завершения процедуры кластеризации.

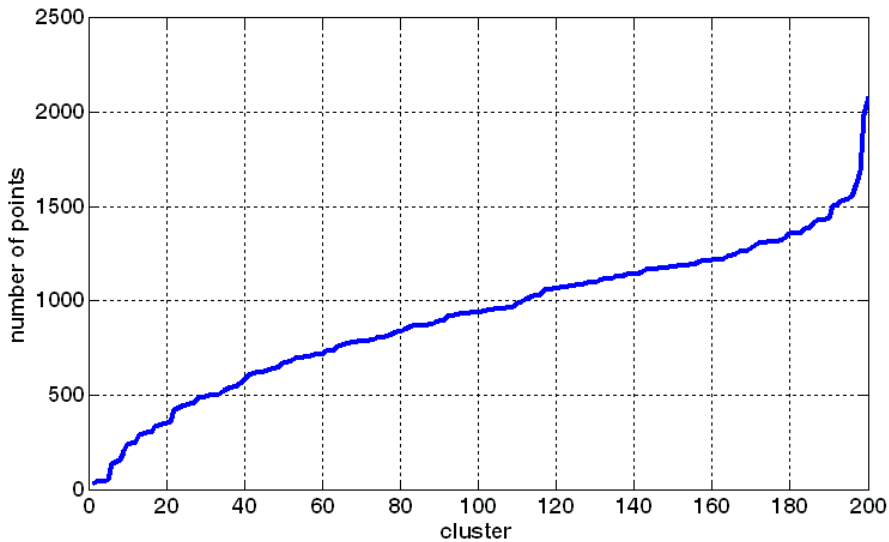


Рис. 3. Распределение числа точек в кластере по $N = 200$ кластерам после завершения процедуры кластеризации.

4.4. Формирование статических паттернов

Формирование статических паттернов S_1, \dots, S_Q из микросостояний s_1, \dots, s_N предполагает решение оптимизационной задачи (1.15) разбиения N микросостояний на Q групп, при котором обеспечивается оптимальное значение критерия (1.14). Для решения этой оптимизационной задачи был использован генетический алгоритм.

В связи с тем, что каждое микросостояние должно входить ровно в один статический паттерн, сформируем хромосому следующим образом:

$$g = (g_1, \dots, g_N), \quad g_i \in \{1, \dots, Q\}, \quad i = \overline{1, N},$$

