NEW APPROACHES TO THE ANALYSIS
AND INTERPRETATION OF THE SHAPE
OF CYCLIC SIGNALS

L. S. Fainzilberg

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

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

INTRODUCTION

Processes repeating in time often occur in engineering [1] and biological [2] objects. Cyclic phenomena are also analyzed in economy [3], thermodynamics [4], and other fields.

Specific signals generated by such processes are normally of complicated shape whose analysis allows obtaining indirectly the information about the state of the object [5]. Properties of cyclic signals and creation of adequate mathematical models of their generation are analyzed in many scientific publications, in particular, in [6–10].

Diagnostic information is often concentrated on small time intervals $\Delta t_i \ll T_0 \forall i = 1, 2, \ldots$ of the definition domain $T_0 = [t_1, t_2]$ of signal $z(t)$, which are distorted by internal and external perturbations in actual practice [11]. Typical examples are electrocardiograms (ECG), rheograms, photoplethysmograms, and other physiological signals whose local fragments contain information about the state of cardiovascular system [12].

Classical methods of orthogonal decompositions

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

in the system of basis functions $\varphi_n(t)$ are hardly suitable for processing of signals with localized information since coefficients $C_n$ of the functional series (1) constructed on the entire domain $T_0 = [t_1, t_2]$ are almost insensitive to variations in parameters that characterize the shape of local fragments and are important for diagnostics.

In spite of the fact that methods oriented to retrieval of localized diagnostic information such as the method of wavelet-based transformations [13] are known, according to experts, available applied information technologies (IT) do not ensure required reliability of decisions being made. Therefore, increasing the efficiency of retrieval of diagnostic information from cyclic signals distorted by perturbations is important.

The purpose of the paper is to investigate new approaches to retrieval of a localized diagnostic information from cyclic signals of complex shape and to show their efficiency on the example of processing of real signals.


1060-0396/20/5604-0665 ©2020 Springer Science+Business Media, LLC 665
ESTIMATING THE PARAMETERS OF THE SHAPE OF AN AVERAGED CYCLE

One of the problems of computer processing of cyclic signals is estimation of an averaged cycle to increase the signal–noise ratio [14]. Cycle averaging allows revealing the variations in signal’s shape, hidden by perturbations, which are important for diagnostics. For example, the SAECG (Signal Averaged ECG) method allows detecting late potentials on ECG [15], which are predictors of dangerous arrhythmia and sudden cardiac death.

Trivial procedures of cycle averaging in a time domain, which are applied in available digital systems [14], have a number of shortcomings. To demonstrate this and to show an advantage of the proposed approach, we will use the model of an ECG of realistic shape [16] and simplify it by assuming that due to internal perturbations the shape of only one informative fragment, \( T \), varies. In this case, we can present the process of generating a sequence of cycles as

\[
\mathbf{z}_m(t) = \sum_{i \in \{P, Q, R, S, T\}} \mathbf{A}_{im} \exp \left[ -\frac{(t - \mu_{im})^2}{2\tilde{b}^2_{im}} \right], \quad m = 1, 2, \ldots, \tag{2}
\]

where

\[
\mathbf{A}_{im} = \begin{cases} \mathbf{A}_i(1 + \alpha_{im}) & \text{if } i = T, \\ \mathbf{A}_i & \text{if } i \neq T, \end{cases} \tag{3}
\]

\[
\mu_{im} = \begin{cases} \mu_i(1 + \delta_{im}) & \text{if } i = T, \\ \mu_i & \text{if } i \neq T, \end{cases} \tag{4}
\]

\[
\tilde{b}_{im} = \begin{cases} b_i^{(1)}(1 + \epsilon_{im}^{(1)}) & \forall t \leq \mu_{im} \text{ if } i = T, \\ b_i^{(1)} & \forall t > \mu_{im} \text{ if } i \neq T, \end{cases} \tag{5}
\]

Here, \( \mathbf{A}_i, \mu_i, b_i^{(1)}, \) and \( b_i^{(2)} \) are parameters that determine the shape of informative fragments \( (P, Q, R, S, T, \) and segment \( S - T) \) of an unobserved useful signal \( z_0(t) \) in the form of asymmetric Gaussian functions; \( \alpha_{im}, \delta_{im}, \epsilon_{im}^{(1)}, \) and \( \epsilon_{im}^{(2)} \) are realizations of independent random variables, which are distributed, with zero expectations, on the intervals

\[
\alpha_{im} \in [-\alpha_i^0, \alpha_i^0], \quad \delta_{im} \in [-\delta_i^0, \delta_i^0], \quad \epsilon_{im}^{(1)} \in [-\epsilon_i^0, \epsilon_i^0], \quad \epsilon_{im}^{(2)} \in [-\epsilon_i^0, \epsilon_i^0],
\]

bounded by the fixed numbers \( \alpha_i^0, \delta_i^0, \) and \( \epsilon_i^0, \) which determine the level of distortions of the shape of \( T \) in cycle generation.

