

МЕТОДОЛОГИЯ ЭКСПЕРТНО-КЛАССИФИКАЦИОННОГО АНАЛИЗА КВАЗИПЕРИОДИЧЕСКИХ СИГНАЛОВ В ЗАДАЧАХ ДИАГНОСТИКИ¹

А.А. Дорофеев, В.В. Гучук, А.А. Десова, Ю.А. Дорофеев

На базе экспертно-классификационных методов анализа данных разработана методология структурного анализа квазипериодических временных рядов для решения диагностических задач. Приведен пример ее применения для анализа пульсового сигнала лучевой артерии человека применительно к решению одной из сложных задач медицинской диагностики. Представлены результаты исследований, подтверждающие принципиальную возможность диагностики сложных заболеваний путем анализа квазипериодических биосигналов — пульсограмм.

Ключевые слова: экспертно-классификационные методы, квазипериодические сигналы, пульсовой сигнал лучевой артерии (пульсограмма), типология формы пульсограммы, типология спектральной плотности пульсограммы, медицинская диагностика.

ВВЕДЕНИЕ

Исходные данные для анализа многих технических, социально-экономических и медико-биологических объектов управления часто задаются в виде квазипериодических временных рядов. Примерами могут служить вибро-акустические данные о работе двигателя внутреннего сгорания, данные о сезонных колебаниях цен, биржевые котировки, данные о пассажиропотоках, изменяющихся со временем и др. Временные ряды характеристик солнечной активности и космического излучения широко используются для прогноза погоды. Немалый интерес и социальную значимость представляет исследование квазипериодических биосигналов (ЭКГ, ЭЭГ, миограммы, пульсограммы и пр.), в первую очередь биосигналов, связанных с вариабельностью сердечного ритма, для решения задач медицинской диагностики. Исследования свидетельствуют об их существенной диагностической значимости при различных патологических состояниях человека, в том числе связанных с заболеваниями сердечно-сосудистой системы, занимающих

первое место в мире по летальности. Создание эффективных методов анализа таких временных рядов дает возможность ранней диагностики сердечно-сосудистых заболеваний и, как следствие, позволяет спасти миллионы человеческих жизней. Кроме того, такие методы анализа широко применяются для оценки работоспособности человека-оператора в различных прикладных задачах.

В настоящей работе на базе экспертно-классификационных методов анализа данных [1, 2] разработана методология структурного анализа квазипериодических временных рядов, которая была реализована и практически применена для анализа пульсового сигнала. Этот сигнал измерялся на лучевой артерии пациента с помощью специально разработанного оптоэлектронного датчика пульса [3]. Аналоговый сигнал датчика оцифровывался и вводился в компьютер, где осуществлялись его визуализация и обработка с помощью специально разработанных алгоритмов, описанных далее.

Разработанная методология, алгоритмы и процедуры носят универсальный характер и могут применяться для исследования, идентификации, диагностики и совершенствования методов принятия решений в различных социально-экономических, организационно-административных, инженерно-технических и медико-биологических системах.

¹ Работа выполнена при частичной финансовой поддержке РФФИ (проекты 08-07-00349, 10-07-00027, 10-07-00210).

1. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ПУЛЬСОВОГО СИГНАЛА

Информация, извлекаемая из пульсового сигнала лучевой артерии, может быть представлена в двух различных видах. Первый вид — значения параметров формы пульсовой волны. Во многих исследованиях было показано наличие взаимосвязи между различными заболеваниями и формой пульсового сигнала, определяемой числом и степенью выраженности локальных пиков в пределах основного периода пульсовой волны. Второй вид — информация, отражающая динамические изменения формы пульсовой волны. Есть основания рассматривать человека как динамическую систему с определенной степенью сбалансированности внутренних биоритмов, как между собой, так и с ритмическими изменениями окружающей среды. Любое нарушение такого баланса является фактором риска возникновения определенных заболеваний. Таким образом, возникает возможность ранней диагностики заболеваний благодаря учету информации о дисбалансе биоритмов. Было отмечено, что степень выраженности их различных колебаний, их амплитудно-фазовые соотношения и степень синхронизации несут существенную информацию о состоянии организма и служат ранними признаками возникающих заболеваний [4]. Поэтому эффективные методы анализа пульсового сигнала должны предусматривать алгоритмы анализа и обработки обоих видов такой информации.

1.1. Базовые характеристики формы пульсовой волны

В общем виде форма волны пульсового сигнала представлена на рис. 1, где a — максимум систолической волны (a -зубец), b — поздняя систолическая волна, v — начало диастолической волны, z — максимум диастолической волны, d — максимум пресистолической (постдиастолической) вол-

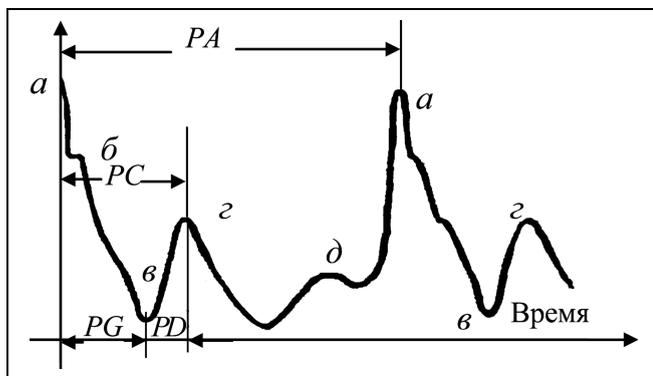


Рис. 1. Форма пульсовой волны лучевой артерии

ны, PA — основной период пульсовой волны, PC — время достижения максимума диастолической волны, PD — время подъема диастолической волны, PG — время падения катакроды.