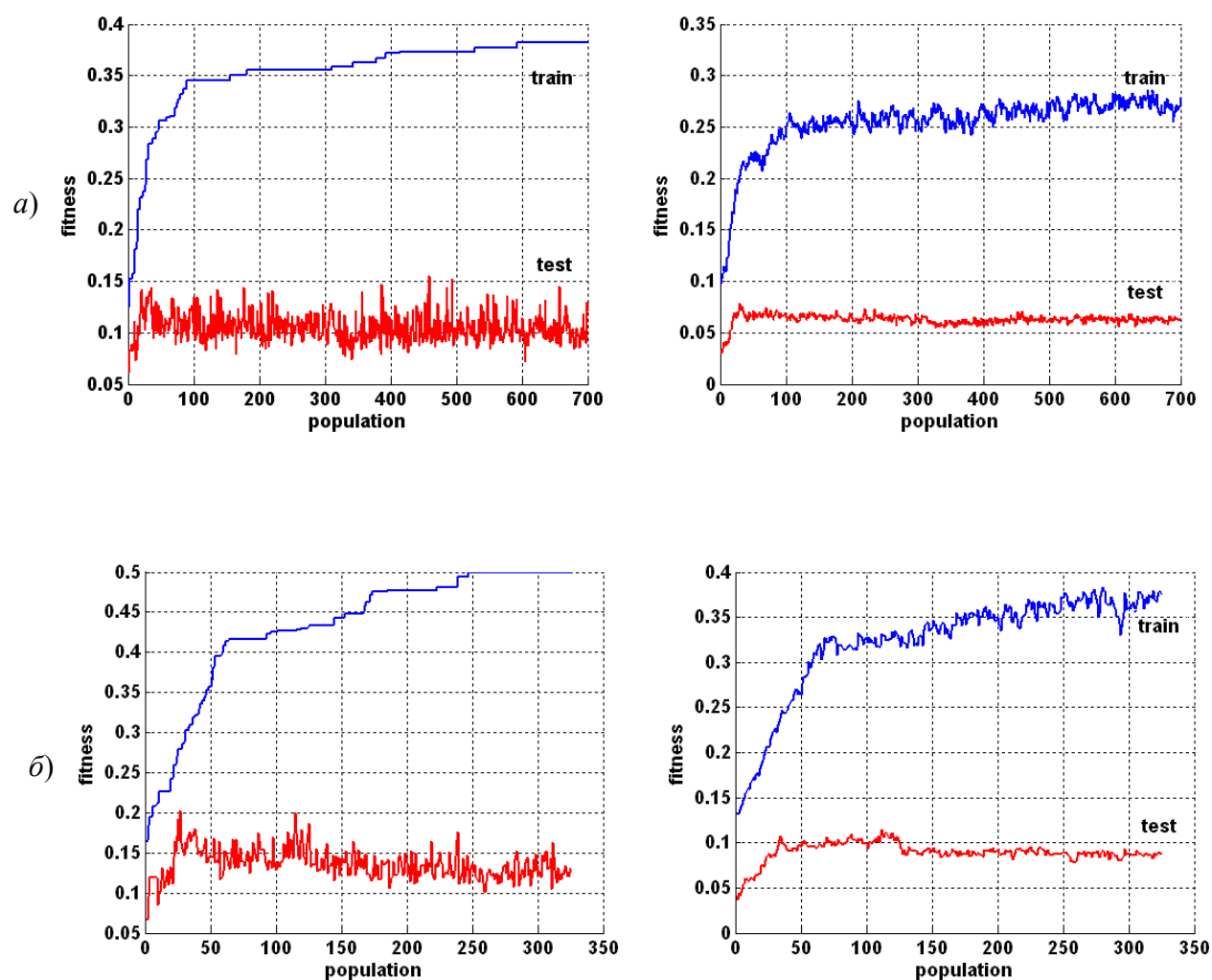
где g_i – i -ый ген хромосомы g , представляющий собой целое число в диапазоне от 1 до Q . Значение $g_i = q$ означает, что микросостояние s_i принадлежит статическому паттерну S_q . Каждая хромосома, таким образом, однозначно определяет разбиение множества микросостояний $\{s_1, \dots, s_N\}$ на статические паттерны S_1, \dots, S_Q . Генетические операции над такими хромосомами определяются аналогично генетическим операциям для бинарных генетических алгоритмов.

Отметим, что никакой априорной информации о числе статических паттернов в пространстве МХП нет, в связи с этим определение числа статических паттернов, при которых наблюдается приемлемая точность классификации, определялось экспериментально. На рис. 4 показаны графики зависимости значений приспособленности лучшей хромосомы в популяции и средней приспособленности в популяции от номера популяции (такта работы генетического алгоритма) для заданного числа статических паттернов Q , рассчитанные для сигналов обучающей и тестовой выборки. Отметим, что лучшие хромосомы выбирались из всех хромосом популяции, при этом лучшие хромосомы на обучающей и тестовой выборках в общем случае являются разными хромосомами.

Исследование поведения на тестовой выборке хромосомы, обладающей лучшей приспособленностью на обучающей выборке, показало, что практически во всех случаях её приспособленность на тестовой выборке не является лучшей, иными словами в популяции практически

всегда существует хромосома с лучшей приспособленностью на тестовой выборке, при этом не обладающей лучшей приспособленностью на обучающей выборке. На рис. 5 представлены графики зависимостей значений приспособленности на обучающей и тестовой выборках для двух хромосом – показывающей лучший результат на обучающей выборке и лучший результат на тестовой выборке при $Q = 15$. Подобная картина наблюдается и при других значениях Q .

В работе использованы следующие параметры генетического алгоритма. Длина хромосомы – $N = 200$, численность популяции – 40, алгоритм формирования родительского пула по методу рулетки, кроссинговер равномерный с вероятностью обмена генами 0.5, элитная стратегия селекции, вероятность мутации – 0.05.



