

УДК 616.12- 073.7

Н.Г. Иванушкина, канд. техн. наук, Е.О. Иванько

Технологии распознавания образов поздних потенциалов предсердий: формирование признаков

Рассмотрены принципы распознавания образов поздних потенциалов предсердий (ППП) в системах электрокардиографии высокого разрешения (ЭКГ ВР). Выполнено формирование признаков образов ППП на основе алгоритма многоуровневого вейвлет-анализа электрокардиосигналов. Проведены численные эксперименты с моделями ППП и реальными сигналами ЭКГ. С помощью обработки детализирующих коэффициентов вейвлет-разложения выбраны диагностические признаки образов, обеспечивающие максимальное расстояние между классами.

The principles of atrial late potentials (ALP) patterns recognition are considered in the systems of high resolution electrocardiography (HR ECG). The features of ALP are formed on the basis of ECG wavelet-analysis algorithm. The numerical experiments with the ALP models and real ECG signals are conducted. Providing the maximum distance between the classes the diagnostic features of ALP features are chosen using detail coefficients of wavelet-decomposition.

Введение

В основе предсердных тахикардий наиболее часто лежит механизм "риентри" ("re-entry") [1-3], представляющий повторный вход волны возбуждения в миокарде. Предвестниками развития данного механизма являются низкоамплитудные поздние потенциалы предсердий (ППП), которые отражают наличие замедленной фрагментированной предсердной деполяризации.

Согласно литературе [1-3] возможна дифференциация механизма повторного входа волны возбуждения на основе различных анатомических и/или функциональных причин в миокарде ("периодическое" и "аперриодическое риентри"), а также размеров замкнутых областей ("макро-" и "микрориентри"). Соответственно, ППП имеют широкий диапазон изменений амплитуды, формы, длительности и времени появления, что вызывает сложности при их обнаружении.

Электрокардиография высокого разрешения является эффективным и перспективным методом повышения качества диагностики и решения проблемы выявления низкоамплитудных компонент ЭКГ. Однако, результаты обнаруже-

ния ППП, полученные при стандартных методиках временного и спектрального анализа электрокардиосигналов, имеют невысокую предсказывающую ценность [1].

Поэтому актуальной проблемой при совершенствовании систем ЭКГ ВР является создание методик выбора дополнительных диагностических признаков ППП на основе моделирования тонких проявлений электрической активности сердца, исследование возможностей новых координатных базисов, а также разработка алгоритмов распознавания образов ППП, позволяющих повысить достоверность диагностики.

Принципы построения алгоритмов распознавания образов ППП

В существующих системах ЭКГ ВР для оценки состояния электрической активности сердца и раннего выявления тахикардий предсердий широко применяются признаки образов ППП в виде параметров P зубца во временной области, а также в виде коэффициентов разложения в частотном координатном базисе [1]. Однако, в этом случае для учёта многообразия проявлений ППП необходимы описания образов ЭКГ большой размерности.

Кроме того, при диагностике в системах ЭКГ ВР принимается, что образы ППП представляют собой детерминированные функции со случайными флуктуациями. Вследствие этого, вектор признаков случайным образом будет смещаться в пространстве при различных измерениях одного и того же образа, что может приводить к перекрыванию границ классов.

Процесс распознавания образов электрокардиосигналов с ППП в системах ЭКГ ВР можно представить в виде двух обобщённых процедур:

- формирование диагностических признаков образов;
- построение классификатора образов.

Формирование признаков можно формализовать как преобразование T исходного сигнала $S(t)$ к вектору признаков Y . В общем случае, в качестве оператора T могут быть применены ортогональные преобразования (Фурье, Адамара, Хаара, Карунена-Лоева и т.д.).

$$Y = T \{S(t)\} \quad (1)$$

После формирования признаков необходимо выполнить переход от пространства Y^m к пространству X^n , $n \leq m$ с меньшей размерностью.

В процессе формирования признаков образцов необходимо учитывать, что при классификации важны те признаки, которые наиболее подобны внутри класса. При кластерном анализе, который является видом классификации без учителя, в качестве меры сходства (подобия) рассматривается расстояние, позволяющее отнести образ к области, характеризуемой некоторым центром кластера. Наименьшее расстояние между признаками внутри класса и, в то же время, наибольшее расстояние между признаками различных классов — принцип, лежащий в основе классификации образов по критерию минимума расстояния [4-5].