Пульсовые сигналы разных людей могут существенно отличаться, как в наборе локальных волн, имеющих в пределах основного периода, так и в их выраженности (значения соответствующих максимумов, минимумов, времени их достижения и др.).

Форма пульсового сигнала зависит от таких факторов, как систолический выброс, интенсивность кровотока, состояние сосудистой стенки, вязкость крови и др. Это создает предпосылки для выявления различных патологий сердечно-сосудистой системы на основании оценки параметров формы пульсовой волны [5].

1.2. Параметры ритмической структуры пульсового сигнала

Квазипериодические биосигналы, к которым относится и пульсовой сигнал, имеют сложную ритмическую структуру, характеризующуюся совокупностью различных колебательных компонент, отличающихся по амплитуде, фазе и частоте. В литературе большое внимание уделяется изучению так называемых «медленных» колебательных процессов, которые характерны для многих функциональных систем организма человека. При изучении «медленных» волн выделяют выраженные компоненты кривой спектральной плотности динамического ряда (рис. 2):

- дыхательные волны HF , с частотой 0,15–0,4 Гц, определяемые дыхательным центром;

- медленные колебания (медленные волны 1-го порядка) LF (0,04–0,15 Гц), определяемые вазомоторной активностью, сосудодвигательным центром, барорефлекторной регуляцией;

- очень медленные колебания (медленные волны 2-го порядка) VLF (менее 0,04 Гц), происхождение которых пока четко не определено, считается, что они в основном связаны с обменными процессами и функционированием нейроэндокринных регуляторных механизмов [6].

Результаты исследований свидетельствуют о том, что параметры «медленных» биологических ритмов имеют выраженную диагностическую значимость как в задачах медицинской диагностики, так и для оценки состояния человека-оператора. В большинстве исследований диагностической значимости «медленных» биоритмов анализируется колебательная структура какого-либо одного параметра сигнала (чаще всего RR -интервала сердечного ритма). Однако каждый сигнал несет в себе информацию о многих физических процессах, протекающих в организме. При этом каждый функционально-значимый элемент единичного коле-

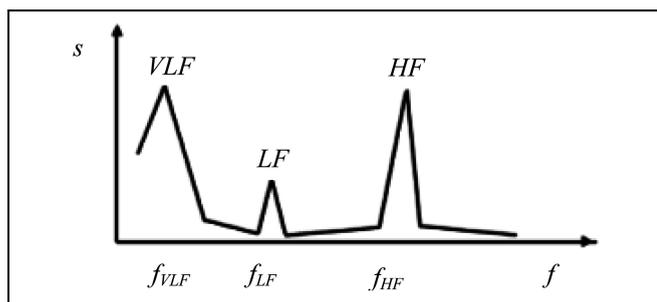


Рис. 2. Компоненты спектральной плотности динамического ряда

бания биосигнала обладает своей индивидуальной совокупностью колебательных компонент. Наши исследования убедительно свидетельствуют о том, что синхронное изучение ритмики отдельных элементов биосигнала и анализ их взаимосвязей позволяет получить существенно большую информацию о состоянии организма человека.

2. АЛГОРИТМЫ ВЫДЕЛЕНИЯ ХАРАКТЕРНЫХ ЭЛЕМЕНТОВ СИГНАЛА

На первом этапе анализа пульсограммы выделяется основной квазипериод пульсовой волны PA по всей длине записи сигнала. Сложность задачи автоматического выделения отдельных квазипериодов обусловлена значительной вариабельностью сигнала, большим разнообразием типов его формы, наличием в пределах основного квазипериода локальных экстремумов (временные: PA , PC и PG и амплитудные: амплитуда в точке a — VA , в точке z — VC и в точке v — VG) и артефактов. Для выделения в автоматическом режиме основной (PA) и дополнительных (PC , PG и др.) волн основного квазипериода были разработаны алгоритмы, базирующиеся на методах классификационного анализа данных [1].

Идею этих алгоритмов поясним на примере выделения основного квазипериода PA . На всей анализируемой записи сигнала выделяются все максимумы, т. е. формируется выборка значений амплитуд типа VA и VC . Затем строится автоматическая классификация этой выборки. В этих целях применяется комплексный алгоритм автоматической классификации [7], из которого далее применяются алгоритмы m -локальной оптимизации заданного критерия J и выбора информативных параметров. Было показано, что в одномерном случае алгоритм m -локальной оптимизации обеспечивает глобально-оптимальное значение критерия J при автоматической классификации на заданное число классов [8]. Подчеркнем, что границы классов определяются автоматически, обеспечивая минимизацию критерия средней по клас-

сам дисперсии классифицируемых точек с учетом функции распределения точек на оси амплитуд.