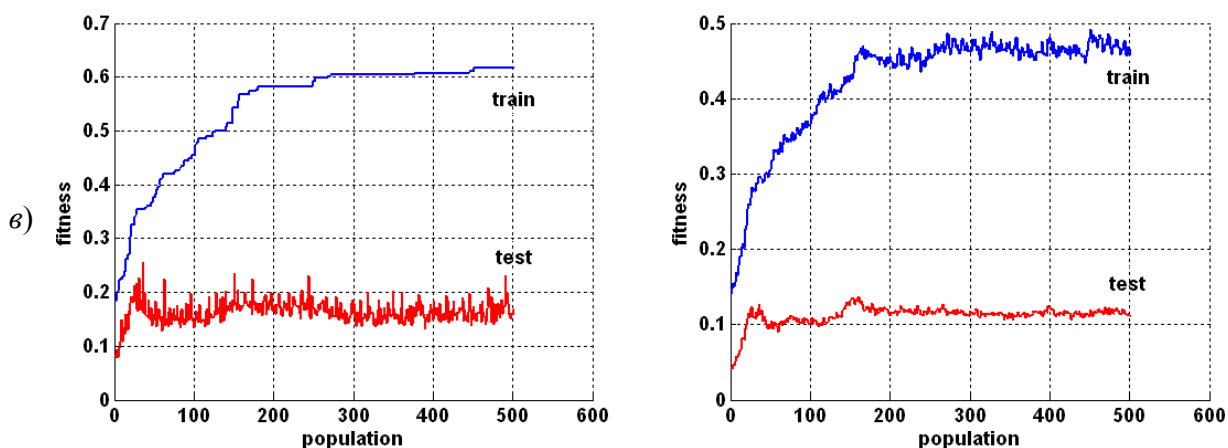


Рис. 4. Графики зависимости значений приспособленности лучшей хромосомы в популяции (слева) и средней приспособленности в популяции (справа) от номера популяции (такта работы генетического алгоритма) для заданного числа статических паттернов $Q = 5$ (а), 10 (б), 15 (в). Значения рассчитаны для сигналов обучающей (*train*) и тестовой (*test*) выборок.

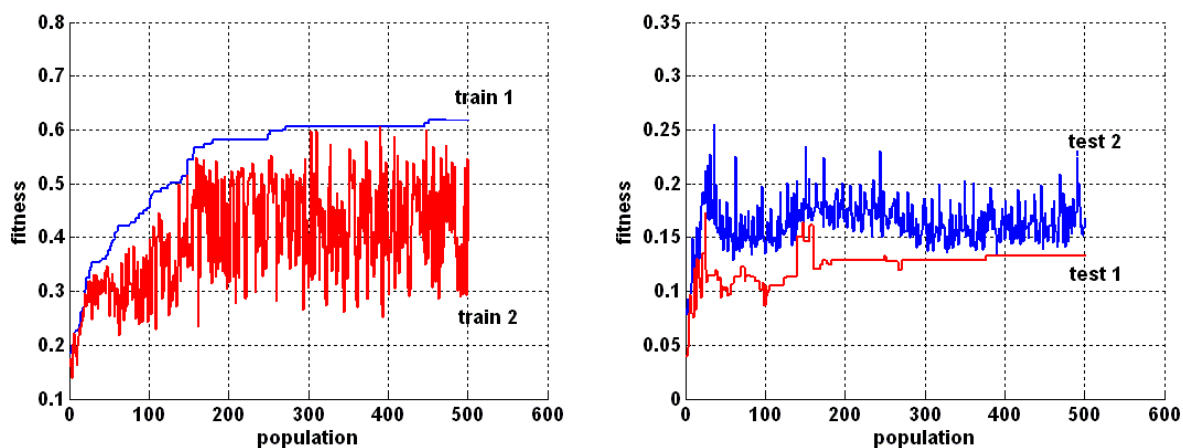


Рис. 5. Графики зависимости значений приспособленности на обучающей (*train*) и тестовой (*test*) выборках для двух хромосом – хромосомы с лучшей приспособленностью на обучающей выборке (графики *train 1* и *test 1*) и хромосомы с лучшей приспособленностью на тестовой выборке (графики *train 2* и *test 2*). Число статических паттернов $Q = 15$.