Figure 1 shows the result of averaging, in the time domain, of five cycles \( z_1(t), \ldots, z_5(t) \) (see Fig. 1a), generated by the model (2)–(5), which are synchronized with respect to the maximum of teeth \( R \). Non-synchronous occurrence of \( T \) due to perturbation \( \delta_{im} \) fuzzifies the tooth shape on the averaged cycle \( \tilde{z}_0(t) \) (see Fig. 1b) and hence causes errors in determining the diagnostic signs that characterize the shape of tooth \( T \).

It is clear that synchronization of the same cycles with respect to the maximum of teeth \( T \) on the averaged cycle \( \tilde{z}_0(t) \) distorts the shape of other fragments: teeth \( P, Q, R, \) and \( S \). The issue described above can probably be overcome by synchronizing the intervals on which separate fragments are concentrated. However, in real situations, teeth \( P, Q, R, \) \( S, \) and \( T \) have no clear boundaries; therefore, it is not easy to synchronize these intervals.

Let us consider an alternative approach to cycle averaging [12], which reduces to transition from a scalar signal \( z(t) \) to its mapping onto the plane with phase coordinates \( \bar{z}(t) \) and \( \dot{z}(t) \), which is the first derivative of the signal. Since differentiation of noisy functions is an ill-posed mathematical problem, special computing procedures [5] are used for practical implementation of the method.

It is obvious that derivative \( \dot{z}(t) \) of any differentiable function \( z(t) \) is invariant with respect to shift of the argument \( t \), i.e., \( \forall \tau = \text{const the condition } \ddot{z}(t) = \ddot{z}(t - \tau) \) is satisfied. Therefore, despite the perturbations \( \delta_{im} \), passage to hodograph \( z(t), \dot{z}(t) \) synchronizes teeth \( T \) automatically on different cycles.
Fig. 1. Results of averaging of ECG cycles.

To each \( m \)th cycle there corresponds certain trajectory on the phase plane \( z(t), \dot{z}(t) \), which, due to random perturbations of \( \epsilon_{im}, \epsilon_{im}^{(1)}, \) and \( \epsilon_{im}^{(2)}, \) deviates the unobserved trajectory of the useful signal \( z_0(t) \). However, since it is assumed that \( \epsilon_{im}, \epsilon_{im}^{(1)}, \) and \( \epsilon_{im}^{(2)}, \) are independent and have zero expectations, for any fixed point \( \theta = \theta_k \) from the domain of definition of the useful signal \( z_0(t) \), coordinates \( z_n^{\ast}[\theta_{mx}], \vec{z}_n^{\ast}[\theta_{mx}] \) of respective points \( \theta_{1k}, \ldots, \theta_{Mk} \) of the signal being observed will be grouped in a local domain of the normalized phase space with center at the point \( Z = z_n^{\ast}[\theta_{mx}], \vec{z}_n^{\ast}[\theta_{mx}] \) and bounded radius \( \rho_Z \) [5].

Signal mapping onto the phase plane \( z(t), \dot{z}(t) \) is a fundamental difference of the proposed method from other well-known approaches [17] that provide signal mapping onto a pseudo-phase plane with coordinates \( z(t), z(t-\tau) \), where \( \tau \) is a time lag. It is such difference that has allowed, for the first time, implementing the procedure of reliable determination of an additional diagnostic feature of an ECG (parameter \( \beta_T \)), which characterizes symmetry of the fragment of repolarization of an averaged phase trajectory.

For efficient estimate of the averaged cycle \( z_0(t) \), it is necessary to eliminate the phase trajectories generated by atypical cycles (for example, extrasystoles). A special feature of this problem is that the shape of an atypical cycle of one signal can be typical for another one and vice versa; therefore, it is impossible to construct the procedure of selection of atypical cycles on the basis of analytical description of their shape.

The proposed method is based on the unique assumption: the number of atypical cycles \( M_N \) is much less than the total number \( M \) of cycles being processed, i.e., \( M_N \ll M \) (otherwise, the definition of "typical cycle" becomes meaningless).

Let us calculate the \( M \times M \) matrix of Hausdorff distances [18]

\[
R_H(Q_i, Q_j) = \max \left\{ \max_{q_k \in Q_i} \min_{q_r \in Q_j} \rho(q_k, q_r), \max_{q_r \in Q_j} \min_{q_k \in Q_i} \rho(q_k, q_r) \right\}
\]

(6)

between all pairs \( Q_i \) and \( Q_j, i=1, \ldots, M, \) \( j=1, \ldots, M, \) of phase trajectories of cycles being observed, where \( \rho(q_k, q_r) = || q_k - q_r || \) is the Euclidean distance between the normalized vectors \( q_k = (z_k, \dot{z}_k) \in Q_i, k=1, \ldots, K, \) and \( q_r = (z_r, \dot{z}_r) \in Q_j, r=1, \ldots, R, \) respectively, of the \( i \)th and \( j \)th cycles presented by \( K \) and \( R \) discrete readings.

