
НОВЫЕ ПОДХОДЫ К АНАЛИЗУ И ИНТЕРПРЕТАЦИИ ФОРМЫ ЦИКЛИЧЕСКИХ СИГНАЛОВ

Аннотация. Предложены новые методы извлечения локализованной диагностической информации из циклических сигналов сложной формы. Показаны преимущества альтернативного метода оценки формы усредненного цикла, основанного на переходе от скалярного сигнала к его отображению на фазовой плоскости. Предложены оригинальные методы оценки динамики параметров, характеризующих форму информативных фрагментов сигнала, основанные на построении выпуклой оболочки фазового портрета перестановочной энтропии и расстоянии Левенштейна.

ВВЕДЕНИЕ

Повторяющиеся во времени процессы часто протекают в технических [1] и биологических [2] объектах. Циклические явления изучают также в экономике [3], термодинамике [4] и других областях.

Специфические сигналы, порождаемые таким процессами, обычно имеют сложную форму, анализ которой позволяет косвенно получать информацию о состоянии объекта [5]. Исследованию свойств циклических сигналов и построению адекватных математических моделей их порождения посвящены многие научные публикации, в частности, [6-10].

Диагностическая информация часто сосредоточена на небольших временных интервалах $\Delta t_i \ll T_0$, $\forall i = 1, 2, \dots$, области определения $T_0 = [t_1, t_2]$ сигнала $z(t)$, которые в реальных условиях искажены внутренними и внешними возмущениями [11]. Типичные примеры – электрокардиограммы (ЭКГ), реограммы, фотоплетизмограммы и другие физиологические сигналы, локальные фрагменты которых несут информацию о состоянии сердечно-сосудистой системы [12].

Классические методы ортогональных разложений

$$z(t) = \sum_{n=1}^N C_n \varphi_n(t) \quad (1)$$

по системе базисных функций $\varphi_n(t)$ мало пригодны для обработки сигналов с локализованной информацией, поскольку коэффициенты C_n функционального ряда (1), построенного на всей области $T_0 = [t_1, t_2]$, почти нечувствительны к изменениям параметров, которые характеризуют форму локальных фрагментов и обладают диагностической ценностью.

Несмотря на то, что существует целый ряд методов, ориентированных на извлечение локализованной диагностической информации, в частности, метод вейвлет-преобразований [13], по мнению специалистов существующие прикладные информационные технологии (ИТ) не обеспечивают требуемую достоверность принятия решений. Поэтому остается актуальной проблема повышения эффективных извлечения диагностической информации из циклических сигналов, искаженных возмущениями.

Цель статьи – исследовать новые подходы к извлечению локализованной диагностической информации из циклических сигналов сложной формы и на примере обработки реальных сигналов продемонстрировать их эффективность.

ОЦЕНКА ПАРАМЕТРОВ ФОРМЫ УСРЕДНЕННОГО ЦИКЛА

Одна из задач компьютерной обработки циклических сигналов – оценка усредненного цикла, которая направлена на повышение соотношения сигнал-шум [14]. Усреднение циклов позволяет обнаружить скрытые возмущения изменения формы сигнала, которые несут диагностическую ценность. Например, метод SAECG (Signal Averaged ECG) позволяет обнаружить на ЭКГ поздние потенциалы (Late Potentials) [15], которые являются предиктором опасных аритмий и внезапной сердечной смерти.

Тривиальные процедуры усреднения циклов во временной области, которые применяют в существующих цифровых системах [14], имеют ряд недостатков. Для того чтобы продемонстрировать это и показать преимущество предложенного подхода, воспользуемся моделью порождения ЭКГ реалистической формы [16], упростив которую будем полагать, что за счет внутренних возмущений варьирует форма только одного информативного фрагмента – зубца T . В этом случае процесс порождения последовательности циклов можно представить в виде

$$z_m(t) = \sum_{i \in \{P, Q, R, S, S-T, T\}} \tilde{A}_{im} \exp \left[-\frac{(t - \tilde{\mu}_{im})^2}{2\tilde{b}_{im}^2} \right], \quad m = 1, 2, \dots, \quad (2)$$

где

$$\tilde{A}_{im} = \begin{cases} A_i(1 + \alpha_{im}), & \text{если } i = T, \\ A_i, & \text{если } i \neq T, \end{cases} \quad (3)$$

$$\tilde{\mu}_{im} = \begin{cases} \mu_i(1 + \delta_{im}), & \text{если } i = T, \\ \mu_i, & \text{если } i \neq T, \end{cases} \quad (4)$$

$$\tilde{b}_{im} = \begin{cases} b_i^{(1)}(1 + \varepsilon_{im}^{(1)}) \quad \forall t \leq \mu_{im}, & \text{если } i = T, \\ b_i^{(2)}(1 + \varepsilon_{im}^{(2)}) \quad \forall t > \mu_{im}, & \text{если } i = T, \\ b_i^{(1)} \quad \forall t \leq \mu_{im}, & \text{если } i \neq T, \\ b_i^{(2)} \quad \forall t > \mu_{im}, & \text{если } i \neq T. \end{cases} \quad (5)$$

Здесь A_i , μ_i , $b_i^{(1)}$, $b_i^{(2)}$ – параметры, определяющие форму информативных фрагментов (зубцов P, Q, R, S, T и сегмента $S-T$) *ненаблюдаемого* полезного сигнала $z_0(t)$ в виде *несимметричных* гауссовых функций, а α_{im} , δ_{im} , $\varepsilon_{im}^{(1)}$, $\varepsilon_{im}^{(2)}$ – реализации независимых случайных величин, которые с нулевыми математическими ожиданиями распределены на интервалах

$$\alpha_{im} \in [-\alpha_i^0, \alpha_i^0], \quad \delta_{im} \in [-\delta_i^0, \delta_i^0], \quad \varepsilon_{im}^{(1)} \in [-\varepsilon_i^0, \varepsilon_i^0], \quad \varepsilon_{im}^{(2)} \in [-\varepsilon_i^0, \varepsilon_i^0], \quad (5)$$

ограниченных фиксированными числами α_i^0 , δ_i^0 , ε_i^0 , определяющими уровень искажений формы зубца T при порождении циклов.