Для выбора числа классов применяется специальная экспертно-компьютерная процедура, входящая в состав комплексного алгоритма [7], посредством которой на первом этапе выбирается в формально определенном смысле оптимальное число классов r_{opt} , а на втором — эксперты (специалисты в области пульсографии) из содержательных соображений корректируют это число. При обработке реальных пульсограмм значение r_{opt} в подавляющем числе случаев находилось в диапазоне 3—7.

При классификации более сложных, многомерных объектов (например, при выделении типов форм пульсового сигнала) во время такой экспертизы анализируется также классификация каждого объекта. Перенесение объекта x_j из j -го класса в l -й считается допустимым, если он расположен вблизи границы между этими классами (в формально определенном смысле).

Типичная гистограмма распределения амплитуд основных квазипериодов PA носит многомодальный характер (у кривой несколько явно выраженных максимумов — «горбов»). Зона (класс) самого правого на оси значений максимума (большие значения амплитуды) заведомо будет соответствовать (в рамках модели этого сигнала) максимумам основного квазипериода для большей части анализируемого сигнала. Отрезки сигнала между двумя амплитудами, попавшими в эту зону (крайне правый класс), и являются претендентами на искомые квазипериоды.

На реальных сигналах часто наблюдаются существенные колебания значений амплитуд основной пульсовой волны. В связи с этим далее анализируется распределение выявленных квазипериодов на временной шкале. Если расстояние между некоторыми соседними a -зубцами (см. рис. 1) оказывается больше $T_c K_a$, то для этой неперiodизированной области выполняется коррекция. Здесь T_c — средняя длительность периода в анализируемом массиве, $K_a = 1,5...2,5$ — коэффициент аритмии.

Процесс коррекции представляет собой итеративную процедуру (алгоритм «амплитудное сито»), а именно: в неперiodизированных областях ищется максимальная амплитуда и ее включают в перечень основных волн (a -зубцов), затем вновь анализируется распределение амплитуд на временной шкале, находится максимальная амплитуда в оставшихся неперiodизированных областях и т. д. Остается проблема, связанная с неопределенностью коэффициента аритмии K_a . Впрочем, как показали тесты, влияние выбора его значения сказывается на качестве периодизации лишь в весьма экзотических случаях.

Предложенный алгоритм периодизации позволил выделить практически все периоды на исследуемой (более 350 сигналов) выборке пульсограмм. Исключения составили отдельные квазипериоды на нескольких пульсограммах, производящие весьма неоднозначное впечатление даже на специалистов по анализу пульсограмм. По аналогичной схеме находятся другие периодические составляющие сигнала в рамках основного квазипериода (в общем случае анализируются распределения не только максимумов, но и минимумов).

В процессе автоматического выделения основных и дополнительных волн оцениваются амплитудные и временные значения базовых параметров. Амплитудные значения вычисляются относительно условного нуля, установленного в измерительном устройстве, а временные параметры отсчитываются относительно момента, соответствующего максимальному значению амплитуды сигнала в пределах основного квазипериода (a -зубца).

Полученные значения базовых параметров основной волны пульсограммы затем используются для формирования динамических рядов, представляющих собой зависимость значений каждого базового параметра от номера периода на временной оси анализируемого сигнала. Эти динамические ряды подвергаются в дальнейшем статистическому и структурному анализу. Предварительно сформированные массивы проверяются на наличие выбросов (они могут возникнуть под воздействием артефактов), и в случае необходимости они корректируются. Контроль выбросов осуществляется путем автоматической оценки относительного отклонения значений каждого идентифицируемого параметра от значений предыдущего и последующего элементов. При отклонении, превышающем заданный порог, данное значение считается аномальным и подвергается коррекции в автоматическом или диалоговом режиме.

3. ЭКСПЕРТНЫЕ КЛАССИФИКАЦИИ ХАРАКТЕРИСТИК ПУЛЬСОГРАММ

3.1. Экспертная классификация формы пульсового сигнала

В процессе анализа пульсовых сигналов (более 350 векторов) была проведена экспертная классификация (типология) формы пульсовых сигналов, позволившая выделить 5 основных классов $\Phi 1$ — $\Phi 5$, представленных на рис. 3. Эти классы отличаются друг от друга количеством и степенью выраженности дополнительных волн в пределах основного квазипериода. Статистический анализ встречаемости различных классов формы для исследуемого массива данных показал, что наиболее распространены пульсовые сигналы классов $\Phi 2$,

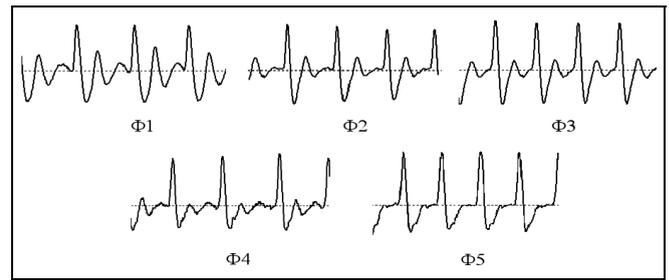


Рис. 3. Примеры пульсовых сигналов различных классов

$\Phi 3$ и $\Phi 4$, встречающиеся соответственно в 25, 37 и 20 % случаев от общего числа объектов, сигналы классов $\Phi 1$ и $\Phi 5$ встречаются достаточно редко, соответственно в 10 и 8 % случаев.