Эвклидово расстояние между образами x и z определяется как

$$D = \|x - z\| = \sqrt{(x - z)'(x - z)}. \quad (2)$$

Для придания больших весов более отдаленным друг от друга объектам используется квадрат евклидова расстояния:

$$D^2 = \|x - z\|^2 = (x - z)'(x - z). \quad (3)$$

Кроме эвклидова расстояния в задачах классификации применяют расстояние Махаланобиса, расстояние Хемминга ("сити-блок" расстояние), расстояние Чебышева, а также другие меры сходства, например, косинус угла, образованного векторами x и z .

Оценивание качества классификации может быть выполнено по критерию максимума отношения расстояний:

$$K = \frac{S_e}{S_{ik}}, \quad (4)$$

где S_e — расстояние между кластерами, S_{ik} — расстояние внутри k -го кластера.

В данной работе в качестве меры сходства при кластеризации признаков ППП использовался квадрат эвклидова расстояния. Параметр S_e определялся как расстояние между центрами кластеров, а S_i — как среднее расстояние от точки, принадлежащей рассматриваемому кластеру, до центра этого кластера:

$$S_e = \|C_k - C_p\|^2, \\ S_{ik} = \frac{1}{N} \sum_{j=1}^N \|x_j - C_k\|^2,$$

где C_k и C_p — координаты центров k -го и p -го кластеров, N — количество точек, отнесенных к k -му классу, x_j — координата j -й точки в пространстве признаков.

При выявлении кластеров признаков ППП в системе MATLAB использовался алгоритм K -средних, минимизирующий показатель качества, определенный как сумма квадратов расстояний всех точек, входящих в кластерную область, до центра кластера.

Формирование признаков образов ППП на основе многоуровневого вейвлет-преобразования

Для получения дополнительных диагностических параметров ППП в системе ЭКГ ВР предлагается формирование признаков образов на основе многоуровневого вейвлет-преобразования.

Применение дискретного вейвлет-разложения (ДВР) при обработке данных ЭКГ позволяет перейти от информации в исходном амплитудно-временном базисе к информации о вейвлет-коэффициентах аппроксимации и детализации и вейвлет-компонентах сигнала в новом базисе. Следовательно, алгоритм процедуры формирования признаков образов ППП на основе многоуровневого вейвлет-преобразования включает в себя этап определения аппроксимирующих и детализирующих коэффициентов, этап обработки коэффициентов для получения вектора признаков с меньшей размерностью, этап селекции и экстракции признаков образов согласно критерию минимальной дисперсии признаков внутри класса и максимальной дисперсии (расстояния) между классами.

При экспериментальном исследовании возможности решения задачи автоматического распознавания образов ППП на основе кластерного анализа с последующей классификацией кардиосигналов была сформирована выборка из 370 Р зубцов ЭКГ. При этом в качестве нормы рассматривались предварительно очищенные от шума реальные записи ЭКГ трёх пациентов с различающейся формой Р зубцов без особенностей: 55 кардиоциклов первого пациента, 106 — второго, 24 — третьего (рис.1). При начальном предположении о наличии двух классов типа "норма" и "патология" кардиоциклы с такими Р зубцами должны быть априорно отнесены к классу 1: "отсутствие ППП". Далее в терминальную часть каждого Р зубца были добавлены смоделированные ППП. Для каждого пациента ППП отличались по амплитуде, форме и количеству. В качестве моделей ППП использовались потенциалы действия, полученные в результате решения уравнений Ходжкина-Хаксли, а также выбранные отрезки модулированной синусоиды [6-7]. Предполагается, что кардиоциклы с такими Р зубцами должны быть отнесены к классу 2: "наличие ППП".