На рис. 1 представлен результат усреднения во временной области пяти циклов $z_1(t)$, ..., $z_5(t)$ (рис. 1, а), порожденных моделью (2)-(5), которые синхронизированы по максимуму зубцов R . Несинхронность появления зубцов T за счет возмущения δ_{im} приводит к «размыванию» формы зубца на усредненном цикле $\hat{z}_0(t)$ (рис. 1, б) и, как следствие, к ошибкам в определении диагностических признаков, характеризующих форму зубца T .

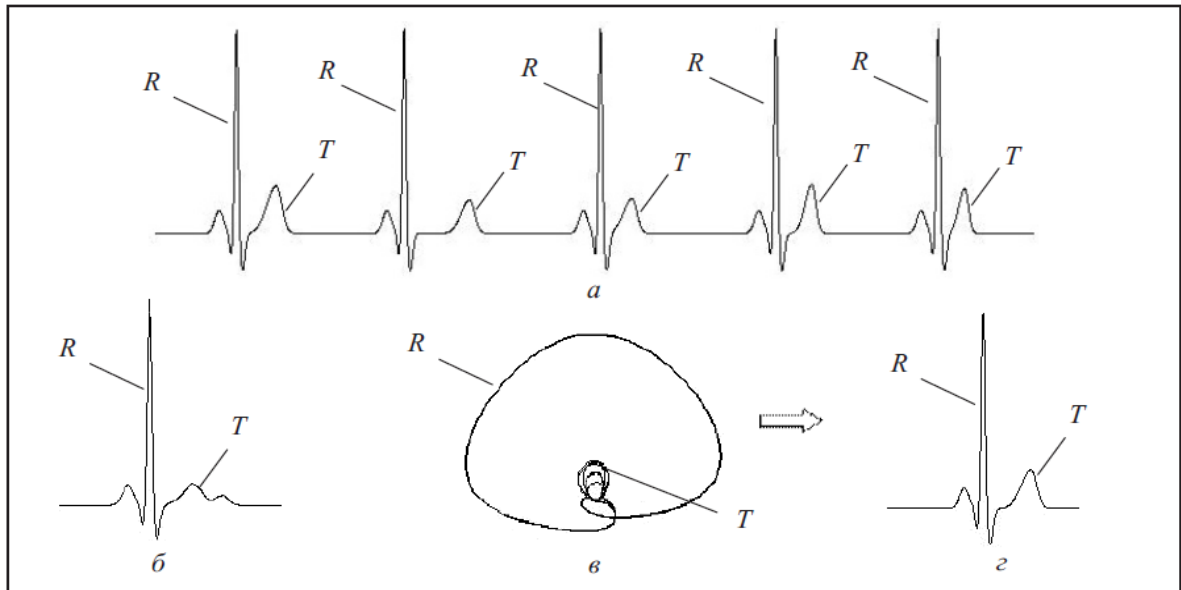


Рис. 1. Результаты усреднения циклов ЭКГ

Понятно, что при синхронизации этих же циклов по максимуму зубцов T на усредненном цикле $\hat{z}_0(t)$ будет искажена уже форма других фрагментов – зубцов P, Q, R, S .

Казалось бы, описанную проблему можно преодолеть, если синхронизировать *интервалы*, на которых сосредоточены отдельные фрагменты. Однако в реальных ситуациях зубцы P, Q, R, S, T не имеют четких границ, а значит не так уж просто синхронизировать указанные интервалы.

Рассмотрим альтернативный подход к усреднению циклов [12], который сводится к переходу от скалярного сигнала $z(t)$ к его отображению на плоскости с фазовыми координатами $z(t), \dot{z}(t)$, где $\dot{z}(t)$ – первая производная сигнала. Поскольку дифференцирование зашумленных функций относится к числу некорректно поставленных математических задач, то для практической реализации метода привлекаются специальные вычислительные процедуры [5].

Очевидно, что производная $\dot{x}(t)$ любой дифференцируемой функции $x(t)$ *инвариантна* относительно сдвига аргумента t , т.е. $\forall \tau = \text{const}$ выполняется условие $\dot{x}(t) \equiv \dot{x}(t - \tau)$. Поэтому, не смотря на действие возмущений δ_{im} , переход к годографу $z(t), \dot{z}(t)$ позволяет *автоматически* синхронизировать зубцы T на различных циклах.

Каждому m -му циклу соответствует своя траектория на фазовой плоскости $z(t), \dot{z}(t)$, которая за счет действия случайных возмущений $\alpha_{im}, \varepsilon_{im}^{(1)}, \varepsilon_{im}^{(2)}$ отклоняется от ненаблюдаемой траектории полезного сигнала $z_0(\cdot)$. Однако, поскольку предполагается, что $\alpha_{im}, \varepsilon_{im}^{(1)}, \varepsilon_{im}^{(2)}$ независимы и имеют нулевые математические ожидания, то для любой фиксированной точки $\theta = \theta_x$ области определения полезного сигнала $z_0(\cdot)$ координаты $z_m^*[\theta_{mx}], \dot{z}_m^*[\theta_{mx}]$ соответствующих точек $\theta_{1x}, \dots, \theta_{Mx}$ наблюдаемого сигнала будут группироваться в локальной области нормированного фазового пространства с центром в точке $Z = z_0^*[\theta_x], \dot{z}_0^*[\theta_x]$ и ограниченным радиусом ρ_Z [5].

Отображение сигнала на фазовой плоскости $z(t), \dot{z}(t)$ *принципиально* отличает предложенный метод от других известных подходов [17], предусматривающих отображение сигнала на псевдофазовой плоскости с координатах $z(t), z(t - \tau)$, где τ – задержка во времени. Именно такое отличие позволило *впервые* реализовать процедуру

надежного определения дополнительного диагностического признака ЭКГ (показателя β_T), характеризующего симметрию фрагмента реполяризации усредненной фазовой траектории.