Из графиков видно, что с ростом числа статических паттернов значение функции приспособленности лучшей хромосомы, полученной в результате работы генетического алгоритма, растёт. Отметим, что в процессе работы генетического алгоритма, несмотря на рост приспособленности лучшей хромосомы и средней приспособленности хромосом в популяции на обучающей выборке, на тестовой выборке эти значения практически перестают изменяться уже после 100-200 итераций генетического алгоритма. Кроме того, значения функции приспособленности, рассчитанные по данным тестовой выборки, для хромосомы, имеющей лучшую приспособленность на обучающей выборке, также практически перестают изменяться, что говорит о неспособности эволюционного процесса к дальнейшему улучшению обобщающих свойств найденного решения. Значения критерия (1.14) после окончания работы генетического алгоритма на обучающей выборке при $Q = 5, 10, 15$ равны соответственно $I = 0.37, 0.50, 0.62$, рассчитанные для этих же хромосом значения приспособленности на тестовой выборке – $I = 0.11, 0.11, 0.14$ соответственно.

Для дальнейшего анализа была отобрана лучшая хромосома ($I^{обуч} = 0.62, I^{тест} = 0.14$). Эта хромосома определяет разбиение $N = 200$ микросостояний на $Q = 15$ статических паттернов. На рис. 6 приведено распределение числа микросостояний по найденным статическим паттернам, рассчитанное для найденного разбиения.

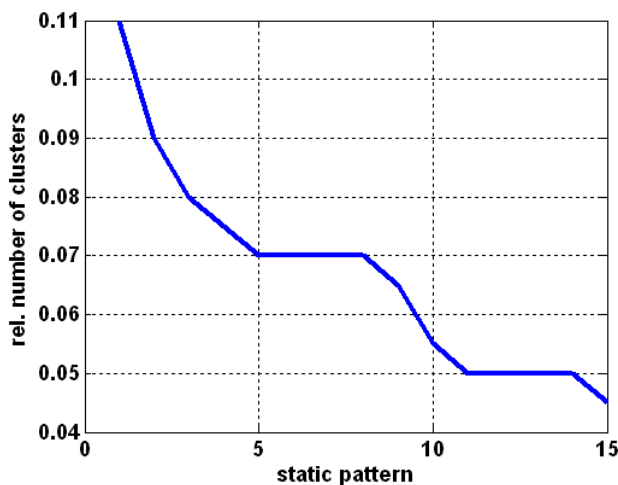


Рис. 6. Распределение числа микросостояний по статическим паттернам для лучшей хромосомы. По оси абсцисс отложен номер статического паттерна, по оси ординат – отсортированное относительно число микросостояний. Общее число микросостояний $N = 200$.

4.5. Выделение наиболее специфичных динамических паттернов

Как только сформированы статические паттерны S_1, \dots, S_Q , далее из них формируются динамические паттерны 2-го порядка $\Sigma_{q_1 q_2} = (S_{q_1}, S_{q_2})$, $q_1 \in \{1, \dots, Q\}$, $q_2 \in \{1, \dots, Q\}$. Для каждого из полученных динамических паттернов на основе сигналов обучающей выборки рассчитывается показатель специфичности (1.9) по отношению к классам k_1 и k_2 , $k_1 \in \{1, \dots, K\}$, $k_2 \in \{1, \dots, K\}$, $k_1 \neq k_2$. В качестве классификатора двух классов использован линейный классификатор вида (1.10). На рис. 7 приведены графики отсортированных значений показателя специфичности для некоторых пар классов в зависимости от порядкового номера динамического паттерна 2-го порядка. Общее число ДП 2-го порядка $Q*Q = 15*15 = 225$.