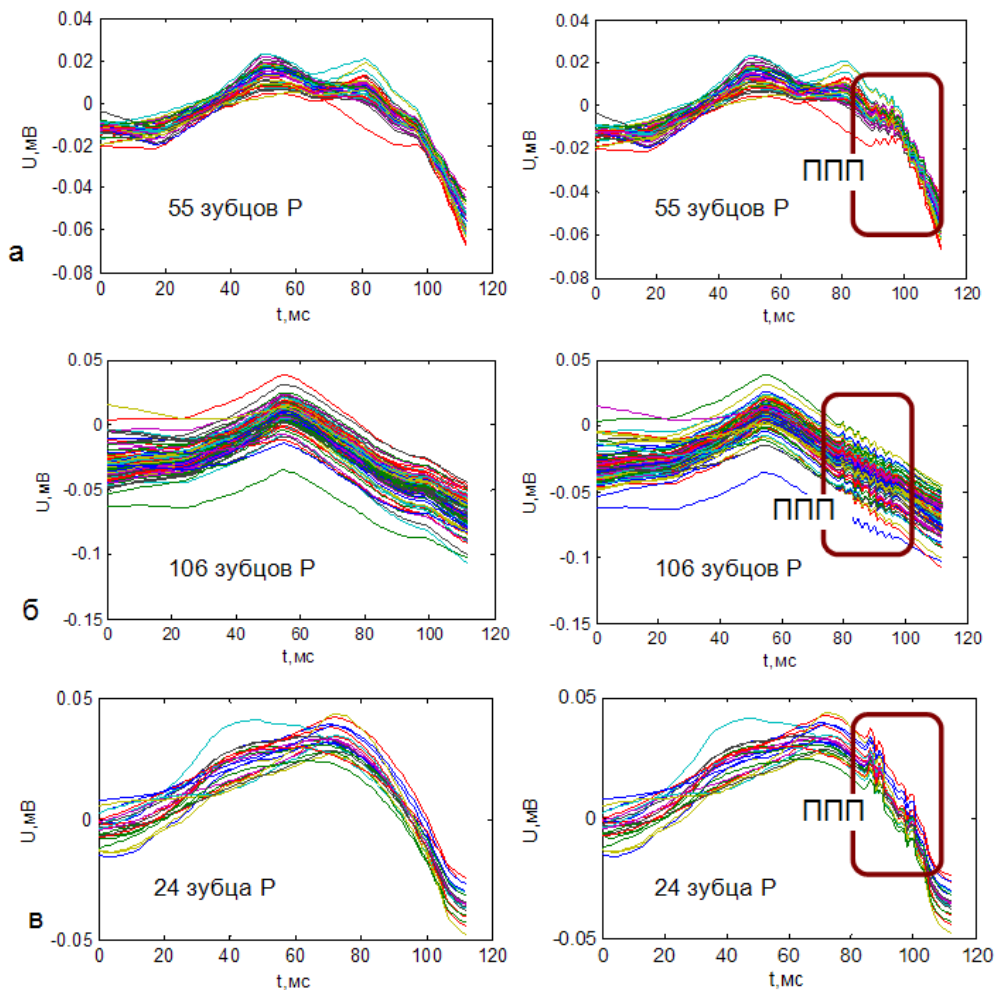


Рис. 1. Формирование выборки данных для классификации:
а) пациент 1; б) пациент 2; в) пациент 3

При помощи ДВР все Р зубцы исследуемой выборки кардиоциклов были разложены до 5-го уровня [8-9]. В качестве вейвлета использовалась вейвлет-функция симлет 4-го порядка. При этом анализируемый сигнал $S_{ЭКГ}$ представляет собой сумму аппроксимирующей составляющей (а5) и детализирующих составляющих (d5, d4, d3, d2, d1):

$$S_{ЭКГ} = a5 + d5 + d4 + d3 + d2 + d1,$$

где d1 соответствует первому уровню детализации.

Из-за высокочастотной низкоамплитудной природы ППП признаки их наличия или отсутствия следует искать на детализирующих уровнях при вейвлет-разложении Р зубца (рис.2).

Экспериментальные исследования различных выборок Р зубцов с наличием и отсутствием ППП показали, что анализ вейвлет-коэффициентов и/или вейвлет-компонентов детализации уровня d1, реже - d2 или совместно d1 и d2, позволяет получить наибольший процент правильных классификаций (97-100% в модельном эксперименте). Вейвлет-коэффициенты остальных

уровней детализации (d3, d4, d5), а тем более уровня аппроксимации (а5), не содержат достаточных для классификации признаков ППП, так как передают характер более высокоамплитудных низкочастотных составляющих кардиосигнала. Следует отметить, что данная особенность вейвлет-преобразования позволяет анализировать информацию о ППП независимо от формы исследуемых Р зубцов, т.е. аппроксимирующей составляющей. Это даёт возможность сравнивать и классифицировать кардиоциклы разных пациентов в одной выборке данных.

Анализ энергетического спектра детализирующих компонент d1-d4 Р зубца без ППП и с наличием ППП подтверждает гипотезу о необходимости анализа именно вейвлет-коэффициентов уровня детализации d1, отражающего наличие высокочастотных составляющих наименьшей амплитуды, которыми и являются ППП. На рис.3 видно, что только на уровне d1 значения энергетического спектра в случае наличия ППП максимально превышают значения энергетического спектра в случае отсутствия ППП.

