

3. Взаимокомпенсация TGF-моды в гемодинамике васкулярного дерева нефронов / Д. Д. Постнов, О. В. Сосновцева, Д. Э. Постнов // Известия Саратовского университета. Сер. Физика. – 2011. – Т. 11., вып. 2. – С. 66–71. – Режим доступа: [http://www.sgu.ru/sites/default/files/journals/izvestiya/pdf/2013/12/13/66-71\\_Postnov.pdf](http://www.sgu.ru/sites/default/files/journals/izvestiya/pdf/2013/12/13/66-71_Postnov.pdf)
4. Емельянова Ю. П. Динамика связанных нефронов и режим широкополосной синхронизации / Ю. П. Емельянова, А. П. Кузнецов, Э. Мозекилде, Я. Л. Лаугесен // Нелинейная динамика. – 2012. – Т. 8, № 5. – С. 875 – 896.
5. Емельянова Ю. П. Особенности синхронизации неидентичных нефронов. Широкополосная синхронизация / Ю. П. Емельянова, А. П. Кузнецов, Е. Мозекилде. – [Электронный ресурс]. – Режим доступа: <http://sgtnd.narod.ru/science/Nephron.pdf>
6. Holstein-Pathlou N-H. Oscillations and chaos in renal blood flow control / N-H. Holstein-Pathlou // JASN. – 1993. – Vol. 4. no. 6. – P. 1275 – 1287.
7. Heterogeneity and weak coupling may explain the synchronization characteristics of cells in the arterial wall / J.C.B. Jacobsen, C. Aalkjaer, V.V. Matchkov [et al.] // Phil. Trans. R. Soc. A. – 2008. – Vol. 366. – P. 3483–3502. doi:10.1098/rsta.2008.0105
8. Nephron blood flow dynamics measured by laser speckle contrast imaging / N-H. Holstein-Rathlou, O.V. Sosnovtseva, A.N. Pavlov [et al.] // American Journal of Physiology - Renal Physiology Published. – 2011. – Vol. 300, no. 2. – P. F319-F329. DOI: 10.1152/ajprenal.00417.2010
9. Сницер А. Р. Определение фильтрационных свойств пористо-упругой среды на основе решения одной краевой задачи для уравнений Био / А. Р. Сницер // Динамические системы. – 2012. – Т. 2 (30), № 3-4. – С. 323 – 335.
10. Устойчивость, структуры и хаос в сетях синхронизации / В. С. Афраимович, В. И. Некоркин, Г. В. Осипов, В. Д. Шалфеев. – Горький: ИПФ АН СССР, 1989. – 256 с.
11. Анисимов А. А. Исследование динамики нефронов на основе вейвлет-анализа и метода эмпирических мод / А. А. Анисимов / Автореферат дисс. на соискание уч. ст. к. ф.-м. н. по специальности 03.01.02 – биофизика. – Саратов. – 2011. – 19 с.
12. Ніколов М. О. Комп'ютерний аналіз кількості функціонуючої паренхіми нирок за даними нефросцинтиграфії з  $^{99m}\text{Tc}$ -ДМСО / М. О. Ніколов, Д. О. Супрунюк, А. Л. Камінська [та інш.] // Український радіологічний журнал. – 2014. – Т. XXII, вип. 2. – С. 98 – 101.
13. Винер Н. Нелинейные задачи в теории случайных процессов / Н. Винер. – М.: Изд. инлит. – 1961. – 160с.

Надійшла до редакції  
18 лютого 2015 року

© Николов Н. А., Коваленко Н.Н., Супрунюк Д. А., Каминская А. Л. Кундин В. Ю., 2015

УДК 612.171.1+004.852

## ОТРАБОТКА ПРИНЦИПОВ АНАЛИЗА СТРУКТУРЫ ЦИКЛИЧЕСКИХ МЕДИКО-БИОЛОГИЧЕСКИХ СИГНАЛОВ ДЛЯ ИХ ОБНАРУЖЕНИЯ, РАСПОЗНАВАНИЯ И КЛАССИФИКАЦИИ

Шачиков А. Д., Шуляк А. П.,

Национальный технический университет Украины "Киевский Политехнический Институт",  
г. Киев, Украина

*Значительные резервы в повышении достоверности диагнозов пациентам содержатся в целесообразном выборе описательных категорий формы сигналов средств диагностики, который может быть сделан на этапе обучения распознающих систем. Дополнительные возможности повышения чувствительности и специфичности распознающих процедур могут быть созданы благодаря определению рационального состава данных для их анализа с*

целью распознавания сигналов. Цель работы – совершенствование инструментария обнаружения, распознавания и классификации циклических медико-биологических сигналов в системах диагностики заболеваний пациентов.

