

УДК 616.12- 073.7

Е.О. Иванько, Н.Г. Иванушкина, канд. техн. наук, Н.А. Матвеева, С.И. Споревой

Технологии распознавания образов поздних потенциалов предсердий: подходы к классификации

В данной статье задача распознавания образов поздних потенциалов предсердий (ППП) решается путем формирования признаков в базисе собственных векторов. Матрица ковариаций формируется для ансамбля вейвлет-коэффициентов детализации, полученных при многоуровневом вейвлет-разложении электрокардиосигналов (ЭКС). Приводятся результаты кластерного анализа при модельном эксперименте по классификации кардиоциклов различных пациентов с наличием и отсутствием ППП на фоне шума.

In this paper the problem of atrial late potentials (ALP) recognition is solved by the formation of their signs in the basis of eigenvectors. Covariance matrix is formed for an ensemble of wavelet detail coefficients obtained by multi-level wavelet decomposition of electrocardiosignals (ECS). The results of cluster analysis in a model experiment on the classification of ECS with and without ALP on the background noise are presented.

Ключевые слова: электрокардиосигнал, поздние потенциалы предсердий, распознавание образов, вейвлет-разложение, собственные числа и вектора, кластерный анализ, классификация.

Введение

Появление низкоамплитудных высокочастотных потенциалов в терминальной части Р зубца ЭКС отражает наличие замедленной фрагментированной предсердной деполяризации, ассоциирующейся с возникновением кругов "риентри" ("re-entry") и развитием предсердных тахикардий [1-2]. Трудности неинвазивной регистрации ППП и их выявления на фоне основного кардиосигнала связаны с тем, что амплитуда шумовых составляющих может значительно превышать амплитуду ППП. Для повышения отношения сигнал/шум в системах электрокардиографии высокого разрешения используется усреднение множественных идентичных кардиоциклов [3].

Учитывая, что стандартные методики обнаружения ППП, основанные на временном и спектральном анализе, имеют невысокую предсказывающую ценность, актуальной является разработка алгоритмов распознавания образов

ППП с использованием дополнительных диагностических признаков, полученных в результате ортогональных преобразований кардиосигнала.

В ходе многочисленных модельных экспериментов было выявлено, что признаки наличия ППП содержатся в коэффициентах детализации первого уровня cD_1 , отражающих картину высокочастотных низкоамплитудных всплесков при дискретном вейвлет-разложении (ДВР) Р зубца до 5-го уровня [4]. Использование в качестве признаков численных характеристик вектора вейвлет-коэффициентов cD_1 при распознавании образов ППП позволило получить хорошо сгруппированные неперекрывающиеся кластеры для классов "отсутствие ППП" и "наличие ППП" [5]. Однако присутствие остаточного шума в усредненной ЭКГ реальной диагностической системы может значительно снизить достоверность диагностики, так как шумовая составляющая и ППП проявляются на одном уровне вейвлет-разложения анализируемого Р зубца [6]. Предлагаемый комплексный метод выявления ППП основан на многоуровневом дискретном вейвлет-разложении исследуемого Р зубца с последующим анализом собственных чисел и векторов ковариационной матрицы коэффициентов детализации первого уровня вейвлет-разложения [6].

Формирование признаков образов ППП

При проведении модельного эксперимента в системе MATLAB в качестве нормы рассматривались предварительно очищенные от высокочастотных составляющих Р зубцы, выделенные из реальных записей ЭКГ различных пациентов. Затем в эти сигналы был добавлен аддитивный белый гауссовский шум с отношением сигнал/шум, равным 60дБ [6]. При начальном предположении о возможности "нормы" или "патологии" кардиоциклы с такими Р зубцами представляют норму, так как ППП в них отсутствуют (рис.1, а). Далее в терминальную часть каждого Р зубца с шумом были добавлены смоделированные ППП. В качестве моделей ППП использовались отрезки модулированной синусоиды, а также потенциалы действия, полученные в результате решения уравнений Ходжкина-Хаксли [7]. При этом для каждого пациента ППП отличались по своей амплитуде, форме и количеству. Анализ кардиоциклов с такими Р

зубцами предполагает обнаружение ППП заранее неизвестного вида на фоне шума (рис. 2, б). Полная выборка данных, сформированная из записей электрокардиограмм (ЭКГ) 5 пациентов, составила 460 кардиоциклов.