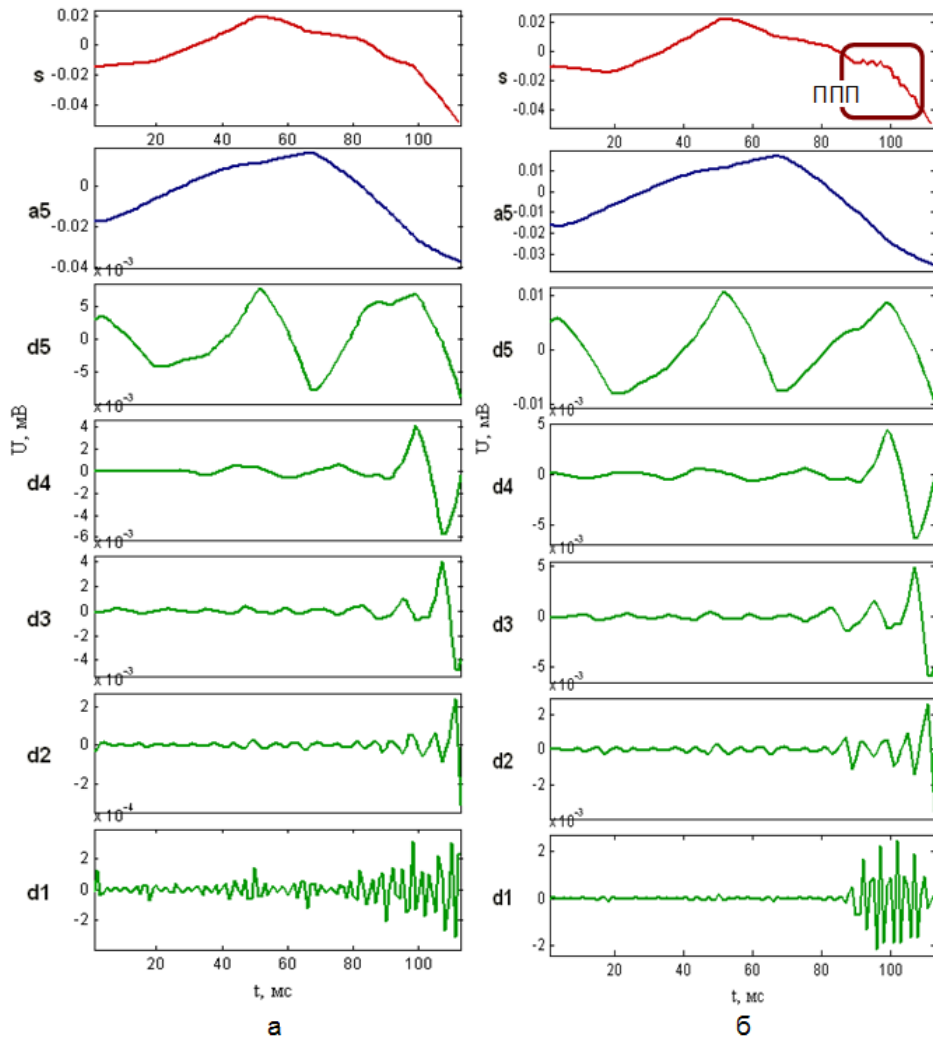


Рис. 2. Дискретное вейвлет-разложение до 5-го уровня: а) Р зубец без ППП; а) Р зубец с наличием ППП