Исследуются базовая модель анализируемых сигналов и соответствующая ей последовательность операций их анализа. Исследование построено как тестовый вариант проверки статистическим путем результативности операций формирования характеристик формы сигналов при распознавании *QRS*-комплексов с использованием данных пациента с достоверной маркировкой типов таких комплексов. Раскрываются дополнительные возможности повышения правильности принимаемых решений за счет определения на этапе обучения распознающей системы рационального состава данных для их анализа с целью распознавания наблюдаемых процессов. Анализатор сигналов построен по общей схеме распознающих систем, обучаемых с учителем. Приводятся дополнительные решения и иллюстрации возможностей повышения правильности решений при распознавании сигналов. Представленный вариант отработки принципов анализа структуры циклических медико-биологических сигналов ранее не рассматривался.

**Ключевые слова:** медико-биологические сигналы, анализ структуры, распознавание, обучение с учителем.

## **Введение**

Среди характеристик анализируемых медико-биологических процессов в диагностике заболеваний значительный вес приобрели характеристики формы наблюдаемых сигналов [1, 2, 3]. Результативность выбора состава и конкретного содержания описательных категорий их формы при распознавании сигналов проявляется в конечном счете в достоверности постановки диагнозов, в чувствительности, специфичности распознающих процедур [1, 3, 4, 5]. В данной работе обсуждаются принципы и операции анализа структуры циклических медико-биологических сигналов с использованием характеристик их формы, которые отделены от проявления факторов масштаба (усиления) и сдвига (изменения постоянной составляющей) в результатах регистрации наблюдаемых процессов. Приводится соответствующая базовая модель, лежащая в основе выделения таких характеристик.

На тестовом примере обработки электрокардиограмм конкретного пациента в задаче распознавания типов *QRS*-комплексов раскрывается предварительная оценка состоятельности рассматриваемых принципов и операций анализа структуры сигналов, включая отбор данных для анализа. Демонстрируются возникающие при этом возможности повышения правильности решений при распознавании сигналов. Анализатор построен в соответствии с общей структурой распознающих систем, обучаемых с учителем.

## **Допущения и базовая модель сигналов**

Понятие формы сигнала как формы кривой на его графике, как характеристики формы для вектора, изображающего эту кривую в многомерном пространстве, может быть выражено представлением об этой кривой как о жесткой конструкции, которая сохраняется при смещении графика по вертикали и его повороте вокруг любой точки на поле графика (вокруг начала координат). Форма сигнала выражается в соотношении его отсчетов; при векторном пред-

ставленні – це орієнтація зображуючого сигнал вектора.

Если форма сигналів вибрана для їх виявлення, розпізнавання і класифікації, то ця характеристика повинна бути відділена при обробці даних від проявлення факторів іншого типу, серед яких, крім названих, - масштаб представлення цієї форми. Змінюваність масштабу, як перешкоджаючий фактор при аналізі сигналів, може бути наслідком варіації властивостей як спостережуваного об'єкта, так і властивостей використовуваних вимірних приладів, - по причині впливу зовнішніх умов або інших факторів. Отже, масштаб представлення процесів на графіку зручно розглядати як окрему характеристику форми, який не повинен впливати на характеристики їх форми.

Цілесловесним є вимога взаємної незалежності характеристик форми, зміщення і масштабу в описанні сигналу в його моделі як основи для подібного аналізу його структури, передбачаючої набір необхідних операцій

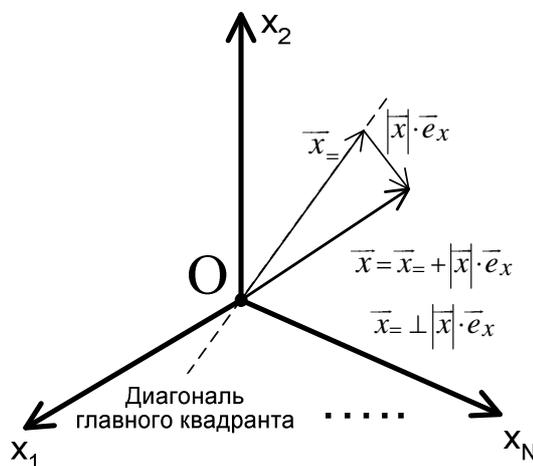


Рис. 1. Базова модель аналізованого сигналу

і алгоритми аналізу в цілому. Таким чином, інваріантність описань форми сопоставляються сигналів і взаємна незалежність компонентів формуваних образів сприяють створенню сприятливих умов для виявлення сигналів, їх розпізнавання і класифікації [4, 1, 5, 3]. При цьому вибрані характеристики форми сигналів повинні достатньо контрастними і стійкими. В відповідності з