3.2. Экспертная классификация кривых спектральной плотности параметра PA

Динамические ряды, сформированные для базовых параметров единичных квазипериодов, в общем случае рассматриваются как случайные процессы, представляющие собой суперпозицию колебаний, трендов и шумов. В работе рассматривается только стационарная (точнее — квазистационарная) модель такого сигнала, поэтому применяются методы анализа стационарных случайных процессов. На первом этапе осуществляется оценка динамических рядов на стационарность и нормальность. Периодические составляющие динамических рядов и их характеристики выявляются с помощью спектрально-корреляционного анализа, прежде всего с помощью анализа функции спектральной плотности. Исследовались динамические ряды длительностью 200—300 циклов (периодов). Число коэффициентов разложения равнялось 30. Пики (выраженные максимумы) спектральной плотности, как правило, были сосредоточены в частотных диапазонах HF , LF и VLF (см. рис. 2). В редких случаях наблюдался локальный максимум кривой спектральной плотности в диапазоне частот между LF и VLF .

Качественный анализ функции спектральной плотности для динамического ряда параметра PA позволил экспертам выделить шесть основных типов (классов) спектральной плотности $C1$ — $C6$ (рис. 4), отличающихся количеством и степенью выраженности указанных выше колебательных компонент.

Для исследуемого массива данных встречаемость этих типов была следующей: $C1$ — 40, $C2$ — 18, $C3$ — 13,6, $C4$ — 6, $C5$ — 12, $C6$ — 10 %, т. е. более чем в 70 % случаев встречаются типы $C1$, $C2$ и $C3$, которые характеризуются сравнительно близкими значениями амплитуд компонент спектральной плотности LF и HF .

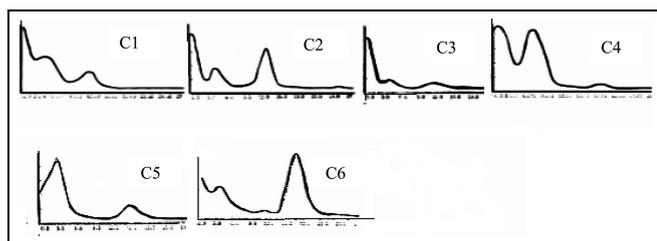


Рис. 4. Примеры различных типов спектров динамического ряда параметра PA

Остальные типы спектров встречаются существенно реже и характеризуются выраженным преобладанием той или иной компоненты. Так, тип С6 характеризуется относительно низким значением VLF и сильно выраженной компонентой HF , тип С5 характеризуется практически полным отсутствием колебательной компоненты LF , тип С4 — существенным преобладанием VLF и LF . Для этих трех типов спектра характерна малая вероятность наличия артериальной гипертензии (около 7%). Аналогичные спектральные характеристики были построены для динамических рядов всех базовых параметров пульсовой волны. Интересно, что для одного и того же исходного сигнала спектральные плотности различных динамических рядов часто сильно отличаются.

4. ФОРМАЛЬНЫЕ КЛАССИФИКАЦИИ ХАРАКТЕРИСТИК ПУЛЬСОГРАММ

4.1. Автоматическая классификация характеристик формы пульсового сигнала

Формализованная классификация типов формы пульсового сигнала проводилась в пространстве основных характеристик кривой спектральной плотности этого сигнала. При расчете спектральных характеристик вычислялись следующие параметры: частоты основных гармоник (f_1, f_2, \dots, f_n); амплитуды функции спектральной плотности для этих гармоник (B_1, B_2, \dots, B_n); а также попарные соотношения амплитуд гармоник вида B_i/B_j , $i, j = 1, \dots, n$. Оценка качества аппроксимации исследуемых сигналов гармоническими составляющими показала, что для исследуемого массива при точности аппроксимации не ниже 85% достаточно рассмотрения первых четырех гармоник, именно они и учитывались в дальнейших исследованиях.

При исследовании структуры спектра пульсового сигнала за основу принималась экспертная типология (классификация) формы сигнала, соответствующие классы ранее были обозначены как $\Phi 1, \Phi 2, \Phi 3, \Phi 4$ и $\Phi 5$. Вначале на базе исходного множества параметров ($B_1, B_2, B_3, B_4; B_i/B_j$,

$i, j = 1, \dots, 4$) формировался набор информативных параметров. В этих целях попарно сравнивались указанные классы в пространстве различных наборов спектральных параметров, выбранных из исходного множества. Для каждой пары классов определялся набор наиболее информативных параметров, в пространстве которых обеспечивалось наилучшее разделение этой пары. В табл. 1 приведены результаты расчетов для наиболее пересекающихся классов, где, как и ранее, n_i — число объектов в i -м классе.

Полученные данные позволили сформировать пространство информативных параметров для проведения автоматической классификации всей выборки данных. Из табл. 1 непосредственно следует, что из десяти изначально возможных параметров в качестве информативных можно использовать только следующие шесть из них: $B_2, B_3, B_2/B_1, B_3/B_1, B_3/B_2, B_4/B_2$.