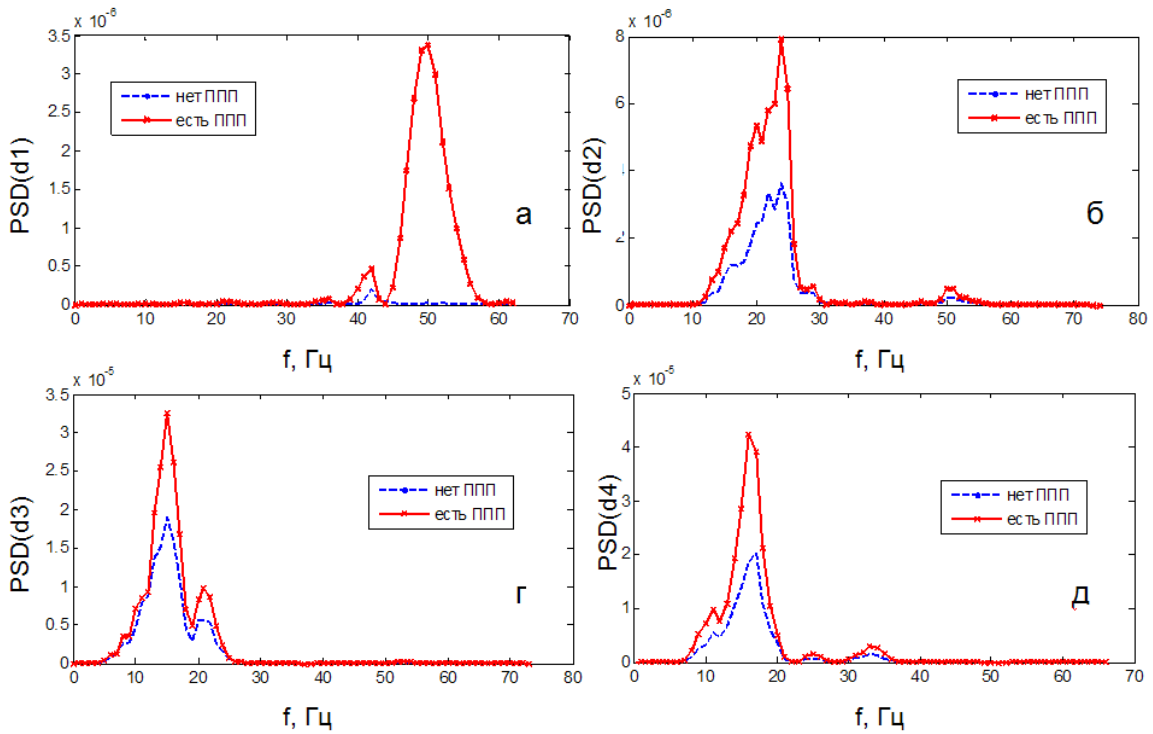


Рис. 3. Энергетический спектр детализирующих вейвлет-компонент Р зубца без ППП и с наличием ППП: а) уровень d1; б) уровень d2; в) уровень d3; г) уровень d4

Выбор признаков ППП и понижение размерности вектора признаков

При выборе d1 в качестве уровня детализации, содержащего информацию о признаках ППП, для рассматриваемой выборки Р зубцов размерность данных для классификации составляет $M \times N$, где $M=370$ — количество Р зубцов в выборке, $N=60$ — количество вейвлет-коэффициентов на уровне d1. Такой набор данных является избыточным и требует дальнейшей обработки вейвлет-коэффициентов для получения вектора признаков с меньшей размерностью.

При формировании вектора признаков вейвлет-коэффициенты уровня детализации d1 рассматривались как дискретный сигнал

$Sd_1(k) = Sd_{1k}$, где $k \in N$. В качестве числовых характеристик определялись: L - энергия абсолютных значений Sd_{1k} (L^2 -норма), abs_mean - среднее абсолютное значение Sd_{1k} , max_abs - максимум отклонения абсолютных значений Sd_{1k} , max_fft - максимальное значение энергетического спектра Sd_{1k} . Каждый из этих параметров существенно отличается для случаев “отсутствие ППП” и “наличие ППП” (рис.3). На рис.4 и рис.5 обозначения “P1”, “P2”, “P3” соответствуют группам Р зубцов без ППП для первого, второго и третьего пациентов, а “P1+ППП”, “P2+ППП”, “P3+ППП” — группам Р зубцов с наличием ППП.