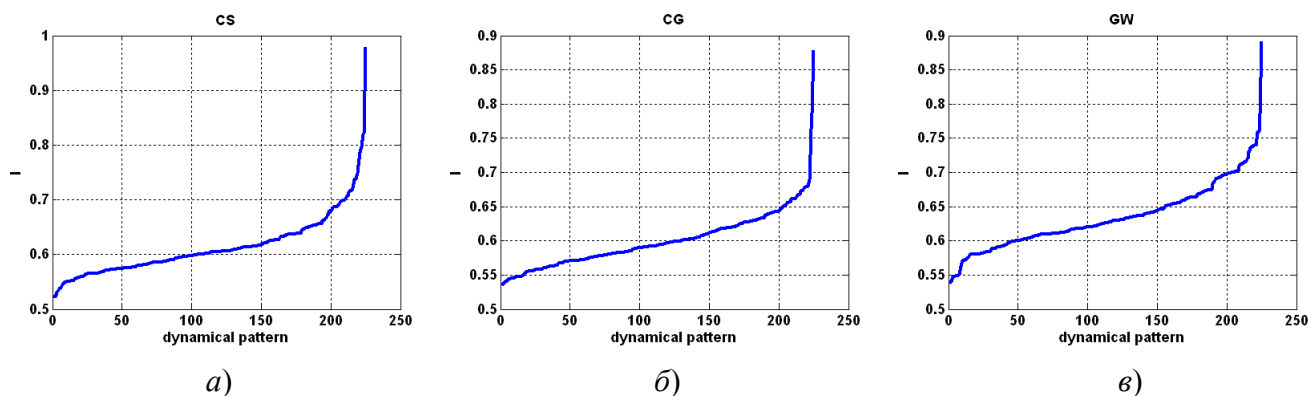


Рис. Графики отсортированных значений показателя специфичности динамических паттернов 2-го порядка для классов 'C', 'S' (а), 'C', 'G' (б), 'G', 'W' (в), рассчитанные для сигналов обучающей выборки. По оси абсцисс отложен порядковый номер динамического паттерна. Число соответствующих статических паттернов $Q = 15$.

В табл. 1 приведены наиболее специфичные ДП Σ_{k_1, k_2}^* для каждой пары классов (k_1, k_2) , рассчитанные по данным обучающей и тестовой выборок, и указаны значения показателя специфичности для них.

Таблица 1

Наиболее специфичные ДП 2-го порядка и значения показателей специфичности

	<i>CS</i>	<i>CG</i>	<i>CW</i>	<i>SG</i>	<i>SW</i>	<i>GW</i>
<i>ДП</i>	(9,9)	(9,3)	(9,9)	(12,12)	(2,11)	(15,12)
<i>обуч.</i>	0.9792	0.8782	0.9527	0.9589	0.8735	0.8913
<i>тест.</i>	0.8728	0.7473	0.8928	0.6751	0.5958	0.7179

Из табл. 1 видно, что наиболее хорошо разделимы классы 'C', 'S', наименее хорошо – классы 'S', 'W'. Этот результат согласуется с нейрофизиологическими представлениями – когнитивные задачи на составление предложений (тип 'S') и исключение лишнего слова из перечня (тип 'W') оба являются вербальными и задействуют схожую мыслительную активность, в то время как при решении задач на пространственное мышление (тип 'C') и на составление предложений (тип 'S') активизируются принципиально различные участки коры головного мозга.

4.6. Формирование вектора характерных признаков и классификация

Вектор характерных признаков z в соответствии с (1.13) формируется из значений индексов активности наиболее специфичных ДП, приведённых в табл. 1. В связи с тем, что для классов 'C', 'S' и 'C', 'W' оказался один и тот же наиболее специфичный ДП, размерность вектора z равна 5. По построению каждый элемент вектора z является оптимальным с точки зрения попарного линейного разделения сигналов на соответствующую пару классов. На рис. 8 приведено рас-

пределение проекций вектора z , рассчитанных для сигналов обучающей выборки, на прямую, соответствующую выбранной паре классов. На рис. 9 приведено аналогичное распределение проекций вектора z на плоскость.