В этом шестимерном пространстве и производилась автоматическая классификация всей выборки на 5 классов. Как и в случае выделения основного квазипериода, для этой цели применялся алгоритм m -локальной оптимизации [9] критерия

$$J_1 = \sum_{i=1}^r \frac{n_i}{n} K(A_i, A_i) \text{ при } m = 5. \text{ Необходимо отме-}$$

тить, что для $m > 5$ не происходило значимого изменения итоговой классификации, как по составу классов, так и по значениям критерия J_1 , при существенном возрастании числа итераций и времени счета. Для выбора начального разбиения использовалась соответствующая процедура комплексного алгоритма [7]. Полученная в итоге формализованная классификация сравнивалась с описанной выше экспертной классификацией. Объект считался ошибочно классифицированным, если в экспертной и формализованной классификациях он относился к разным классам (каждому классу A_i

Таблица 1

Наборы информативных параметров для некоторых пар классов

Пары классов	Число объектов в классе	Информативные параметры
$\Phi 2-\Phi 4$	$n_2 = 106, n_4 = 78$	$B_3, B_2/B_1, B_3/B_1$
$\Phi 3-\Phi 4$	$n_3 = 132, n_4 = 78$	$B_2, B_2/B_1, B_4/B_2$
$\Phi 2-\Phi 3$	$n_2 = 106, n_3 = 132$	$B_2, B_3/B_1, B_3/B_2$
$\Phi 4-\Phi 5$	$n_4 = 78, n_5 = 23$	$B_2, B_4/B_2$
$\Phi 1-\Phi 4$	$n_1 = 16, n_4 = 78$	$B_3, B_3/B_1, B_3/B_2$
$\Phi 1-\Phi 5$	$n_1 = 16, n_5 = 23$	$B_3, B_3/B_1, B_3/B_2$

формализованной классификации присваивался номер того класса экспертной классификации Φ_j , объектов из которого было больше всего в классе A_i). Погрешность (процент ошибочно классифицированных объектов) формализованной классификации относительно экспертной составила 12,7 %. Наилучшее соответствие было у классов A_4 и Φ_4 (погрешность около 4 %), а наихудшее — у классов A_2 и Φ_2 (погрешность около 15 %). Полученные результаты свидетельствуют о хорошем соответствии формализованной и экспертной классификаций — более 87 %.

4.2. Автоматическая классификация кривых спектральной плотности параметра РА

Ранее экспертный анализ типов спектральной плотности длительностей основного квазипериода (PA) исходного пульсового сигнала (далее для краткости — типов спектра) позволил выделить шесть основных типов ($C1$ — $C6$), отличающихся степенью выраженности (амплитудами) трех основных колебательных компонент HF , LF и VLF . Формальная классификация типов спектра проводилась в пространстве параметров, выбираемых из основных характеристик соответствующей спектральной кривой, а именно: значений амплитуд для этих гармоник и их взаимных соотношений — HF/LF , LF/VLF , VLF/HF . Вначале на базе этого множества параметров проводилось формирова-

ние набора информативных параметров. Для этой цели, как и в п. 4.1, попарно сравнивали классы экспертной классификации в пространстве различных наборов параметров, выбранных из исходного множества. Для каждой пары классов определялся набор наиболее информативных параметров, в пространстве которых обеспечивалось наилучшее разделение этой пары классов. Оказалось, что для всех пар классов наиболее информативны одни и те же два расчетных параметра — LF/VLF и HF/LF . В качестве примера на рис. 5 приведено сравнительное распределение объектов на плоскости этих параметров.

Анализ таких распределений на плоскости информативных параметров выявил области расположения объектов, относящихся к различным типам кривой спектральной плотности, что позволило оценить диапазоны значений информативных параметров для каждого из них (табл. 2).

Как и в п. 4.1, для формальной классификации применялся алгоритм m -локальной оптимизации

[7] критерия $J_1 = \sum_{i=1}^r \frac{n_i}{n} K(A_i, A_i)$ в двумерном про-

странстве параметров LF/VLF и HF/LF при $m = 5$. Число классов равнялось 6, оно выбиралось из соображений удобства сопоставления экспертной и формализованной классификаций. Для выбора начального разбиения применялась соответствующая процедура комплексного алгоритма [7].