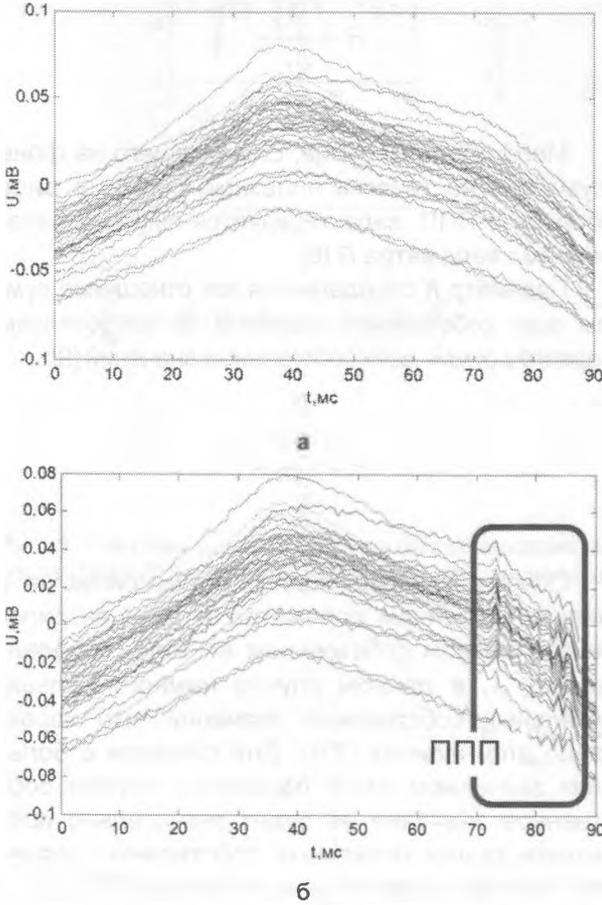
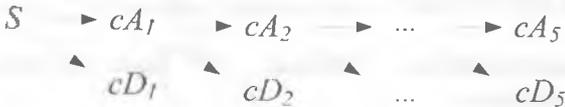


Рис. 1. Ансамбль Р зубцов ЭКГ одного из пациентов: а) отсутствие ППП; б) наличие ППП

Для Р зубцов исходных кардиоциклов были получены наборы вейвлет-коэффициентов при ДВР до 5-го уровня с использованием вейвлет-функции “симлет” 4-го порядка:



При ДВР векторы вейвлет-коэффициентов получают сверткой исследуемого кардиосигнала S с фильтром нижних частот для определения коэффициентов аппроксимации cA_1 и с фильтром высоких частот для определения коэффициентов детализации cD_1 [8]. На следующем шаге раскладываются коэффициенты аппроксимации cA_1 на две части, с получением cA_2 и cD_2 и т.д. до необходимого уровня разложения, в нашем случае 5-го.

Пусть i -му Р зубцу в ансамбле соответствует набор вейвлет-коэффициентов первого уровня детализации cD_1 , который в векторном виде представляется как $W_i = [w_1 \ w_2 \ w_3 \ \dots \ w_N]$, где N — количество элементов вектора коэф-

фициентов детализации, соответствующее данному уровню вейвлет-разложения. Тогда для ансамбля из M Р зубцов набор вейвлет-коэффициентов первого уровня детализации cD_1 будет иметь вид матрицы:

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1N} \\ w_{21} & w_{22} & \dots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \dots & w_{MN} \end{bmatrix}$$

где w_{ij} — значение j -го коэффициента первого уровня детализации для i -го Р зубца.