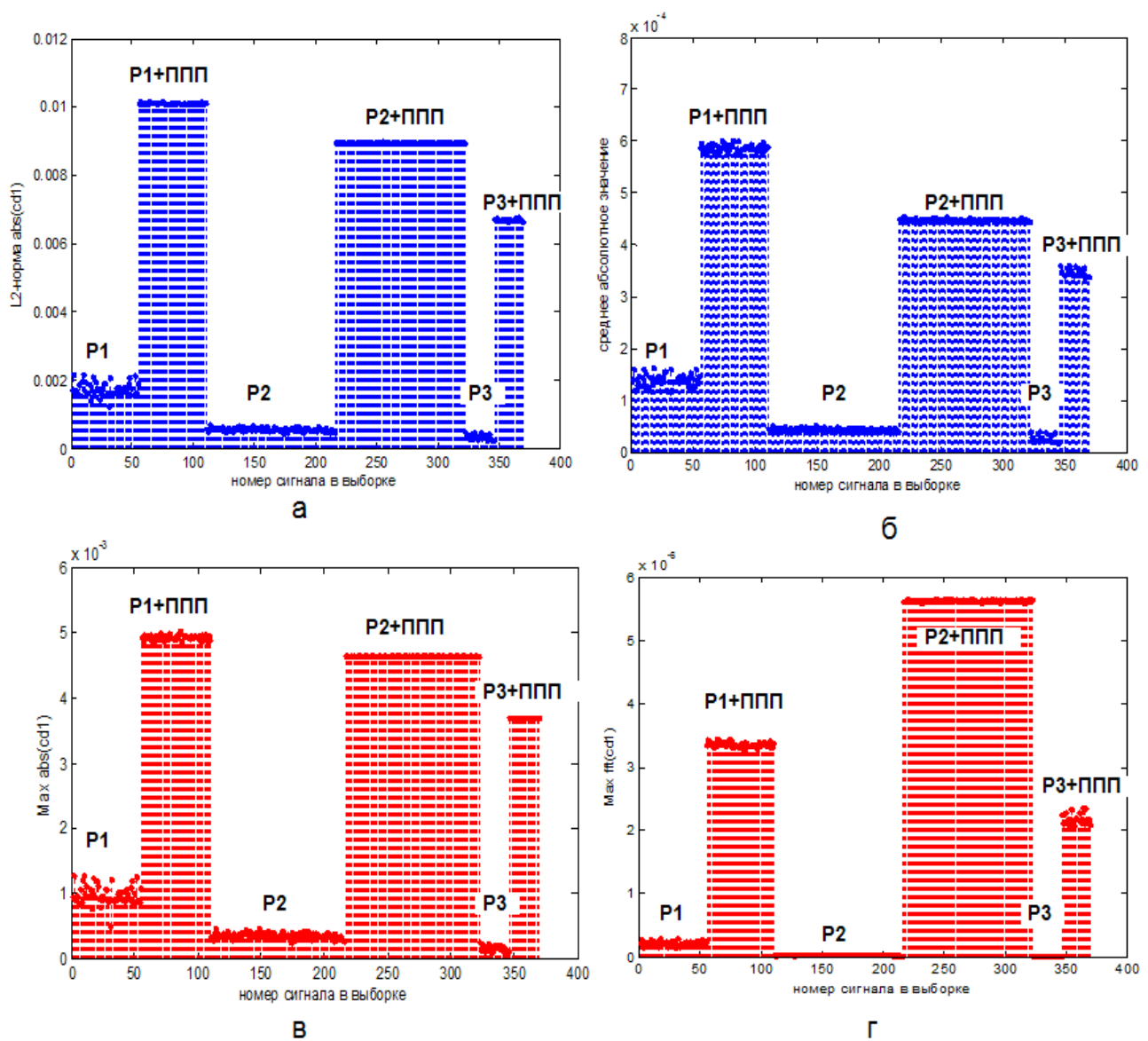


Рис. 4. Числовые характеристики вейвлет-коэффициентов уровня детализации d1: а) L2-норма; б) среднее абсолютное значение; в) максимум отклонения абсолютных значений; г) максимальное значение энергетического спектра

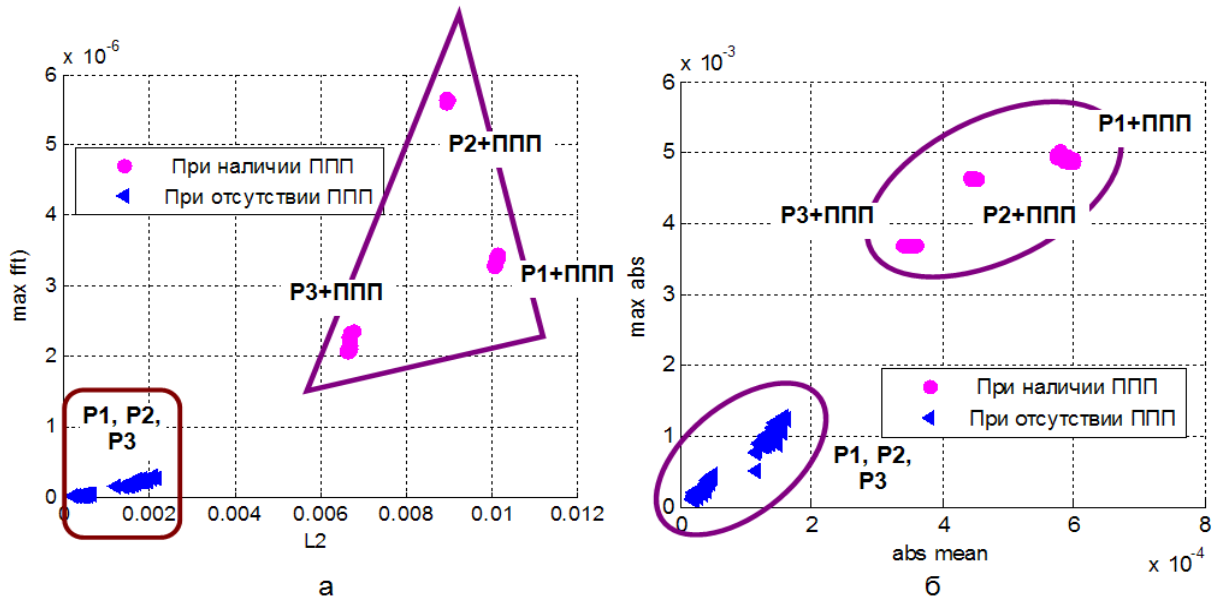


Рис. 5. Результаты кластеризации числовых характеристик вейвлет-коэффициентов в случае отсутствия и наличия ППП: а) L^2 -норма и максимальное значение энергетического спектра; б) среднее абсолютное значение и максимум отклонения абсолютных значений