вказаними принципами, аналізована характеристика форми сигналу не повинна залежати як від постійної складової в його складі, так і від його інтенсивності (характеристики масштабу представлення аналізованого графіка), тому в складі сигналу виділяємо ці компоненти і об'єднуємо їх в одній моделі (рис. 1).

Формально це виражається математичною моделлю такого виду:

$$\vec{x} = \vec{x}_\parallel + |\vec{x}| \cdot \vec{e}_x, \quad \vec{x} - \text{вектор відліку сигналу } x(t) \text{ на рівнодискретній сітці}$$

часу  $\{t_i\}$ ,  $i = \overline{1, N}$ ,  $\vec{x}_\parallel$  - вектор постійної складової сигналу (вектор з однаковими компонентами, орієнтований по діагоналі головного квадранта  $N$ -мерного простору  $Ox_1x_2\dots x_N$ ),  $\vec{e}_x$  - орт цього вектора, його координати - косинуси напрямлюючих кутів вектора відносно координатних осей, сума квадратів яких дорівнює одиниці. Нехай частота відліку сигналу вибрана достатньо успішно для фіксації його форми.

З такої моделі випливає, що відділення характеристики форми сигналу від

других составляющих его описания можно выполнить в две таких операции.

1. Удаление из сигнала характеристики масштаба – приведение длин (модулей) рассматриваемых векторов к одному значению. Уравнивать их удобно на уровне единицы по формуле:  $\vec{x} \rightarrow \vec{x}/|\vec{x}|$ , где  $|\vec{x}| = \sqrt{(\vec{x}, \vec{x})}$ . Энергии сигналов, представленных такими векторами, также уравниваются. Кроме того, становится корректным использование корреляционных интегралов для сравнения формы сигналов.

2. Удаление из сигнала фактора сдвига (постоянных составляющих сравниваемых векторов):  $\vec{x} \rightarrow \vec{x} - \bar{x}$ ,  $\bar{x} = \sum_{i=1}^N x_i$ , где  $N$  – количество отсчетов, подвергаемых сравнению. Сохранение ненулевой постоянной составляющей одной и той же величины, как показали исследования, нецелесообразно.

Рассмотренные базовая модель и порядок выделения в сигнале характеристики его формы использованы в разработке структуры анализатора и алгоритма анализа сигналов.

Предварительные исследования на записях ЭКГ пациентов [6] показали, что описания  $\bar{e}_x$  формы сигнала, выделяемые в соответствии с приведенной базовой моделью, являются для кардиограмм характерными, т.е. повторяющимся и относительно стабильными для всех *QRS*-комплексов. На тестовой записи было установлено, что статистического среднего характеристики  $\bar{e}_x$  на обучающей выборке пациента достаточно, чтобы с высокой степенью правильности обнаруживать и распознавать *QRS*-комплексы в этой записи – правильно позиционировать на кардиограмме окна принятия решений и определять типы этих комплексов, вычисляя корреляционные интегралы с использованием указанных эталонов. Результативность их применения нашла также предварительное подтверждение на повторной, контрольной записи ЭКГ того же пациента [6], снятой через некоторое время.

### Структура анализатора и алгоритм анализа сигналов

Строение анализатора представлено для случая циклического входного процесса – последовательности однотипных явлений без их взаимного перекрытия во времени, как, например, в электрокардиографии [2]. Более сложный вариант, с возможным перекрытием характерных элементов сигнала, что свойственно, например, электроэнцефалографии [2], не обсуждается. Иллюстрации анализа данных в работе представлены для электрокардиографии на примере обработки данных конкретного пациента в одном отведении [6]. Рассматривается задача распознавания типов *QRS*-комплексов. Длительность тестовой записи – 1800 с, 2650 кардиоциклов, их количество по типам:  $N$  – 2360,  $A$  – 179,  $V$  – 110.