Ковариационная матрица вейвлет-коэффициентов, соответствующих ансамблю из M Р зубцов имеет вид:

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1N} \\ C_{21} & C_{22} & \dots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{M1} & C_{M2} & \dots & C_{MN} \end{bmatrix}$$

где элементы C_{ij} матрицы C определены следующим образом:

$$C_{ij} = \sum_{k=1}^M (w_{ki} - m_k)(w_{kj} - m_k),$$

где m_k — среднее значение вейвлет-коэффициентов cD_1 , рассчитанное для Р зубца k -го кардиоцикла.

Для собственных векторов V_1, V_2, \dots, V_N матрицы C , представляющих собой ортонормированный базис вейвлет-коэффициентов cD_1 , определённых для ансамбля Р зубцов ЭКГ, выполняется условие

$$CV = \lambda V,$$

где $V = [V_1, V_2, \dots, V_N]$ — собственные векторы матрицы C , $\lambda_1, \lambda_2, \dots, \lambda_N$ — собственные значения, причем $\lambda_1 > \lambda_2 > \dots > \lambda_N$

Для нахождения признаков образов ППП было составлено 46 ансамблей, в каждый из которых входило по 10 наборов вейвлет-коэффициентов cD_1 . В результате разложения в базисе собственных векторов были получены наборы собственных чисел $\lambda_1, \lambda_2, \dots, \lambda_N$ и собственных векторов $V = [V_1, V_2, \dots, V_N]$ матриц ковариации, найденных для каждого ансамбля вейвлет-коэффициентов. В полученном базисе главные собственные векторы, соответствующие наибольшему собственному значению $\lambda_1 > \lambda_2 > \dots > \lambda_K$ ($K < N$), несут основную информацию об исследуемом сигнале.

Величина $L = \frac{\lambda_1}{\sum_{k=1}^N \lambda_k} \cdot 100\%$ используется

для оценки доли информации, заключенной в компоненте, определяющей по собственному

вектору V_1 . При разложении в базисе собственных векторов вейвлет-коэффициентов cD_1 сигнала с наличием ППП на фоне шума видно, что $L_1=62\%$, т.е. наибольшая доля информации содержится в первой компоненте (рис.2, б). В случае отсутствия ППП, т.е. при анализе вейвлет-коэффициентов cD_1 , содержащих только шум, $L_1=18\%$ (рис.2, а).

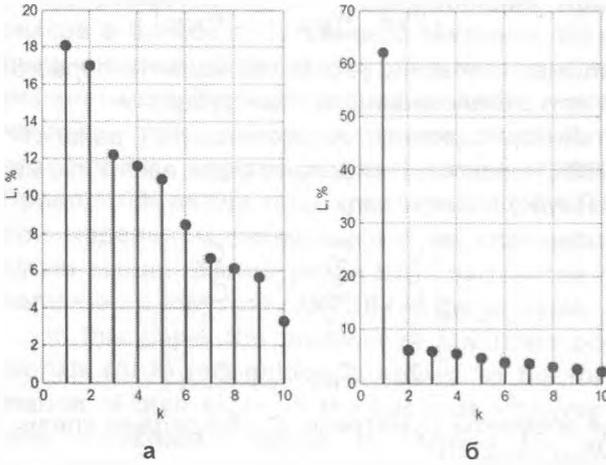


Рис. 2. Параметр L для первых 10-ти собственных чисел: а) отсутствие ППП на фоне шума; б) наличие ППП на фоне шума

Таким образом, собственный вектор V_1 , соответствующий наибольшему собственному числу λ_1 , определяет главную компоненту, содержащую признаки ППП. Эта компонента для случая отсутствия ППП на фоне шума имеет вид, свойственный помехам (рис.3, а), тогда как при наличии ППП она содержит характерные всплески, отражающие особенности поздних потенциалов конкретного пациента в базисе собственных векторов (рис.3, б-е).