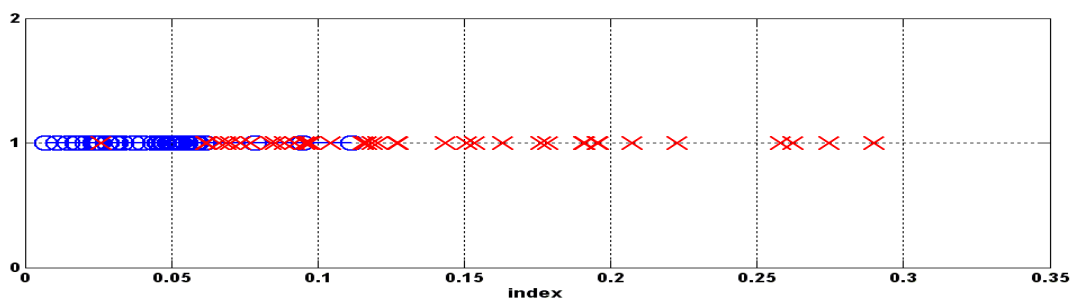


Рис. 8. Распределение проекций вектора z , рассчитанных для сигналов обучающей выборки, на прямую, соответствующую паре классов 'C', 'S'. Фактически по оси абсцисс отложены значения индексов ДП (9,9) для сигналов класса 'C' (помечены 'o') и сигналов класса 'S' (помечены 'x').

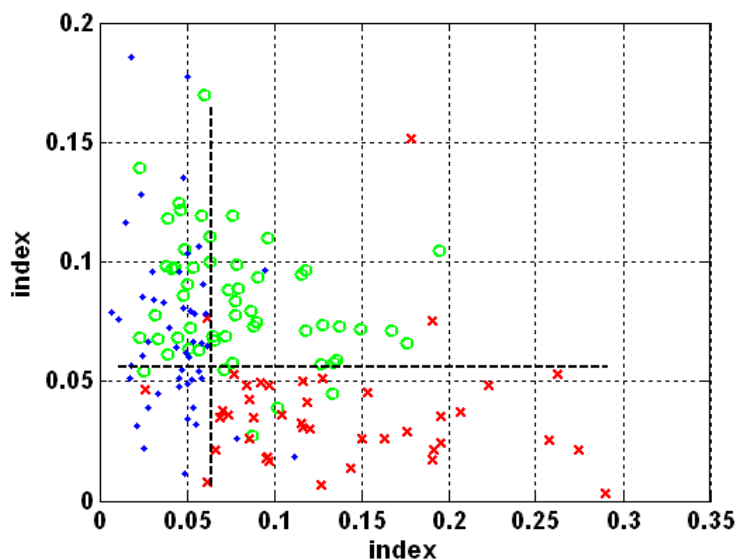


Рис. 9. Распределение проекций вектора z , рассчитанных для сигналов обучающей выборки, на плоскость индексов ДП (9,9) – ось абсцисс, и (12,12) – ось ординат. ДП (9,9) является наиболее

лее специфичным для классов 'C', 'S', ДП (12,12) – для классов 'S', 'G'. Символами 'o', 'x', '•' помечены индексы классов 'G', 'S', 'C' соответственно.

Из рис. 9 видно, что классы 'C', 'S' хорошо разделимы вертикальной прямой (по значениям индексов ДП (9,9)), а классы 'G', 'S' хорошо разделимы горизонтальной прямой (по значениям индексов ДП (12,12)). Однако классы 'C', 'G' в указанной плоскости разделимы плохо – наиболее специфичный ДП для этих классов в соответствии с табл. 1 – (9,3).

Векторы характерных признаков $z^{k,p}$, рассчитанные для всех сигналов обучающей выборки, используются для обучения многослойного персептрона, предназначенного для их классификации. Объёмы обучающей и тестовой выборок соответственно равны $P^{обуч} = 198$, $P^{тест} = 87$. Обучение продолжалось по методу Левенберга-Маркардта в течение 2000 эпох.