Таблица 1. Оценка качества классификации по критерию максимума отношения расстояний (4)*

Параметр	L	max_fft	abs_mean	max_abs
среднее расстояние от точки до центра кластера C_1 внутри кластера 1 (отсутствие ППП), S_{i1}	$3.3295 \cdot 10^{-7}$	$5.1464 \cdot 10^{-13}$	$2.1046 \cdot 10^{-9}$	$1.0324 \cdot 10^{-7}$
среднее расстояние от точки до центра кластера C_2 внутри кластера 2 (наличие ППП), S_{i2}	$1.0540 \cdot 10^{-6}$	$1.0083 \cdot 10^{-12}$	$6.2897 \cdot 10^{-9}$	$1.4162 \cdot 10^{-7}$
расстояние между центрами кластеров C_1 и C_2 , S_e	$6.6057 \cdot 10^{-5}$	$2.0546 \cdot 10^{-11}$	$1.6540 \cdot 10^{-7}$	$1.6710 \cdot 10^{-5}$
S_e / S_{i1}	198.3992	39.9231	78.5898	161.8559
S_e / S_{i2}	62.6727	20.3769	26.2970	117.9918

*L — L^2 -норма, max_fft — максимальное значение энергетического спектра, abs_mean — среднее абсолютное значение, max_abs — максимум отклонения абсолютных значений. Все параметры определены для вектора вейвлет-коэффициентов d1 уровня детализации.

Применение любой из предложенных числовых характеристик вейвлет-коэффициентов d1 уровня детализации в качестве признака ППП, а также их комбинаций в виде вектора признаков $[X_1 \ X_2]$, позволило правильно разделить кардиоциклы исходной выборки на 2 класса: “класс 1 — отсутствие ППП” и “класс 2 — наличие ППП” (рис.6). Однако, анализ рис.4, 5 и

данных табл.1 показывает, что такие параметры, как максимум отклонения абсолютных значений и L^2 -норма вектора вейвлет-коэффициентов, обеспечивают лучшее разделение классов в пространстве признаков, чем максимальное значение энергетического спектра и среднее абсолютное значение вектора вейвлет-коэффициентов.