Индивидуальные особенности пациентов с анатомическими и/или функциональными нарушениями в миокарде обуславливают широкий диапазон изменений амплитуды, формы, длительности и времени появления ППП. Проведенные численные эксперименты показали, что при априорной неопределенности о характеристиках шума и выявляемого низкоамплитудного сигнала использование в качестве признаков ППП таких числовых характеристик, как энергия (L^2 -норма), среднее абсолютное значение или среднеквадратическое отклонение главной компоненты, приводит к перекрыванию классов "норма" и "патология" при распознавании образов ППП. Поэтому для вычисления признаков ППП предлагается использовать параметры, основанные на использовании собственных значений матрицы ковариации, определенной для ансамбля вейвлет-коэффициентов детализации cD_1 .

Одним из параметров, используемых для оценки высокочастотной составляющей P зубца в пространстве собственных векторов, является отношение первого собственного значения к сумме всех собственных значений [8]:

$$R = \frac{\lambda_1}{\sum_{i=1}^N \lambda_i}$$

Морфология P зубца, содержащего на фоне шума диагностически полезный сигнал в виде всплесков ППП, характеризуется высокими значениями параметра R [6].

Параметр K определяется как отношение суммы всех собственных значений за исключением первого к сумме всех собственных значений [9]:

$$K = \frac{\sum_{i=2}^N \lambda_i}{\sum_{i=1}^N \lambda_i}$$

Основная информация о форме сигналов с малым значением параметра K сконцентрирована в первом собственном векторе, следовательно, λ_1 в данном случае намного больше остальных собственных значений, что характерно для наличия ППП. Для сигналов с большим значением этого параметра первое собственное значение не будет значительно превышать сумму остальных собственных значений, что наблюдается при отсутствии ППП [6].

Также в качестве диагностического признака ППП в координатном базисе собственных векторов можно использовать параметр S — след диагональной матрицы собственных чисел, определяемый как сумма всех собственных значений ковариационной матрицы вейвлет-коэффициентов детализации cD_1 .

Кластеризация данных в пространстве признаков

Анализ группировки точек в трёхмерном пространстве признаков $[S, K, R]$ свидетельствует о наличии двух кластеров данных. При этом кластер с плотной упаковкой точек соответствует данным, относящимся к P зубцам с отсутствием ППП, представляющим класс "норма". Достаточно большой, по сравнению с первым случаем, разброс признаков внутри второго кластера, представляющего класс "патология", объясняется индивидуальным видом ППП для каждого из пациентов. При этом классы не перекрываются в пространстве признаков (рис.4).