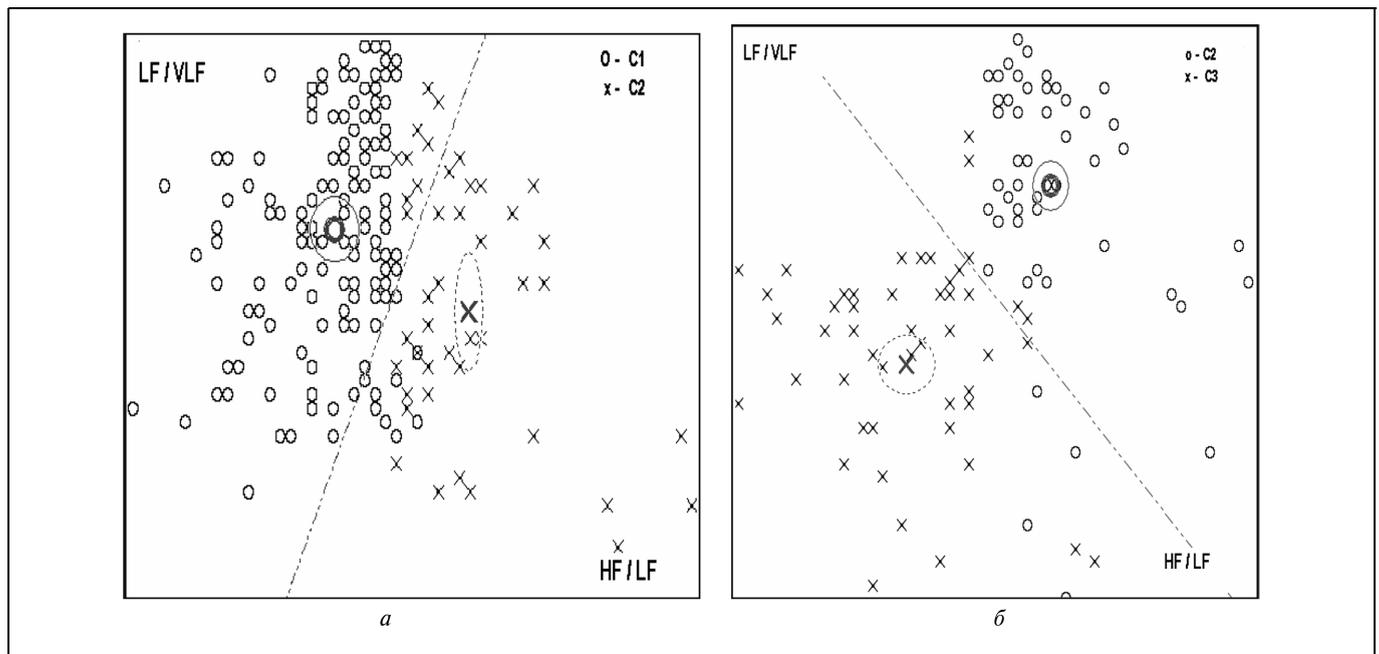


Рис. 5. Распределение объектов двух пар типов спектра на плоскости LF/VLF — HF/LF :
 a — $C2$ и $C2$; b — $C2$ и $C3$



Таблица 2

Диапазоны значений информативных параметров для различных типов функции спектральной плотности

Тип спектра	Число объектов	VLF/LF	LF/HF
C1	145	1,00—1,4	0,99—1,35
C2	64	1,00—3,0	0,72—0,99
C3	50	1,50—6,0	0,88—1,38
C4	26	0,85—2,5	1,06—3,60
C5	40	0,50—1,0	1,00—1,90
C6	36	0,90—1,1	0,77—0,99

Полученная в итоге формальная классификация сравнивалась с экспертной классификацией. Погрешность (процент ошибочно классифицированных объектов) итоговой формализованной классификации относительно экспертной составила 7,2 %. Наилучшее соответствие было у классов A2, C2 и A6, C6 (погрешность около 3 %), а наихудшее у A4, C4 (погрешность около 16 %). Полученные результаты свидетельствуют о хорошем соответствии формализованной и экспертной классификаций типа кривой спектральной плотности — около 93 %. Аналогичные формальные классификации были построены для спектральных кривых динамических рядов и других квазипериодов (*PC*, *PG* и др.).

5. МЕТОДЫ КЛАССИФИКАЦИОННОГО АНАЛИЗА ПАРАМЕТРОВ ПУЛЬСОВОГО СИГНАЛА В ЗАДАЧАХ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

5.1. Ступенчатая схема решения задач диагностики

Опираясь на типологию исследуемого сигнала (например, пульсового), полученную автоматическим, экспертным или смешанным способами, был разработан новый подход к решению задач диагностики (например, медицинской) на базе информации, получаемой при анализе исходного сигнала. Для решения такого класса задач было предложено использовать «ступенчатую схему распознавания» [9], основная идея которой состоит в следующем.

Пусть имеется обучающая выборка из n объектов. Каждый объект — это набор значений информативных параметров исследуемого сигнала для конкретного пациента, а также информация о его состоянии (в случае медицинской диагностики — это информация о том, страдает человек определенным заболеванием или нет). Ступенчатая схема помимо этой информации подразумевает также

наличие информации о принадлежности каждого объекта обучающей выборки к тому или иному классу некоторой другой (не диагностической) классификации. В качестве такой классификации может быть выбрана классификация типа формы исходного пульсового сигнала и (или) различные типологические классификации кривых спектральной плотности динамических рядов параметров пульсового сигнала. Если имеется только одна дополнительная классификация такого рода, то используется двухступенчатая схема диагностики; если же таких классификаций l , то используется l -ступенчатая схема.

Для примера рассмотрим случай двухступенчатой схемы, когда в качестве дополнительной классификации рассматривается классификация формы спектральной плотности длительностей основного квазипериода (*PA*) исходного пульсового сигнала.

Подчеркнем, что для эффективной диагностики необходимо, чтобы в обучающей выборке было достаточное (со статистической точки зрения) число объектов каждого класса заболеваний не только для всей выборки (как для классической, одноступенчатой схемы), но и для каждого типа (класса) формы пульсового сигнала в отдельности. Пусть получены значения информативных параметров пульсового сигнала для исследуемого пациента (объекта, подлежащего диагностике). Тогда на *первом этапе* определяется (распознается) принадлежность диагностируемого объекта к одному из ранее выделенных типов формы спектральной плотности исследуемого сигнала. В работе для этой цели применялся алгоритм распознавания образов с учителем метода потенциальных функций [10]. Пусть для определенности он был отнесен к первому типу формы спектральной плотности — C1. На *втором этапе* производится распознавание наличия у объекта диагностируемого заболевания, т. е. распознается принадлежность объекта к классу «больных» (в отличие от другого класса «здоровых», а точнее, не больных этим заболеванием). В построении диагностического (решающего) правила участвуют только те объекты обучающей выборки, тип формы пульсового сигнала которых совпадает с типом формы спектральной плотности диагностируемого объекта (в нашем примере — C1). Для этой цели также используется алгоритм распознавания образов по методу потенциальных функций [10].