Для эффективной оценки усредненного цикла $z_0(t)$ требуется исключить фазовые траектории, порожденные нетипичными циклами (например, экстрасистолами). Особенность этой задачи состоит в том, что форма нетипичного цикла одного сигнала может быть типична для другого и наоборот, а значит невозможно построить процедуру селекции нетипичных циклов на основе аналитического описания их формы.

Предлагаемый метод основан на единственном предположении: число нетипичных циклов M_N значительно меньше общего числа M обрабатываемых циклов, т.е. $M_N \ll M$ (в противном случае определение «типичный цикл» теряет смысл).

Вычислим $M \times M$ матрицу хаусдорфовых расстояний [18]

$$R_H(Q_i, Q_j) = \max \left\{ \max_{q_k \in Q_i} \min_{q_r \in Q_j} \rho(q_k, q_r), \max_{q_r \in Q_j} \min_{q_k \in Q_i} \rho(q_k, q_r) \right\} \quad (6)$$

между всеми парами Q_i и Q_j , $i = 1, \dots, M$, $j = 1, \dots, M$ фазовых траекторий наблюдаемых циклов, где $\rho(q_k, q_r) = \|q_k - q_r\|$ – евклидово расстояние между нормированными векторами $q_k = (z_k, \dot{z}_k) \in Q_i$, $k = 1, \dots, K$ и $q_r = (z_r, \dot{z}_r) \in Q_j$, $r = 1, \dots, R$ соответственно i -го и j -го циклов, представленных K и R дискретными отсчетами.

Номер строки этой матрицы, сумма элементов которой минимальна, определит траекторию

$$Q_0 = \arg \min_{1 \leq i \leq M} \sum_{j=1}^M R_H(Q_i, Q_j), \quad (7)$$

которую будем называть *опорной траекторией* в нормированном фазовом пространстве.

Упорядочим по возрастанию элементы строки матрицы хаусдорфовых расстояний, которая соответствует Q_0 . Полученная последовательность расстояний

$$\mathfrak{R} = R_H(Q_0, Q_\mu), \quad \mu \in [1, M] \quad (8)$$

дает представление о вариабельности формы траекторий наблюдаемых сигналов по отношению к Q_0 . Если наблюдаемый сигнал содержит *только* типичные циклы, то расстояния $R_H(Q_0, Q_\mu)$ возрастают более-менее плавно с ростом μ . Если же на сигнале появляются нетипичные циклы, форма которых существенно отличается от Q_0 , то последовательность (8) содержит один или несколько выраженных скачков. Определив положение первого из таких скачков, можно найти пороговое значение R_H^0 и тем самым выделить подмножество «надежных» траекторий

$$\Omega_N = \{Q_\mu : R_H(Q_0, Q_\mu) < R_H^0\}, \quad (9)$$

подлежащих усреднению.

Модельные эксперименты показали, что усреднение траекторий (рис. 1, с) в фазовом пространстве с последующим возвращением во временную область (рис. 1, d) обеспечивает приемлемую для практики точность восстановления ненаблюдаемого полезного сигнала $z_0(\cdot)$, а характерная форма фазовых траекторий (рис. 2) позволяет упростить процедуру интерпретации ЭКГ.

Клинические испытания подтвердили, что оценка показателя β повышает чувствительность и специфичность ЭКГ-диагностики даже в тех случаях, когда анализ ЭКГ в 12 традиционных отведениях оказывается неинформативным [19].

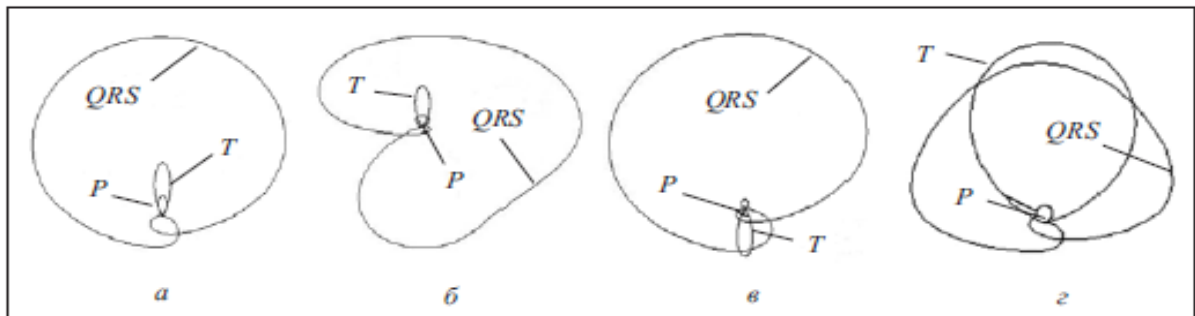


Рис. 2. Форма фазовой траектории усредненного цикла ЭКГ: нормальный цикл (а), патологический зубец Q (б), отрицательный зубец T (в), коронарный зубец T (г)