В основе построения анализатора – общая структура распознающей системы

с обучением принятию решений с учителем. В составе анализатора сигналов (рис. 2) – два основных блока: формирования априорных данных распознающей системы и обработки вновь поступивших реализаций кардиограмм (с распознаванием типов *QRS*-комплексов). Блоки пополнены формирователем оценок качества распознавания сигналов.



Рис. 2. Структура анализатора сигналов

Первый блок обеспечивает реализацию этапа обучения распознающей системы, второй – работу по распознаванию вновь поступивших сигналов. И если в первом случае данные ЭКГ в одном отведении (отведении I) сопровождаются маркировкой типов *QRS*-комплексов (запись полагается полностью достоверной), то во втором – маркировка отсутствует и должна быть восполнена анализатором.

Оценка качества распознавания в первом блоке основана на сопоставлении исходной и полученной анализатором маркировок – с вычислением частот правильных и ошибочных решений. В первом случае имеются в виду оценки чувствительности распознающих процедур анализатора, отражающей результативность не только самих таких процедур, но и качество формируемых в анализаторе портретов сигнала, а также степень их соответствия портретов процедурам.

В целом, общий алгоритм анализа сигналов, отвечающий представленной структуре анализатора, состоит в следующем.

1. Портреты сигналов – характеристики их формы вида  $\vec{e}_x$ . Вначале они выделяются, затем проводится их анализ. Выделение характеристики  $\vec{e}_x$  реализуется в два этапа – локализации комплексов и выделения из них составляющих формы. Фиксация положения комплексов производится по разметке  $R-R$ -интервалов в записи. Взаимная фазовая привязка комплексов выполняется путем совмещения  $R$ -пиков комплексов.

2. Далее для всех комплексов применяется одинаковое окно анализа. Его ширина выбрана для захвата при анализе всех комплексов записи в их основной части. Соотношение ширины окна слева и справа от  $R$ -пика взято неизменным – в соответствии типовыми параметрами комплексов [2]. Возможность выделения характеристики  $\vec{e}_x$  формы комплексов обеспечивается стабилизацией ну-

левой линии записи (рис. 3) и удалением постоянной составляющей из сигнала в каждом окне. По полученным результатам, стабилизация нулевой линии тестовой записи (обучающей выборки) при использовании одной и той же процедуры принятия решений существенно повышает правильность распознавания типов *QRS*-комплексов – с 78%, 49% и 64% для *N*, *A* и *V*- типов соответственно до 77%, 86% и 90%.

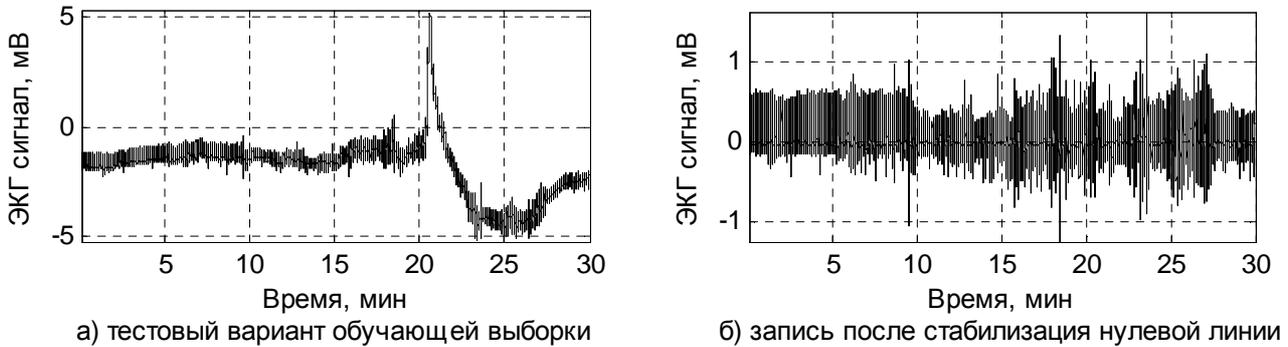


Рис. 3. Стабилизация нулевой линии ЭКГ (стандартное отведение I)

3. Далее решается задача масштабирования наблюдаемого процесса в каждом окне, что приводит к получению требуемой характеристики  $\bar{e}_x$  формы сигнала. Масштабирование анализируемых сигналов, как показала обработка тестовых данных, повышает правильность обнаружения и распознавания *QRS*-комплексов в записях ЭКГ. Для анализа формы *QRS*-комплексов на этапе обучения они сортируются и подсчитываются по типам в соответствии с классификацией, принятой в базе данных [6], – типам *N* (*Normal beat*), *A* (*Atrial premature beat*), *V* (*Premature ventricular contraction*). Для каждого типа *QRS*-комплексов формируется комплекс-среднее (рис. 4), который используется в дальнейшем как эталон и входит в состав сформированных априорных данных распознающей системы.