Следует обратить внимание, что на каждом из этих этапов (определение типа формы спектральной плотности сигнала и диагностика) могут использоваться, вообще говоря, различные наборы информативных параметров. Эти наборы формируются в процессе обучения независимо для

первого и второго этапов, а именно: на первом этапе выбирается такой набор параметров, который обеспечивает в пространстве этих параметров хорошую делимость объектов, имеющих различные типы формы спектральной плотности пульсового сигнала. А на втором этапе информативным будет считаться такой набор параметров, который обеспечит наиболее высокую эффективность диагностики (малый процент ошибок на обучающей выборке) при достаточно простом решающем правиле. В процессе такого выбора широко применяется алгоритм выбора информативных параметров, входящий в комплексный алгоритм [7].

5.2. Дифференциальная диагностика ранней стадии артериальной гипертензии у подростков

Эта задача решалась на базе обширного экспериментального материала, полученного в ходе клинических обследований по выявлению ранней стадии артериальной гипертензии в детском и подростковом возрасте. Обследования проводились в клинике функциональной патологии Научного центра здоровья детей РАМН. Использовалась выборка из 360 пациентов в возрасте от 9 до 16 лет. Все пациенты по основному диагнозу были разделены на два класса: *класс 1* — первичная артериальная гипертензия (160 чел.), *класс 2* — отсутствие артериальной гипертензии, но наличие различных видов психосоматической функциональной патологии, в том числе ночной энурез, длительный субфебрилитет, цефалгия, ожирение и ряд других заболеваний при нормальном артериальном давлении (200 чел.). Подчеркнем, что во второй класс входили не здоровые люди, а пациенты с другими, нередко достаточно серьезными заболеваниями, что существенно осложняет задачу дифференциальной диагностики артериальной гипертензии, в особенности ее ранних стадий. Клинические диагнозы устанавливались лечащи-

ми врачами по принятым методикам — исследовалась динамика артериального давления, делались клинические анализы и выполнялся ряд других исследований.

Вначале было сформировано пространство информативных признаков, которое использовалось на первом этапе ступенчатой схемы для идентификации формы пульсового сигнала пациентов выборки. В качестве показателей формы пульсового сигнала использовались амплитуды компонент *VLF*, *LF* и *HF* кривой спектральной плотности (см. рис. 2) для динамических рядов параметров *PA*, *PC*, *PD* и *PG* (см. рис. 1); взаимные соотношения компонент для этих динамических рядов и взаимное соотношение однотипных компонент для динамических рядов разных параметров. Исследовались агрегированные характеристики формы спектральной плотности для динамического ряда длительностей единичных колебаний пульсового сигнала по параметрам *VLF*, *LF* и *HF*.

Для оценки эффективности ступенчатой схемы анализировалось распределение объектов по диагнозу (классы 1 и 2) в пределах каждого типа формы спектральной плотности (табл. 3).

Полученные результаты показывают, что, используя лишь агрегированные данные о некоторых типах спектральной плотности, можно делать вероятностные оценки наличия или отсутствия артериальной гипертензии, а именно — при спектре типа *C2* можно с вероятностью около 0,8, а при спектре типов *C3* и *C6* — с вероятностью около 0,7 говорить об отсутствии артериальной гипертензии. Правда, у остальных типов спектров нет столь выраженной зависимости от наличия или отсутствия артериальной гипертензии, здесь требуется привлечение, в рамках второго этапа ступенчатой схемы, алгоритмов распознавания образов для построения соответствующих диагностических правил.

Для выбора информативных параметров для второго этапа ступенчатой схемы проводилась оценка информативности различных показателей при решении задачи дифференциальной диагностики артериальной гипертензии отдельно для каждого типа формы спектральной плотности. Значение информативности *I* определялось как мера «вклада» каждого показателя в решающее правило для конкретной диагностической классификации массива объектов в пределах каждого типа формы спектральной плотности.

Для всех шести типов формы спектральной плотности наиболее информативными оказались одни и те же пять показателей, определяемых ритмической структурой пульсового сигнала (с несколькими отличающимися значениями оценок информативности, но с одной и той же их ранжиров-

Таблица 3

Распределение объектов по диагнозу в пределах каждого типа формы

Тип формы спектральной плотности	Число объектов	Доля объектов класса 1 (гипертензия)	Доля объектов класса 2 (нет гипертензии)
1	145	0,45	0,55
2	64	0,17	0,83
3	50	0,31	0,69
4	26	0,50	0,50
5	40	0,44	0,56
6	36	0,30	0,70