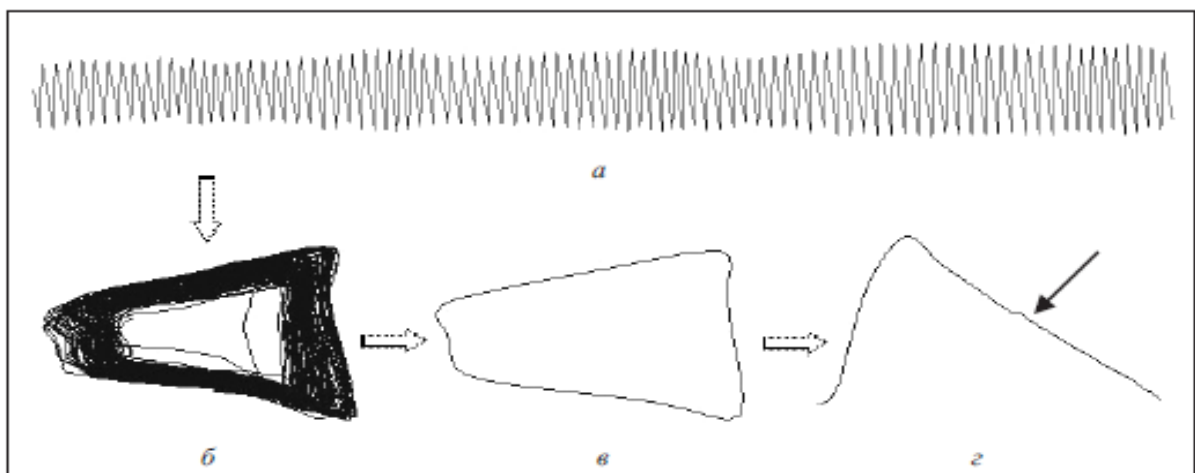


Рис. 3. Результат усреднения пальцевой фотоплетизмограммы: реальная фотоплетизмограмма (а), фазовые траектории фотоплетизмограммы (б), усредненная фазовая траектория (в), пульсовая волна во временной области с фрагментом отраженной волны (г)

Предложенный подход позволяет также провести эффективное усреднение циклов пальцевой фотоплетизмограммы (рис. 3) и распознать на усредненном сигнале характерный фрагмент (отмечен стрелкой), соответствующий моменту появления отраженной волны. Такая информация дает возможность интегрально оценивать жесткость артериальных сосудов по скорости распространения пульсовой волны [20].

ФАЗОВЫЙ ПОРТРЕТ ПЕРЕСТАНОВОЧНОЙ ЭНТРОПИИ

Диагностическую ценность несет не только форма усредненного цикла, но и изменчивость этой формы в процессе регистрации сигнала [21]. Анализ хаотичности $R-R$ -интервалов давно уже получил известность в кардиологической практике [22]. Наши исследования показали [23], что энтропийные оценки хаотичности параметров, характеризующих форму фрагментов ЭКГ, позволяют получать важную дополнительную диагностическую информацию.

В основе многих математических методов исследования хаотичности лежит известное выражение шенноновской энтропии

$$H = -\sum_{i=1}^N p_i \log_2 p_i, \quad (10)$$

которая количественно оценивающая неопределенность системы, находящейся в одном из N состояний с вероятностями $p_i, i = 1, \dots, N$.

Для оценки хаотичности циклического сигнала будем анализировать последовательность параметров, характеризующих форму отдельных циклов, например, последовательность значений показателя β , отражающих динамику симметрий зубцов T :

$$\beta_1, \beta_2, \dots, \beta_M, \quad (11)$$

Казалось бы для анализа хаотичности последовательности (11) можно прямо воспользоваться формулой (10), оценив вероятности (частоты) попадания значений β в заданные интервалы. Однако это не так! Легко показать, что величина (10) *инвариантна* относительно перестановок элементов ряда (11). Поэтому при непосредственном использовании формулы (10) фактически будет оцениваться не хаотичность последовательности (11), а неопределенность случайной величины β , порождающей эту последовательность, что не одно и то же.

Будем сканировать последовательность (11) скользящим окном, содержащим $W \ll M$ точек и в каждом l -м окне, $l=1, \dots, L$, оценим процентное приращении энтропии H_l по отношению к энтропии H_1 в первом окне, т.е. вычислим отношение

$$h_l = \frac{H_l}{H_1} \cdot 100\%, \quad l=1, \dots, L, \quad (12)$$

полагая, что $H_1 \neq 0$.

Принципиально для вычисления H_l можно воспользоваться любыми известными энтропийными оценками, подробный анализ которых приведен в работе [23], в частности модернизированной перестановочной энтропией. В этом случае отношение (12) приобретает вид

$$h_l = \frac{-\sum_{j=1}^5 p(\pi_{j_l}) \log p(\pi_{j_l})}{H_1} \cdot 100\%, \quad l=1, \dots, L, \quad (13)$$

где $p(\pi_j)$ – частота появления одного из пяти характерных паттернов π_1, \dots, π_5 (рис. 4) в l -м окне, а

$$h_l = -\sum_{j=1}^5 p(\pi_{j_l}) \log p(\pi_{j_l}) \quad (14)$$

– перестановочная энтропия, вычисленная в первом окне в предположении, что $h_1 \neq 0$.

Классы паттернов однозначно определяются по трем последовательным значениям $\beta_{i-1}, \beta_i, \beta_{i+1}$, $i=2, \dots, M-1$ в пределах l -го окна, $l=1, \dots, L$ следующим образом:

класс π_1 , если $(\beta_i - \beta_{i-1}) > \varepsilon_0 \wedge (\beta_i - \beta_{i+1}) > \varepsilon_0$,

класс π_2 , если $(\beta_{i-1} - \beta_i) > \varepsilon_0 \wedge (\beta_{i+1} - \beta_i) > \varepsilon_0$,

класс π_3 , если $(\beta_i - \beta_{i-1}) > \varepsilon_0 \vee (\beta_{i+1} - \beta_i) > \varepsilon_0 \vee (\beta_{i+1} - \beta_{i-1}) > \varepsilon_0$,

класс π_4 , если $(\beta_{i-1} - \beta_i) > \varepsilon_0 \vee (\beta_i - \beta_{i+1}) > \varepsilon_0 \vee (\beta_{i-1} - \beta_{i+1}) > \varepsilon_0$,

класс π_5 , если не выполняется ни одно из приведенных соотношений,

где ε_0 – заданный порог нечувствительности к локальным изменениям сигнала.