5. На этапе обработки вновь поступившей реализации характеристики  $\bar{e}_x$  формы сигнала выделяются в том же самом порядке. Результат сравнивается с каждым эталоном. Сравнение выполняется на основе расчета корреляционного интеграла, корректность применения которого была подготовлены нормировкой (масштабированием) сигналов – энергии сравниваемых процессов между собой равны и их значения единичны. Общая схема анализа сигналов при их распознавании отрабатывалась с использованием простейшего критерия принятия решений – выбирался тот вариант (*N*, *A* или *V*), для которого корреляционный интеграл имел наибольшее значение. Обсуждение вариантов решений с использованием взаимно-ортогональных разложений сигналов [7] и с выделением периодических негармонических «непрерывных» и импульсных компонент, адаптивных форме сигналов [8], вынесено за рамки данной работы.

6. Для оценки качества распознавания сигналов по данным обучения подсчитывались правильные и ошибочные решения для трех возможных исходов. Результат представлялся в виде диаграмм (рис. 5).

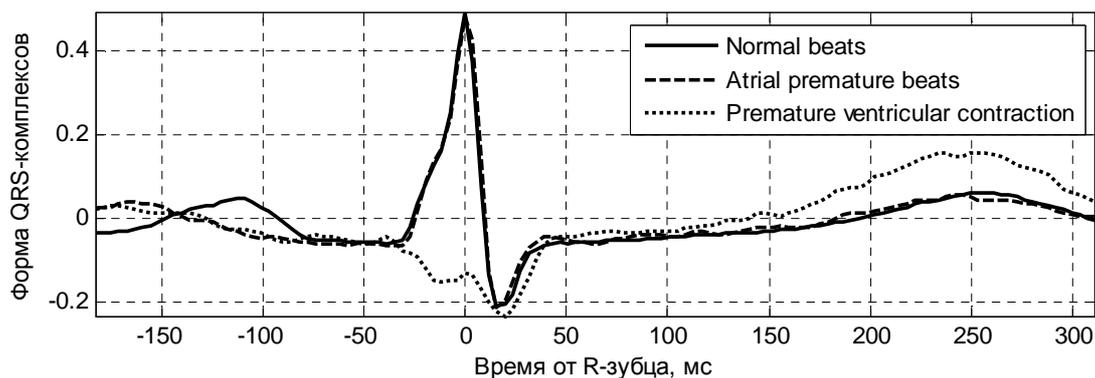


Рис. 4. Эталоны  $QRS$  - комплексов  $N$ ,  $A$  и  $V$  - типов в окне анализа

При оценке качества распознавания  $QRS$  - комплексов текущей реализации ЭКГ рассчитываются два значения приведенного корреляционного интеграла – среднее по результатам обучения и текущее для каждого комплекса, например, для  $N$  - типа (рис. 6).