Число слоёв и нейронов в слоях многослойного персептрона подбиралось экспериментально. Наилучшие результаты по точности работы персептрона на тестовой выборке наблюдались при числе слоёв $\mu = 3$ и распределении нейронов по слоям: $N_1 = 8$, $N_2 = 5$, $N_3 = 4$. Активационные характеристики скрытых слоёв – сигмоиды, выходного слоя – линейная. Достигнутые значения показателя точности (1.16) на обучающей и тестовой выборках соответственно $I^{обуч} = 0.81$, $I^{тест} = 0.56$. Доля числа правильно распознанных примеров каждого класса (значения n_k/N_k в формуле (1.16)) обученным нейросетевым классификатором на обучающей и тестовой выборках приведена в табл. 2.

Таблица 2

Доля числа правильно распознанных примеров каждого класса обученным нейросетевым классификатором на обучающей и тестовой выборках.

Класс	'C'	'G'	'S'	'W'
обуч.	0.9792	0.6182	0.8810	0.7547

<i>тест.</i>	0.7619	0.2917	0.5263	0.6522
--------------	--------	--------	--------	--------

На рис. 10 приведено распределение значений выходов обученного нейросетевого классификатора при подаче на вход сигналов обучающей и тестовой выборок различных классов.

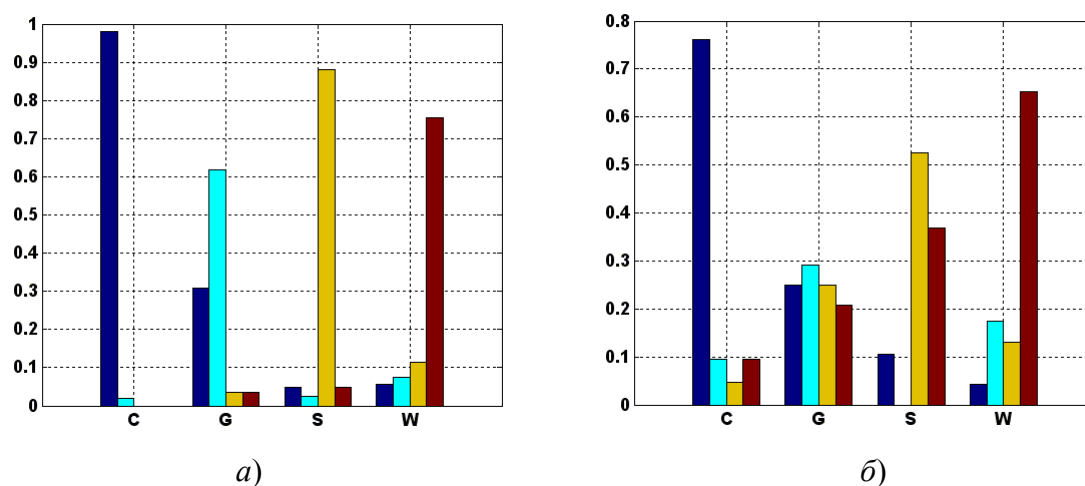


Рис. 10. Распределение значений выходов обученного нейросетевого классификатора при подаче на вход сигналов обучающей (а) и тестовой (б) выборок различных классов. Каждый рисунок содержит 4 гистограммы (по числу классов), каждая из которых содержит 4 столбика, соответствующие классам 'C', 'G', 'S', 'W'. Высота i -го столбца k -го класса представляет собой относительное число сигналов k -го класса, отнесённых классификатором к i -му классу (значения n_i/N_k).

Из табл. 2 видно, что наиболее точно распознаются сигналы класса 'C', что говорит о высокой специфичности сигналов ЭЭГ этого класса. Наименьшая точность распознавания наблюдается для сигналов класса 'G', при этом, как видно из рис. 10, ошибки распознавания сигналов класса 'G' на тестовой выборке практически равномерно распределились "в пользу" всех остальных классов, что говорит о невысокой специфичности сигналов ЭЭГ данного класса по отношению ко всем остальным классам. Ошибки распознавания сигналов класса 'S' в основном приходятся

"в пользу" класса 'W', что соответствует нейрофизиологическим представлениям о деятельности мозга при решении когнитивных задач этих типов.