Надлежащим выбором порога ε_0 и ширины W скользящего окна можно добиться плавности изменения функции (13) и, применив процедуру численного дифференцирования, оценить производную $\dot{h}(l)$ в l -й точке. Это позволило перейти от ряда дискретных значений скользящей энтропии $h(l)$ к фазовому портрету на плоскости

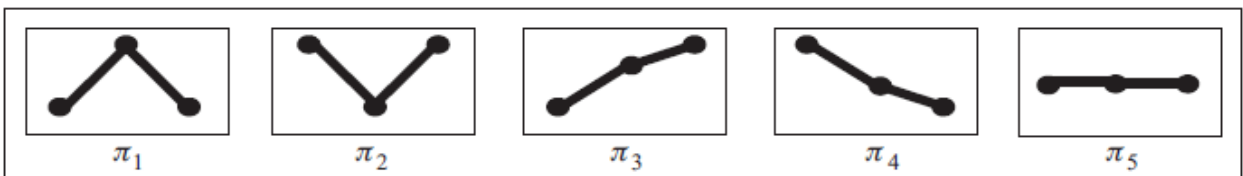


Рис. 4. Классы паттернов модернизированной перестановочной энтропии

$h(l), \dot{h}(l)$ и построить его выпуклую оболочку (рис. 5).

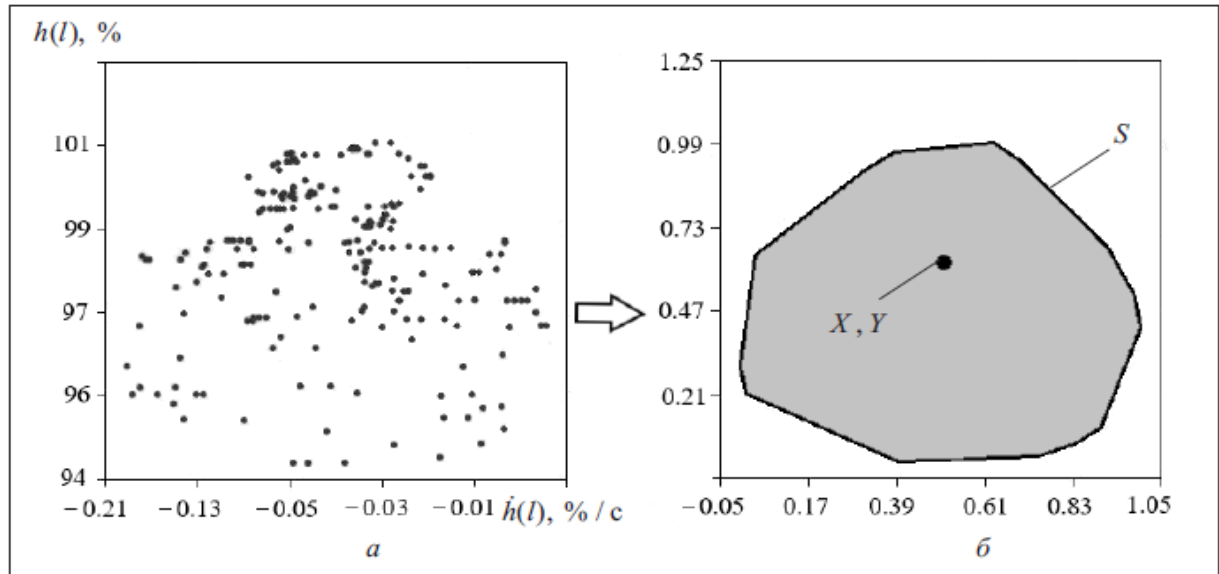


Рис. 5. Фазовый портрет скользящей энтропии (а) и его выпуклая оболочка (б)

Для интегральной оценки хаотичности β предлагается использовать площадь S выпуклой оболочки, координаты X, Y центра тяжести и ряд других показателей.

Аналогичным образом оценивают интегральные показатели, характеризующие хаотичность других параметров формы циклического сигнала, в частности амплитуд и продолжительностей зубцов P, Q, R, S, T электрокардиограммы.

Следует отметить, что энтропия, в отличие от дисперсии, характеризует не столько разброс, сколько *разнообразие* значений случайной величины, и поэтому позволяет более полно анализировать поведение сложных медико-биологических систем.

Например, при обработке ЭКГ больного, на которой наблюдался эффект электрической альтернации сердечного ритма (предиктор внезапной сердечной смерти) площадь S выпуклой оболочки фазового портрета перестановочной энтропии была почти в 4 раза меньше, чем у здорового пациента, в то время как в обоих случаях традиционные показатели (среднеквадратические отклонения RR -интервалов) практически совпадали [24].

Фазовый портрет перестановочной энтропии оказался полезным также при анализе тонких изменений ЭКГ в процессе капельного введения лекарственных препаратов и подтвердил статистически значимые отличия при обработке ЭКГ здоровых людей с разным уровнем адаптационных возможностей.

ИНТЕРПРЕТАЦИЯ ЦИКЛИЧЕСКИХ СИГНАЛОВ НА ОСНОВЕ РАССТОЯНИЯ ЛЕВЕНШТЕЙНА

Анализ динамики изменения формы циклических сигналов можно проводить в рамках лингвистического подхода, основанного на переходе от k -й реализации $z_k(t)$ сигнала, $k = 1, 2, \dots$, к слову $S_k = \alpha_1 \alpha_2 \dots \alpha_K$, представляющему собой *конечную* цепочку *символов* $\alpha_j \in \Psi, j = 1, \dots, K$ некоторого алфавита Ψ «названий» фрагментов. В общем случае лингвистический подход предполагает [25], что множество $\{S_k\}$ всевозможных слов (не обязательно конечное) образует формальный язык, для которого строится *грамматика* в виде четверки

$$G = \langle \Omega_0, \Omega_T, P_G, \omega_0 \rangle, \quad (15)$$

где Ω_0 – множество *нетерминальных* символов (переменных), Ω_T – множество терминальных символов (констант), $\Omega_T \cup \Omega_0 = \Psi$, $\Omega_T \cap \Omega_0 = \emptyset$, P_G – множество правил подстановки, а $\omega_0 \in \Omega_0$ – *начальный* (корневой) нетерминальный символ.

В большинстве работ, посвященных лингвистическому анализу сигналов, предполагается, что алфавит Ψ *заранее известен*, а конструирование грамматик, адекватных множеству наблюдаемых сигналов, осуществляется человеком на основании *неформальных* знаний эксперта в предметной области.