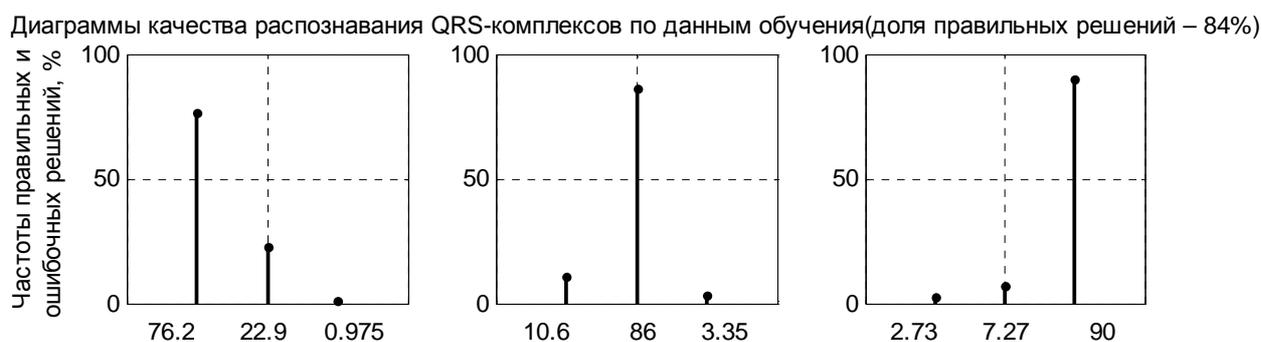


Рис. 5. Качество распознавания  $QRS$  - комплексов по данным обучения

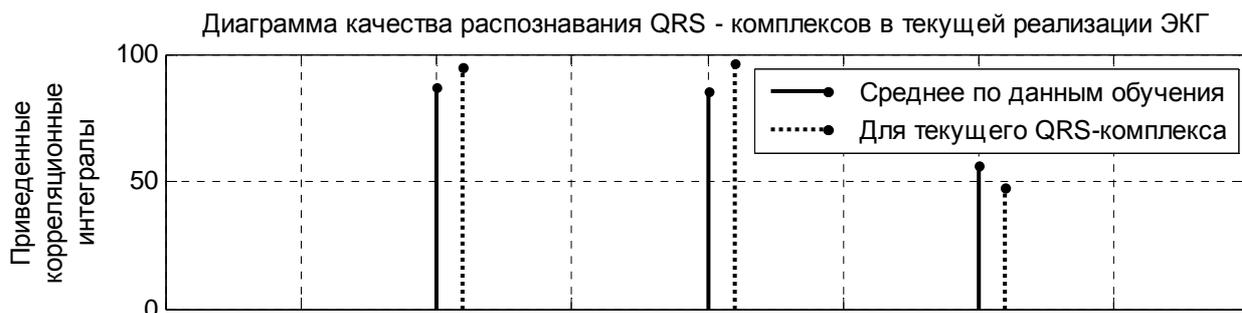


Рис. 6. Качество распознавания  $QRS$  - комплексов в текущей реализации ЭКГ

При этом для каждого из трех возможных решений используются нормированные значения корреляционного интеграла, приведенные к шкале  $[0;1]$ . Пересчет интеграла выполняется по формуле  $z \in [-1;1] \rightarrow z^* \in [0;1] = (1/2) \cdot (z + 1)$ , где  $z$  - корреляционный интеграл, вычисляемый обычным образом,  $z^*$  - корреляционный интеграл, значения которого приведе-

ны к шкале  $[0;1]$ . Если сравниваются противоположные по знаку сигналы одной и той же формы и энергии, то  $z = -1$ , а  $z^* = 0$ ; для ортогональных векторов  $z = 0$ ,  $z^* = 1/2$ ; при полном совпадении сигналов по форме  $z = 1$ ,  $z^* = 1$ .

Операции приведенного перечня упорядочены в соответствующем алгоритме, реализованном в среде MATLAB. В алгоритм внесен ряд дополнительных операций, предлагаемых ниже к рассмотрению в числе результатов данной работы.

### Дополнительные решения для повышение результативности распознавания сигналов

Введение дополнительных операций в процедуры анализа сигналов направлено на повышение результативности распознающих алгоритмов. Такие операции связаны с оптимизацией положения и размера окон анализа сигналов и с выбором состава элементов сопоставляемых портретов.

Все вопросы рассматриваются в рамках принятия решений по характеристикам формы сигналов. Приведенные иллюстрации направлены на подтверждение выигрыша качества распознавания сигналов за счет введения предлагаемых дополнительных операций. Ниже раскрывается их содержание и приводится иллюстрация возникающих преимуществ.

**1. Оптимизация положения и размера окон анализа сигналов** – состоит в поиске положения начала и ширины окна анализа в окне наблюдения. С изменением этих параметров выполнялся сбор статистики правильных и ошибочных решений и оценивалась правильность решений по комплексам по всем трем типам одновременно – с учетом их разного количества во всей реализации. В каждом случае рассчитывалась взвешенная сумма долей правильных решений по типам комплексов:  $\bar{p} = n_N / (n_N + n_A + n_V) p_N^+ + n_A / (n_N + n_A + n_V) p_A^+ + n_V / (n_N + n_A + n_V) p_V^+$ , где  $n_N$ ,  $n_A$ ,  $n_V$  – количество комплексов типов  $N$ ,  $A$  и  $V$  в обучающей выборке,  $p_N^+$ ,  $p_A^+$ ,  $p_V^+$  – доля правильных решений. Рассматривался также критерий оценки правильности решений по всей их совокупности:  $\bar{p}_1 = (p_N^+ + p_A^+ + p_V^+) / 3$ . В качестве обучающей выборки вначале рассматривалась вся исходная запись кардиограммы. Полученный результат о положении и размере окна отражен на рис. 7.