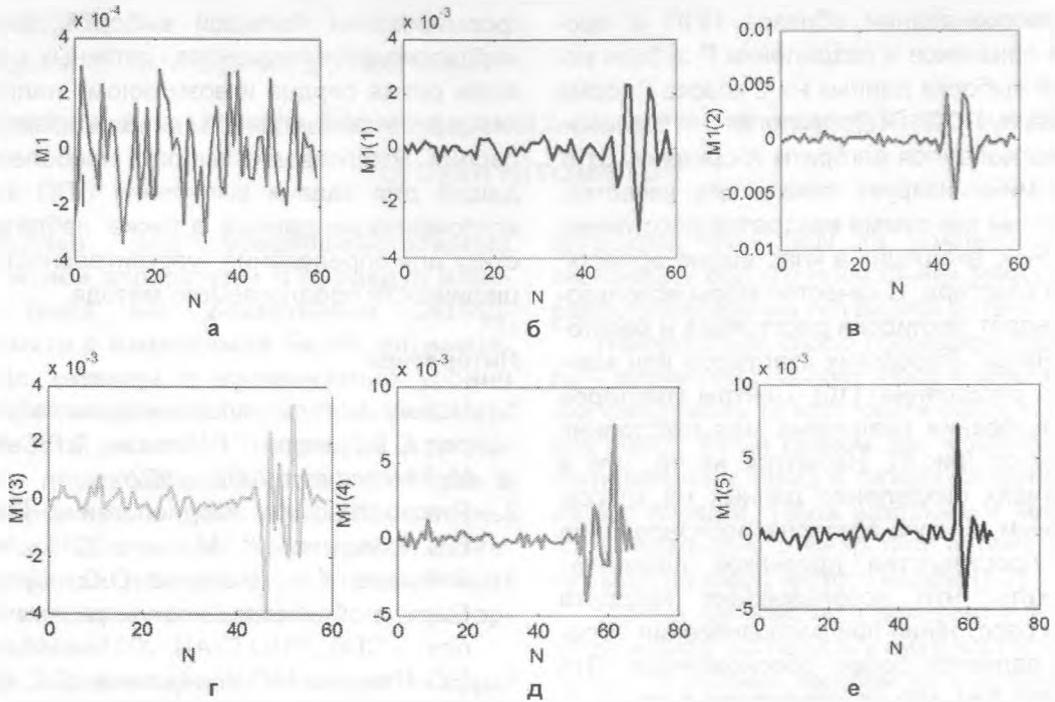


Рис. 3. Главная компонента, соответствующая наибольшему собственному значению: а) для случая отсутствия ППП на фоне шума; б-е) для случаев наличия ППП различных видов на фоне шума

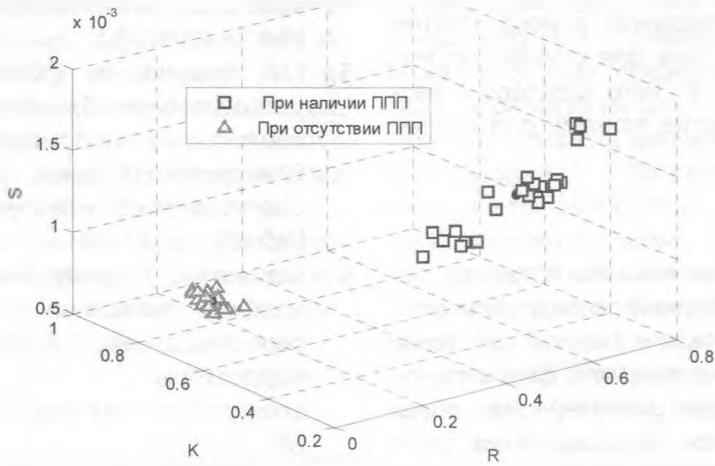


Рис. 4. Результаты кластеризации данных в пространстве признаков: □ — патология, △ — норма.

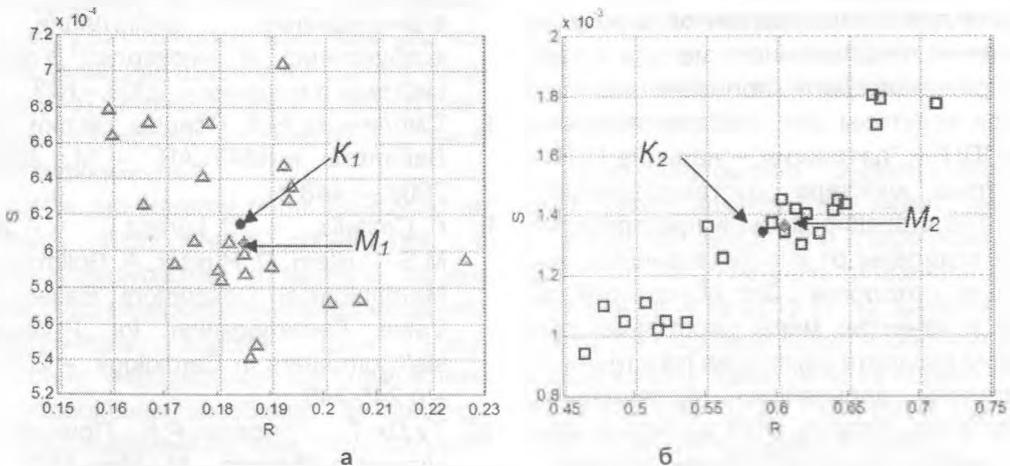


Рис. 5. Расположение центров кластеров “норма” (а) и “патология” (б): K_1 и K_2 — при использовании в качестве меры квадрата евклидова расстояния, M_1 и M_2 — “city-block” (городских кварталов или манхэттенского расстояния)