Следуя [26, 27], будем осуществлять переход от наблюдаемого сигнала $z_k(t)$ к слову $S_k = \alpha_1 \alpha_2 \dots \alpha_K$, анализируя разности значений показателей на смежных циклах. Для каждого параметра, характеризующего форму фрагментов циклического сигнала, введем индикаторную переменную, которая принимает значения в зависимости от знака приращений показателя i -го цикла по отношению к $i-1$ -му, $i = 2, \dots, M$. Для последовательности (11) индикаторную переменную можно представить в виде

$$V_i^{(\beta)} = \begin{cases} +1, & \text{если } \beta_i - \beta_{i-1} \geq 0, \\ -1, & \text{если } \beta_i - \beta_{i-1} < 0, \end{cases} \quad i = 2, \dots, M. \quad (16)$$

Аналогичным образом вводится другие индикаторные переменные, в частности переменная $V_i^{(RR)}$, значения которой определяют знаки приращений продолжительностей циклов ($R-R$ -интервалов). В этом случае каждый цикл будет кодировать один из символов алфавита $\Psi = \{a, b, c, d\}$ так, как показано в таблице 1.

Таблица 1. Принцип кодирования цикла ЭКГ

Значение индикаторной переменной $V_i^{(\beta)}$	+1	-1	+1	-1
Значение индикаторной переменной $V_i^{(RR)}$	+1	+1	-1	-1
Символ	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>

Последовательность символов, полученных в соответствии с таблицей 1, формирует M -разрядное слово S_k , однозначно кодирующее k -й сигнал, содержащий M циклов.

Предлагаемый метод основан на вычислении расстояния Левенштейна $L(S_1, S_2)$ между двумя словами (кодограммами) S_1, S_2 , состоящими из M_1 и M_2 символов соответственно. Напомним, что расстояние Левенштейна равно минимальному числу операций редактирования типа вставки, удаления и замены символа для преобразования слова S_1 в слово S_2 [28], а один из известных алгоритмов вычисления расстояния $L(S_1, S_2)$ состоит в следующем.

Формируем матрицу D размерностью $M_1 + 1, M_2 + 1$, первые строку и столбец которой заполняем так:

$$\begin{aligned} D(i, 0) &= i, \quad i = 0, \dots, M_1, \\ D(0, j) &= j, \quad j = 0, \dots, M_2. \end{aligned} \quad (17)$$

Остальные элементы $D(i > 0, j > 0)$ последовательно заполняем в соответствии с правилом:

$$D(i, j) = \min \{ D(i, j-1) + 1, D(i-1, j) + 1, D(i-1, j-1) + m(S_1(i), S_2(j)) \}, \quad (18)$$

где

$$m(S_1(i), S_2(j)) = \begin{cases} 0, & \text{если } S_1(i) = S_2(j), \\ 1, & \text{если } S_1(i) \neq S_2(j). \end{cases} \quad (19)$$

В результате расстояние Левенштейна $L(S_1, S_2)$ определяет число $D(M_1, M_2)$.

Будем принимать решение о принадлежности циклического сигнала $z(t)$ одной из двух групп Ω_1 и Ω_2 на основе сравнения расстояний Левенштейна между кодограммой S_t обрабатываемого сигнала и эталонными кодограммами $S_0^{(1)}$ и $S_0^{(2)}$, построенными по обучающим выборкам сигналов в указанных группах, т.е.

$$z(t) \in \Omega_1, \text{ если } L(S_t, S_0^{(1)}) \leq L(S_t, S_0^{(2)}), \quad (20)$$

$$z(t) \in \Omega_2, \text{ если } L(S_t, S_0^{(1)}) > L(S_t, S_0^{(2)}). \quad (21)$$

Алгоритм построения эталонных кодограмм $S_0^{(1)}$ и $S_0^{(2)}$ состоит в следующем. Пусть, например, в результате клинических исследований зарегистрировано Q_1 и Q_2 ЭКГ представителей двух групп: группе Ω_1 больных и группе Ω_2 условно здоровых пациентов.

Каждую из Q_1 ЭКГ первой группы кодируем словами $S_q^{(1)}$, $q=1, \dots, Q_1$ в соответствии с таблицей 1. Воспользовавшись процедурами (17)-(19), определим расстояния Левенштейна $L_{\mu\nu}(S_\mu^{(1)}, S_\nu^{(1)})$ между каждой парой $S_\mu^{(1)}, S_\nu^{(1)}$, $\mu=1, \dots, Q_1$, $\nu=1, \dots, Q_1$ кодограмм и сформируют квадратную $Q_1 \times Q_1$ матрицу расстояний $L_{\mu\nu}(S_\mu^{(1)}, S_\nu^{(1)})$, $\mu=1, \dots, Q_1$, $\nu=1, \dots, Q_1$ между всеми парами кодограмм первой группы:

$$\Lambda = \begin{pmatrix} L_{11}, & L_{12}, & \dots, & L_{1Q_1} \\ L_{21}, & L_{22}, & \dots, & L_{2Q_1} \\ & & \dots & \\ L_{Q_1 1}, & L_{Q_1 2}, & \dots, & L_{Q_1 Q_1} \end{pmatrix}. \quad (22)$$

Тогда эталонную кодограмму $S_0^{(1)}$ первой группы определит строка матрицы (22), сумма элементов которой минимальна, т.е.

$$S_0^{(1)} = \arg \min_{1 \leq \mu \leq Q_1} \sum_{\nu=1}^{Q_1} L_{\mu\nu}. \quad (23)$$

Эталонное слово $S_0^{(2)}$ второй группы определяют аналогичным образом на основе анализа суммы элементов матрицы расстояний Левенштейна $L_{\mu\nu}(S_\mu^{(2)}, S_\nu^{(2)})$, $\mu=1, \dots, Q_2$, $\nu=1, \dots, Q_2$, построенной для всех пар кодограмм второй группы, т.е.

$$S_0^{(2)} = \arg \min_{1 \leq \mu \leq Q_2} \sum_{\nu=1}^{Q_2} L_{\mu\nu}. \quad (24)$$

Исследование диагностических возможностей предложенного метода проводилось на основе обработки базы реальных ЭКГ, зарегистрированных в отделении ишемических болезней сердца НИИ кардиологии им. В.Д. Стражеско АМН Украины (Киев) и четырех клиник Германии – Essen University Hospital (Essen), Katholical Hospital "Phillpusstift" (Essen), Heart and Diabetes Center of North Rhein-Weasfalia (Bad-Oeynhausien), German Heart Center (Berlin).