Окно отмечено вертикальными линиями на рисунке. При использовании такого окна как окна принятия решений вместо всего исходного окна анализа  $QRS$  - комплексов правильность их распознавания в среднем возросла – с 76%, 86% и 90% для  $N$ ,  $A$  и  $V$  - типов соответственно до 80%, 90% и 87%.

**2. Выбор рационального состава элементов сопоставляемых портретов** основывался на алгоритме последовательной отбраковки отсчетов характеристики  $\vec{e}_x$  для формы комплексов на этапе обучения. Вначале из всего исходного окна анализа формы сигнала удалялся отсчет эталона, худший по информативности, затем среди оставшихся определялся следующий и т.д.

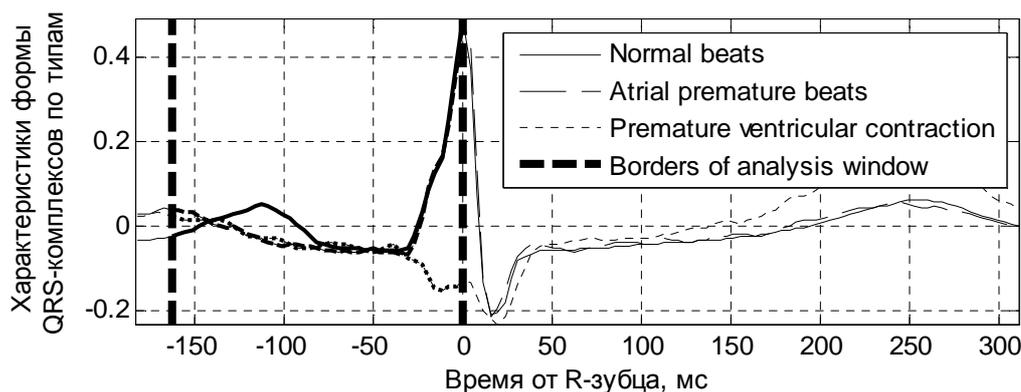


Рис. 7. Оптимальное окно анализа сигналов по данным обучающей выборки

Критерием выбора худшего отсчета являлся наибольший рост правильности распознавания сигнала (по всем типам комплексов совместно) при исключении этого отсчета из всех, принимаемых во внимание. Локализацию исключаемых отсчетов в динамике характеризует рис. 8.

Найденный оптимальный состав точек для сравнения *QRS*-комплексов, выбранный алгоритмом из соображения обеспечения максимума достоверности решений, представлен на рис. 9.

Из рисунков следует, что в первую очередь «закрываются» временные позиции, на которых разные эталоны близки друг другу. Это позволяет внести в состав характеристик информативности отсчетов их различимость.

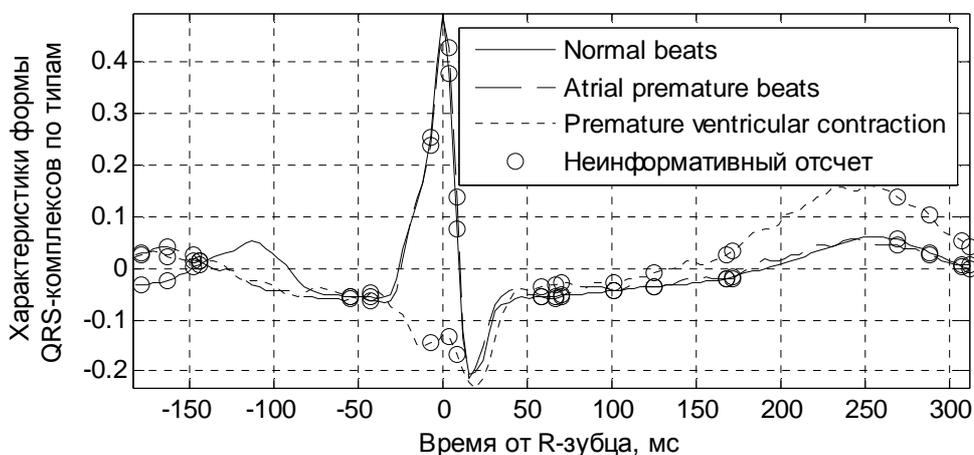


Рис. 8. Исключаемые отсчеты на начальной стадии их отбраковки

Рис. 10 отражает динамику изменения правильности решений на тестовой выборке по мере сокращения состава отсчетов в характеристике  $\bar{e}_x$  формы сигнала по предложенному алгоритму отбора признаков. График имеет характерные участки повышения результативности сокращения состава признаков в связи с исключением отсчетов, усложняющих распознавание, и, ближе к окончанию реализации процедуры, – утраты этой результативности вследствие того, что удаляются начинают существенные отсчеты (элементы формы сигналов).