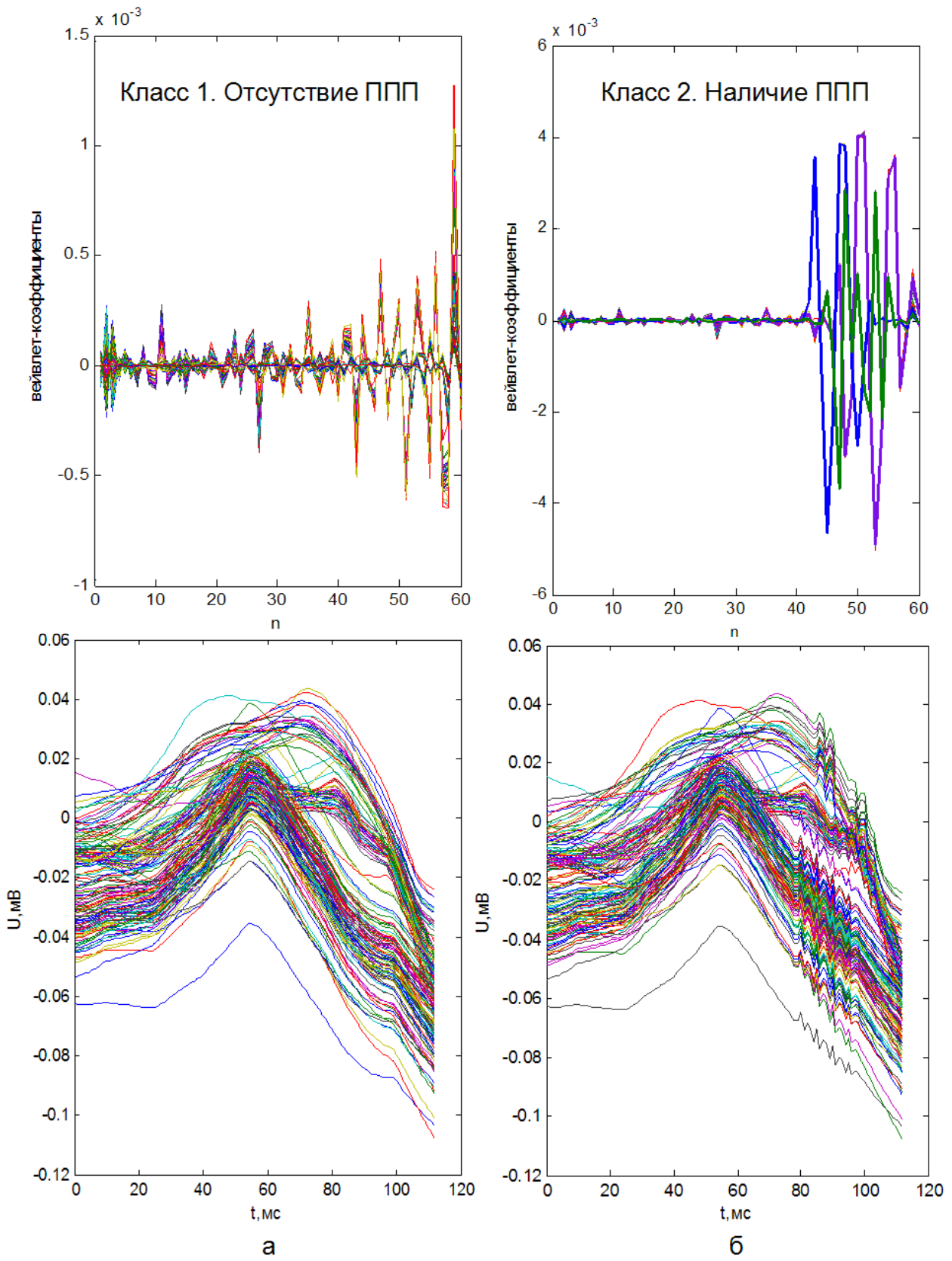


Рис. 6. Разделение выборки данных на 2 класса:
 а) класс 1: "отсутствие ППП"; б) класс 2: "наличие ППП"

Выводы

При распознавании образов электрокардио-сигналов с ППП в системах ЭКГ ВР предлагается использовать в качестве признаков численные характеристики вектора вейвлет-коэффициентов d1 уровня детализации ДВР.

Анализ результатов, полученных для экспериментальной выборки данных при кластеризации образов ППП в пространстве их признаков, показал, что:

- Независимо от формы Р-зубца, отличающейся для разных пациентов, для всех вариантов рассмотренных признаков точки кластера “отсутствие ППП” в пространстве признаков хорошо локализованы, что подтверждается малыми значениями среднего расстояния от точки до центра данного кластера.
- Точки кластера “наличие ППП” хорошо сгруппированы только для каждого отдельного пациента и разнесены в пространстве признаков для разных пациентов. Это объясняется тем, что при формировании выборки данных выбирались индивидуальные проявления ППП для каждого из пациентов.
- Предложенный метод формирования признаков с помощью ДВР позволяет получить хорошо сгруппированные неперекрывающиеся кластеры для классов “отсутствие ППП” и “наличие ППП”.

Литература

1. *Новые методы электрокардиографии.* Под ред. С.В. Грачева, Г.Г. Иванова, А.Л. Сыркина. -М.:Техносфера, 2007.— 552с.
2. *Руководство по кардиологии/*Под ред. В.Н.Коваленко – К.: Морион, 2008.– 1404 с.
3. *Л.Т.Малая, А.И.Дядык, А.Э.Багрий.* Патогенез фибрилляции предсердий. Сообщение 1// Український терапевтичний журнал — 2002.— Т.4, №2.— С. 58-65.
4. *Ту Дж.Т., Гонсалес Р.К.* Принципы распознавания образов. - М.: Мир, 1978. - 411 с.
5. *Абакумов В.Г., Рибін О.І., Сватош Й.* Біомедичні сигнали. Генезис, обробка, моніторинг - К.: Нора-принт, 2001. - 516 с.
6. *Е.О. Иванько, Н.Г. Иванушкина, Ю.В. Прокопенко.* Моделирование процессов возникновения циркуляции волны возбуждения в миокарде// Управляющие системы и машины – 2009.– №3.– с.36-41.
7. *Metin Akay (Ed.).* Time-Frequency and Wavelets in Biomedical Signal Processing. — IEEE Press, 1997. — 652 pp.
8. *Смоленцев Н.К.* Основы теории вейвлетов. Вейвлеты в MATLAB. — М.: ДМК Пресс, 2008.— 448 с.
9. *Е.О. Иванько, Н.Г. Иванушкина, Ю.С. Синькоп.* Многоуровневый анализ электрокардиограмм для выявления поздних потенциалов предсердий// Электроника и связь: тематический выпуск “Электроника и нанотехнологии”. – 2009.– №4-5.– с.160-164.