Клинический материал составлял $Q_1 = 100$ записей ЭКГ больных хронической формы ишемической болезни сердца (ИБС), диагноз которым был предварительно установлен по результатам коронарографии (группа Ω_1), и $Q_2 = 100$ записей ЭКГ здоровых добровольцев, включенных в контрольную группу (группа Ω_2).

Эталонные кодограммы, построенные для указанных групп, имели вид

$$S_0^{(1)} = \text{adc b d a d c a d a b d a b c a d a b d a d c b d a b},$$

$$S_0^{(2)} = \text{cbcdcabdcaabddcaadcaa}.$$

На основе обработки имеющихся данных установлено, что принятия решения по правилам (20), (21) обеспечивает чувствительность $S_E = 72\%$ и специфичность $C_p = 79\%$.

Важно отметить, что на ЭКГ верифицированных больных отсутствовали традиционные признаки ишемии миокарда (элевация либо депрессия ST сегмента). А это значит, что традиционная электрокардиография отнесла бы *все* обрабатываемые ЭКГ (в том числе, ЭКГ больных), к группе здоровых. В то же время на основе сравнения расстояний Левенштейна $L(S_t^{(2)}, S_0^{(1)})$ и $L(S_t^{(2)}, S_0^{(2)})$ удалось провести приемлемую классификацию представителей групп Ω_1 и Ω_2 на таком «сложном» для диагностики клиническом материале.

Анализ расстояний Левенштейна между кодограммами оказался также полезным для наглядной иллюстрации динамики изменения ЭКГ при медикаментозном лечения сердечно-сосудистых патологий и для решения других задач.

Предложенный подход может быть естественным образом обобщен на случаи, когда в формировании кодограмм ЭКГ используют не только продолжительностей $R-R$ - интервалов и значения показателя β_T , но и другие информативные показатели, в частности, амплитуды зубцов R и T . Этот подход можно также использовать для анализа и интерпретации других циклических сигналов.

ЗАКЛЮЧЕНИЕ

Рассмотрены оригинальные методы компьютерной обработки циклических сигналов сложной формы, направленные на извлечения локализованной диагностической информации в условиях возмущений. Показано, что переход от скалярного сигнала к его отображению на фазовой плоскости позволяет повысить точность определения усредненного цикла.

Фазовый портрет ЭКГ позволил также впервые реализовать процедуру надежного определения дополнительного диагностического признака – показателя, характеризующего симметрию фрагмента реполяризации, который обеспечил повышение чувствительности и специфичность ЭКГ-диагностики даже в тех случаях, когда традиционный анализ оказывается неинформативным

Предложены оригинальные методы оценки динамики параметров, характеризующих форму информативных фрагментов сигнала, которые основаны на построении выпуклой оболочки фазового портрета перестановочной энтропии и расстояния Левенштейна. Клинические испытания подтвердили диагностическую ценность этих методов при обнаружении эффекта электрической альтернации сердечного ритма, капельном введении лекарственных препаратов и ряде других медицинских приложений.

СПИСОК ЛИТЕРАТУРЫ

1. Membhard A.D., Sinha J.K., Pinkerton A.J., Elbhah K. Fault diagnosis of rotating machines using vibration and bearing temperature measurements. *Diagnostyka*. 2013. Vol. 14, No. 3. P. 45-51.
2. Frank D.W., Evans J.A., Gorman M.R. Time-Dependent Effects of Dim Light at Night on Re-Entrainment and Masking of Hamster Activity Rhythms. *Journal of Biological Rhythms*. 2010. Vol. 25, No. 2. P. 103–112. DOI:[10.1177/0748730409360890](https://doi.org/10.1177/0748730409360890).
3. Benkemoune R. Charles Dunoyer and the Emergence of the Idea of an Economic Cycle. *History of Political Economy*. 2009. Vol. 41, No. 2. P. 271-295. <https://doi.org/10.1215/00182702-2009-003>.
4. Wu C. Thermodynamic Cycles: Computer-Aided Design and Optimization. New York : Marcel Dekker, 2003. 312 p.