Выигрыш в правильности распознавания *QRS* - комплексов в данном случае, как показала статистика, имеет тот же порядок, что и в предыдущем.

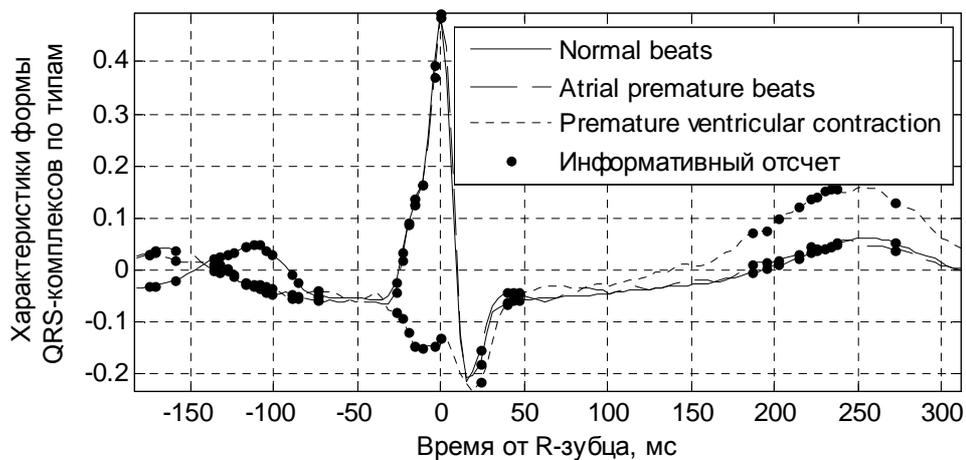


Рис. 9. Оптимальный состав информативных точек

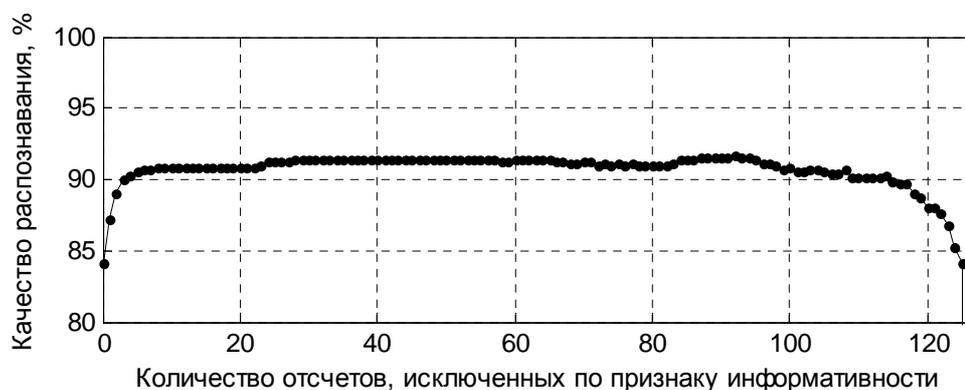


Рис. 10. Изменение правильности решений при отбраковке отсчетов сигнала

Приведенные иллюстрации подтверждают состоятельность рассмотренной схемы анализа структуры сигналов в целях их обнаружения, распознавания и классификации и целесообразность дальнейшего углубленного системного анализа обнаруженных и представленных в работе эффектов влияния преобразований сигналов на качество принимаемых решений.

### Выводы

1. Представлен тестовый вариант обработки на реальных данных электрокардиографии принципов анализа структуры циклических медико-биологических сигналов для обеспечения их обнаружения, распознавания и классификации. В целом подтверждены состоятельность базовой модели и предлагаемой последовательности преобразований сигналов при их анализе, что указывает на целесообразность проведения полноформатного обучения рассмотренного анализатора с получением оценок достигаемой результативности принятия решений по используемым критериям.

2. Применительно к анализу ЭКГ показано, что описания формы сигналов, выделенные на основе представленной базовой модели, для таких сигналов ха-