При распознавании образов ППП в пространстве признаков и разделении Р зубцов исследуемой выборки данных на 2 класса (“норма – отсутствие ППП” и “патология — наличие ППП”) использовался алгоритм K-средних. Этот алгоритм минимизирует показатель качества, определенный как сумма квадратов расстояний от всех точек, входящих в кластерную область, до центра кластера. В качестве меры использовались квадрат евклидова расстояния и расстояние “city-block” (городских кварталов или манхэттенское расстояние) [10]. Центры кластеров при использовании различных мер расстояния не совпадают (рис.5). Несмотря на то, что в обоих случаях разделение данных на классы было верным, анализ паттерна распределения точек в пространстве признаков позволяет предположить, что использование квадрата евклидова расстояния при распознавании образов ППП является более обоснованным. Это объясняется тем, что по сравнению с расстоянием “city-block” квадрат евклидова расстояния придаёт большие веса более отдаленным друг от друга объектам [10], а для исследуемых данных характерно образование одного плотно сгруппированного кластера для случая “норма” и второго отдаленного от него кластера с разнесенными в пространстве точками для случая “патология”.

Выводы

Применение математического аппарата ортогональных преобразований посредством объединения вейвлет-анализа и метода собственных векторов позволило получить диагностически важные признаки при распознавании образов поздних потенциалов предсердий на фоне шумовой составляющей ЭКГ.

Независимо от форм Р зубцов и видов ППП, отличающихся для разных пациентов, в результате применения предложенного метода в пространстве признаков были получены неперекрывающиеся кластеры для классов “норма – отсутствие ППП” и “патология – наличие ППП”. При этом точки кластера, соответствующего “норме”, плотно сгруппированы в пространстве признаков и отдалены от кластера данных, характерных для “патологии”. Это обуславливает применение в качестве меры расстояния при кластеризации квадрата евклидова расстояния.

Направлением дальнейших исследований по распознаванию образов ППП является проведение клинических исследований с целью

формирования большой выборки данных из кардиосигналов пациентов, склонных к нарушениям ритма сердца и возможному наличию потенциалов замедленной деполяризации предсердий. Это позволит выбрать наиболее подходящий для задачи выявления ППП алгоритм классификации данных, а также набрать статистику для определения чувствительности и специфичности предлагаемого метода.

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. Зайченко К.В., Жаринов О.О., Кулин А.Н. Съём и обработка биоэлектрических сигналов — СПб.: РИО ГУАП, 2001.—140 с.
4. Е.О.Иванько, Н.Г.Иванушкина, Ю.С.Синекоп. Многоуровневый анализ электрокардиограмм для выявления поздних потенциалов предсердий// Электроника и связь: тематический выпуск “Электроника и нанотехнологии”. – 2009 – №4-5.– с.160-164.
5. Н.Г.Иванушкина, Е.О.Иванько. Технологии распознавания образов поздних потенциалов предсердий: формирование признаков.// Электроника и связь: тематический выпуск “Электроника и нанотехнологии”. – 2010 – №5(58).– с.177-184.
6. Иванько К.О., Иванушкина Н.Г. Комплексний метод виявлення пізніх потенціалів передсердь на основі аналізу власних підпросторів вейвлет-образів електрокардіосигналів // Наукові вісті. – 2010 – №6.– с.11-18.
7. Е.О.Иванько, Н.Г.Иванушкина, Ю.В.Прокopenко. Моделирование процессов возникновения циркуляции волны возбуждения в миокарде// Управляющие системы и машины – 2009.– №3.– с.36-41.
8. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс, 2008.— 448 с.
9. F. Castells, J. Lorenz, A.M. Climent, M.S. Guillem, D. Husser, A. Bollmann, J. Millet. Morphological Descriptors Based on Eigen Value Decomposition for P-Wave Analysis//Computers in Cardiology. – 2008. – №35. – p.245-248.
10. Ту Дж.Т., Гонсалес Р.К. Принципы распознавания образов. - М.: Мир, 1978. - 411 с.