кой). Далее приведены полученные показатели, а также средневзвешенные оценки их информативности (оценки для каждого типа формы спектральной плотности усреднялись с весом, равным доле объектов этого типа от объектов всей выборки): $HF(PC)/LF(PC)$, $I_1 = 0,54$; $LF(PC)$, $I_2 = 0,41$; $VLF(PG)$, $I_3 = 0,25$; $HF(PC)/HF(PA)$, $I_4 = 0,22$; $HF(PC)/HF(PG)$, $I_5 = 0,18$. Таким образом, для данной диагностической задачи наиболее информативно отношение спектральных мощностей дыхательной волны и медленной волны первого порядка для параметра PC (длительность дикротической волны). Для сравнения была оценена информативность аналогичного показателя для параметра PA (длительность основного квазипериода), наиболее часто используемого в исследованиях ритмической структуры электрокардиограмм. Ее значение составило лишь 0,01. Это подтверждает важность исследования ритмической структуры биосигналов, обусловленной совокупностью колебательных компонент различных функционально-значимых элементов в пределах основного квазипериода.

На основе полученных наборов информативных параметров для обоих этапов была реализована двухступенчатая схема построения решающих правил для задачи дифференциальной диагностики артериальной гипертензии у детей и подростков. Как уже говорилось, на обоих этапах применялся алгоритм распознавания образов по методу потенциальных функций [10]. Каждое решающее правило определяется значениями коэффициентов при каждом из информативных признаков и определяет в пространстве этих признаков поверхность, разделяющую объекты двух распознаваемых классов.

Оценка эффективности полученных решающих правил на экзаменационной выборке показала достаточно хорошие результаты — с достоверностью 0,87 правильно распознаются объекты первого класса (наличие артериальной гипертензии) и с вероятностью 0,92 — второго класса (ее отсутствие).

ЗАКЛЮЧЕНИЕ

Перспективность использования пульсограмм для диагностики и оценки состояния человека, в том числе многих заболеваний, не связанных с сердечно-сосудистой системой, в настоящее время не вызывает сомнений. Их диагностическая информативность позволит решить проблему оперативной диагностики ряда болезней, выявление которых пока требует весьма длительного времени. Несмотря на многовековую историю пульсовой диагностики, практика ее применения достаточно

ограничена, что связано с неформализованностью процесса такой диагностики. Наибольшие трудности здесь возникают при создании методов выделения из квазипериодических биосигналов информативных параметров и эффективных алгоритмов построения решающих правил (диагностики).

Представленные результаты исследований подтвердили принципиальную возможность диагностики сложных заболеваний на базе анализа пульсограмм (на примере ранней стадии артериальной гипертензии в детском и подростковом возрасте).

ЛИТЕРАТУРА

1. Бауман Е.В., Дорофеюк А.А. Классификационный анализ данных // Тр. междунар. конф. по проблемам управления. — М.: СИНТЕГ, 1999. — Т. 1. — С. 62–67.
2. Дорофеюк А.А., Покровская И.В., Чернявский А.Л. Экспертные методы анализа и совершенствования систем управления // Автоматика и телемеханика. — 2004. — № 10. — С. 172–188.
3. Десова А.А., Дорофеюк А.А., Максимов Д.Ю. Портативная компьютерная система регистрации, обработки и хранения пульсовых сигналов // Датчики и системы. — 2008. — № 4. — С. 29–34.
4. Теоретические и прикладные аспекты анализа временной организации биосистем. — М.: Наука, 1976. — 350 с.
5. Логвинов В.С. Метод диагностики по параметрам колебательных и волновых процессов в сердечнососудистой системе // Пульсовая диагностика тибетской медицины: Сб. науч. тр. — Новосибирск: Наука, 1988. — С. 90–93.
6. Емельянов И.П. Структура биологических ритмов человека в процессе адаптации // Сб. науч. тр. — Новосибирск: Наука, 1986. — С. 52–61.
7. Дорофеюк Ю.А. Комплексный алгоритм автоматической классификации и его использование в задачах анализа и принятия решений // Таврический вестник информатики и математики / Междунар. периодическое издание КНЦ НАН Украины. — 2008. — № 1. — С. 171–177.
8. Глобально-оптимальные алгоритмы классификационного анализа сложноорганизованных данных / Е.В. Бауман, А.А. Дорофеюк, Ю.А. Дорофеюк, Н.Е. Киселева / Сб. тр. IV междунар. конф. по проблемам управления (МКПУ-IV) / ИПУ РАН. — М., 2009. — С. 344–349.
9. Дорофеюк А.А. Алгоритмы автоматической классификации // Автоматика и телемеханика. — 1971. — № 12. — С. 78–113.
10. Браверман Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. — М.: Наука, 1983. — 464 с.

Статья представлена к публикации членом редколлегии А.С. Манделем.

Дорофеюк Александр Александрович — д-р техн. наук, профессор, зав. лабораторией,
☎(495) 334-75-40, ✉adorof@ipu.ru,

Десова Аэлита Арсеньевна — ст. научный сотрудник, канд. техн. наук, ☎(495) 334-56-19, ✉adesova@mail.ru,

Гучук Владимир Всеволодович — ст. науч. сотрудник, канд. техн. наук, ☎(495) 334-89-51, ✉polma@bk.ru,

Дорофеюк Юлия Александровна — науч. сотрудник, ☎(495) 334-75-40, ✉tigr886@bk.ru,

Институт проблем управления им. В.А.Трапезникова РАН, г. Москва.