Заключение

Результаты применения предложенного алгоритма для выделения характерных признаков многомерных нестационарных сигналов ЭЭГ и их классификации показали его работоспособность на имеющихся экспериментальных данных. Точность нейросетевого классификатора в построенном пространстве признаков составила $I^{обуч} = 0.81$ на обучающей выборке и $I^{тест} = 0.56$ на тестовой. Установлено, что наиболее специфичными являются сигналы ЭЭГ класса 'C' (соответствующие ментальной активности при решении задач на пространственное мышление), точность распознавания сигналов этого класса достигает 98% на обучающей выборке и 76% на тестовой. Наименее специфичными являются сигналы класса 'G' (соответствующие задачам на планиметрическое мышление), которые вносят наибольший вклад в результирующую ошибку работы классификатора.

Несмотря на достигнутый результат, предложенный алгоритм требует проведения ряда исследований, которые можно объединить в следующие группы:

- исследование различных пространств моментных характерных признаков;
- исследование процедуры кластеризации векторов МХП и формирования множества микросостояний;
- исследование процедуры формирования статических паттернов из микросостояний;
- исследование возможности и результатов применения динамических паттернов высоких порядков;
- исследование различных вариантов формирования вектора характерных признаков на основе индексов активности динамических паттернов;

- исследование процедуры обучения многослойного персептрона для классификации выделенных векторов характерных признаков.

Результатом этих исследований должны стать рекомендации по выбору числа микросостояний и числа статических паттернов, определение порядка вводимых в рассмотрение динамических паттернов, достаточного для обеспечения требуемой точности классификации, а также ряд других выводов по применимости и эффективности предложенного алгоритма формирования характерных признаков для классификации многомерных нестационарных временных рядов.

Работа выполнена при финансовой поддержке Совета по Грантам Президента РФ для поддержки молодых российских учёных (Грант МК-4245.2009.8), а также в рамках гранта РФФИ 10-08-00816 и ФЦП "Научные и научно-педагогические кадры инновационной России" на 2009-2013 годы.

Литература

1. M.B. Priestley. *Nonlinear and Nonstationary Time Series Analysis*. Academic Press, 1988.
2. *Nonlinear and Nonstationary Signal Processing*. Edited by W. J. Fitzgerald. University of Cambridge. 2001.
3. R. Dahlhaus. Fitting time series models to nonstationary processes // *The Annals of Statistics*. 1997. Vol. 25. № 1. Pp. 1-37.
4. G. McLachlan. *Discriminant Analysis and Statistical Pattern Recognition*. New York: John Wiley&Sons, 1992.
5. Осовский С. *Нейронные сети для обработки информации*. – М.: Финансы и статистика, 2004.
6. Wolpaw J.R., Birbaumer N., McFarland D.J., Pfurtscheller G., Vaughan T.M. Brain-computer interfaces for communication and control // *Clinical Neurophysiology*, 2002, V.113, P.767–791.

7. В.И. Скругин, А.О. Роик, Р.А. Наумов, Трофимов А.Г. Алгоритм классификации сигналов ЭЭГ на основе анализа в частотно-временной области // Сборник научных трудов XII Всероссийской научно-технической конференции "Нейроинформатика-2010". – М.: НИЯУ МИФИ, 2010. Т.1. С.266-276.
8. В.И. Скругин, А.О. Роик, Р.А. Наумов, Трофимов А.Г. Алгоритм выделения пространственно-частотных паттернов в структуре сигнала ЭЭГ // Материалы избранных научных трудов по теме: «Актуальные вопросы нейробиологии, нейроинформатики и когнитивных исследований». – М.: НИЯУ МИФИ, 2010. С.213-223.
9. С.А. Айвазян, Мхитарян В.С. Прикладная статистика и основы эконометрики: Учебник для вузов. – М.: ЮНИТИ, 1998.
10. Емельянов В. В., Курейчик В. В., Курейчик В. М. Теория и практика эволюционного моделирования. – М: Физматлит, 2003.
11. Зенков Л.Р. Клиническая электроэнцефалография (с элементами эпилептологии). Руководство для врачей. – М.: МЕДпресс-информ, 2004.
12. Витязев В.В. Вейвлет-анализ временных рядов: Учеб. пособие. – СПб.: Изд-во С.-Петербургского университета, 2001.