5. Файнзильберг Л.С. Информационные технологии обработки сигналов сложной формы. Теория и практика. Киев: Наукова Думка, 2008. 333 с.
6. Лупенко С.А. Детерминированные и случайные циклические функции как модели колебательных явлений и сигналов: определение и классификация. *Электронное моделирование*. 2006. Т. 28, № 4. С. 29–45.
7. Драган Я.П. Математичне й алгоритмічно-програмне забезпечення комп'ютерних засобів статистичного опрацювання стохастичних коливань (ритмічних процесів). Вісник Нац. ун-ту «Львівська політехніка»: Інформаційні системи та мережі. 2008. № 621. С. 124–130.
8. Zvarich V. N., Marchenko V. G. Linear autoregressive processes with periodic structures as models of information signals. *Radioelectronics and Communications Systems*. 2011. Vol. 54, No. 7. P. 367–372.
9. Шачиков А.Д., Шуляк А.П. Отработка принципов анализа структуры циклических медико-биологических сигналов для их обнаружения, распознавания и классификации. *Вісник НТУУ «КПІ». Серія Приладобудування*. 2015. Вип. 49, № 1. С. 169-179.
10. Lytvynenko I.V. The problem of segmentation of the cyclic random process with a segmental structure and the approaches to its solving. *Journal of Hydrocarbon Power Engineering*. 2016, Vol. 3, Issue 1. P. 30-37.
11. Поворознюк А.И., Филатова А.Е Проектирование нелинейного фильтра в задаче структурной идентификации биомедицинских сигналов с локально сосредоточенными признаками. *Системні дослідження та інформаційні технології*. 2014. № 1. С. 69-80.
12. Файнзильберг Л.С. Компьютерная диагностика по фазовому портрету электрокардиограммы. Киев: Освита України, 2013. 191 с.
13. Bruns A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *Journal of Neuroscience Methods*. 2004. Vol 137, Issue 2. P. 321-332. <https://doi.org/10.1016/j.jneumeth.2004.03.002>.
14. Zywiencz S., Borovsky D., Goetsch G., Joseph G. Methodology of ECG Interpretation in the Hanover Program. *Methods of Information in Medicine*. 1990. No. 29. P. 375.
15. Liao Y.C., Lin Y.J., Chung F.P et al. Risk stratification of arrhythmogenic right ventricular cardiomyopathy based in signal averaged ECGs. *Int. Journal of cardiology*. 2014. Vol. 174, No. 3. P. 628-633. doi: 10.1016/j.ijcard.2014.04.169
16. Файнзильберг Л.С. Имитационные модели порождения искусственных электрокардиограмм в условиях внутренних и внешних возмущений. *Journal of Qafqaz University. Mathematics and Computer Science*. 2012. № 34. С. 92-104.
17. Фрумин Л.Л., Штарк М.Б. О фазовом портрете электрокардиограммы. *Автоматрия*. 1993. № 2. С. 51-54.
18. Fainzilberg L.S. Restoration of a Standard Sample of Cyclic Waveforms with the Use of the Hausdorff Metric in a Phase Space. *Cybernetics and Systems Analysis*. 2003. Vol. 39, Issue 3. P. 338-344.
19. Файнзильберг Л.С. Основы фазографии. Киев: Освита України, 2017. 264 с.
20. Jayanthi T., Anburajan M. Pulse wave velocity and its usefulness in the estimation of hypertension. *Asian Journal of Pharmaceutical and Clinical Research*. 2017. Vol 10, Issue 4 P. 181-187. DOI: 10.22159/ajpcr.2017.v10i4.16447.
21. Файнзильберг Л.С., Минаева Е.Н. Оценка функционального состояния сердечно-сосудистой системы по величине разброса фазовых траекторий одноканальной ЭКГ. *Кибернетика и вычислительная техника*. 2014. Вып. 175. С. 5-19, 88.
22. Wesfraid E., Billat V. Randomness and changes of heart rate and respiratory frequency during high altitude mountain ascent without acclimatization. *Physica A: Statistical Mechanics and its Applications*. 2012. Vol 391, Issue 4. P. 1575-1590. <https://doi.org/10.1016/j.physa.2011.08.067>

23. Файнзильберг Л.С., Ориховская К.Б., Ваховский И.В. Оценка хаотичности формы фрагментов одноканальной ЭКГ. *Кибернетика и вычислительная техника*. 2016. Вып. 183. С. 4-24.
24. Fainzilberg L., Orikhovska K., Vakhovskyi I. Analysis of Subtle Changes in Biomedical Signals Based on Entropy Phase Portrait. *Биомедицинская инженерия и электроника*. 2017. № 3. С. 44-66.
25. Trahanias P., Skordalakis E. Syntactic pattern recognition of the ECG. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990. Vol. 12, Issue 7. P. 648-657. DOI: [10.1109/34.56207](https://doi.org/10.1109/34.56207)
26. Uspenskiy V.M. Diagnostic System Based on the Information Analysis of Electrocardiogram. *Proceedings of Mediterranean Conference on Embedded Computing (MECO 2012)*. Montenegro (June 19–21). 2012. P. 74–76.
27. Колеснікова О.В., Кривенко С.С. Інформаційний аналіз електрокардіосигналів: обґрунтування і можливості. *Збірник наукових праць Першої Міжнародної науково-практичної конференції «Інформаційні системи та технології в медицині» (ISM–2018)*. Харків: ХНУРЕ, 2018. С. 161-163.
28. Wagner R.A., Fischer M.J. The String-to-String Correction Problem. *Journal of the ACM*. 1975. Vol. 22, Issue 2. P. 177-183. DOI: [10.1145/321879.321880](https://doi.org/10.1145/321879.321880)

Надійшла до редакції 26.06.2019

Л.С. Файнзильберг

НОВІ ПІДХОДИ ДО АНАЛІЗУ ТА ІНТЕРПРЕТАЦІЇ ФОРМИ ЦИКЛІЧНИХ СИГНАЛІВ

Анотація. Запропоновано нові методи вилучення локалізованої діагностичної інформації з циклічних сигналів складної форми. Показано переваги альтернативного методу оцінювання форми усередненого циклу, який базується на переході від скалярного сигналу до його відображення на фазовій площині. Запропоновано оригінальні методи оцінювання динаміки параметрів, що характеризують форму інформативних фрагментів сигналу, основані на побудові опуклої оболонки фазового портрета переставної ентропії та відстані Левенштейна.

Ключові слова: циклічний сигнал, фазовий портрет, переставна ентропія, відстань Левенштейна.

L.S. Fainzilberg

NEW APPROACHES FOR ANALYSIS AND INTERPRETATION OF THE CYCLIC SIGNAL'S SHAPE

Abstract. New methods for extracting localized diagnostic information from cyclic signals of complex shape have been proposed. The advantages of an alternative method for estimating the shape of an averaged cycle based on the transition from a scalar signal to its mapping on the phase plane are shown. Original methods for estimating the dynamics of parameters characterizing the shape of informative signal fragments based on the construction of the convex hull of the phase portrait of the permutation entropy and the Levenshtein distance are proposed.

Keywords: cyclic signal, phase portrait, permutation entropy, Levenshtein distance.

Файнзильберг Леонид Соломонович,

доктор техн. наук, профессор, главный научный сотрудник Международного научно-учебного центра информационных технологий и систем НАН и МОН Украины, Киев, e-mail: fainzilberg@gmail.com.