The number of the row of this matrix with the minimum sum of elements determines the trajectory

\[
Q_0 = \arg \min_{1 \leq i \leq M} \sum_{j=1}^{M} R_H(Q_i, Q_j).
\]

(7)

which we will call reference trajectory in the normalized phase space.

667
Fig. 2. The shape of phase trajectory of an averaged ECG cycle: normal cycle (a), pathological tooth Q (b), negative tooth T (c), coronary tooth T (d).

Fig. 3. The result of averaged finger photoplethysmogram: real photoplethysmogram (a), phase trajectories of the photoplethysmogram (b), averaged phase trajectory (c), pulse wave in the time domain with a fragment of the reflected wave (d).

Let us arrange in increasing order elements of the row of the matrix of the Hausdorff distances (6) corresponding to the reference trajectory (7). The obtained sequence of distances

$$R = R_H (Q_0, Q_\mu), \; \mu \in [1, M],$$

shows variability of the shape of trajectories of the observed signals with respect to $Q_0$. If the signal being observed contains only typical cycles, then distances $R_H (Q_0, Q_\mu)$ increase more or less smoothly with growth of $\mu$. If atypical cycles whose shape substantially differs from $Q_0$ occur on the signal, then the sequence (8) contains one or several expressed saltuses. Having determined the position of the first such saltus, it is possible to find the threshold value $R_H^0$ and thus to select a subset of "reliable" trajectories

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

subject to averaging.

Modeling experiments have shown that averaging of reliable trajectories (9) (see Fig. 1c) in the phase space with subsequent return to the time domain (see Fig. 1d) ensures practically acceptable accuracy of recovery of an unobserved useful signal $z_0 (\cdot)$, and the characteristic shape of phase trajectories (Fig. 2) simplifies the procedure of ECG interpretation.

Clinical trials have confirmed that estimate of parameter $\beta$ increases sensitivity and specificity of ECG-diagnostics even when ECG analysis in 12 traditional derivations appears non-informative [19].

The proposed approach allows also an efficient averaging of cycles of finger photoplethysmogram (Fig. 3) and recognition, on the averaged signal, of the characteristic fragment (marked by an arrow) that corresponds to the time of occurrence of the reflected wave. Such information allows integral estimation of the rigidity of arterial vessels based on the rate of propagation of pulse wave [20].
THE PHASE PORTRAIT OF PERMUTATION ENTROPY

Not only the shape of averaged cycle but also variability of this shape during signal’s registration [21] is important for diagnostics. Chaoticity analysis of $R-R$-intervals has been known in cardiological practice for a long time [22]. The studies have shown [23] that entropy estimates of chaoticity of parameters that characterize the shape of ECG fragments allow obtaining important additional diagnostic information.

Many mathematical methods of chaoticity analysis are based on the well-known expression of Shannon entropy

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

which quantitatively estimates uncertainty of the system that is in one of $N$ states with probabilities $p_i$, $i=1,...,N$.

To estimate chaoticity of a cyclic signal, let us analyze the sequence of parameters that characterize the shape of separate cycles, for example, the sequence of values of parameter $\beta$, which reflect the dynamics of symmetry of teeth $T$:

$$\beta_1, \beta_2, \ldots, \beta_M.$$  \hspace{1cm} (11)

For chaoticity analysis of sequence (11), we could probably use formula (10) and estimate the probabilities (frequencies) of values $\beta$ occurring on the given intervals. However, that is not the case! It can be easily shown that quantity (10) is invariant with respect to permutations of elements of the series (11). Therefore, in case of immediate use of formula (10), not the chaoticity of sequence (11) will be actually estimated but uncertainty of the random variable $\beta$ that generates this sequence, which is not identical.

Let us scan sequence (11) by a sliding window containing $W \ll M$ points, and in each $l$th window, $l=1,\ldots,L$, let us estimate the percentage increment of entropy $H_l$ with respect to entropy $H_1$ in the first window, i.e., let us calculate the ratio

$$h_l = \frac{H_l}{H_1} \cdot 100 \%, \ l=1,\ldots,L,$$  \hspace{1cm} (12)

on the assumption that $H_1 \neq 0$.

In essence, to estimate $H_l$, we can use any available entropy estimates whose detailed analysis can be found in [23], in particular, the upgraded permutation entropy. In this case, we can write the ratio (12) as

$$h_l = \sum_{j=1}^{5} \frac{p(\pi_j)}{h_1} \cdot 100 \%, \ l=1,\ldots,L,$$  \hspace{1cm} (13)

where $p(\pi_j)$ is the frequency of occurrence of one of the five characteristic patterns $\pi_1, \ldots, \pi_5$ (Fig. 4) in the $l$th window, and

$$h_1 = \sum_{j=1}^{5} p(\pi_{j1}) \log p(\pi_{j1})$$  \hspace{1cm} (14)

is the permutation entropy calculated in the first window on the supposition that the denominator (14) in ratio (13) is not equal to zero, i.e., $h_1 \neq 0$.

Classes of patterns can be determined uniquely based on three sequential values: $\beta_{i-1}$, $\beta_i$, and $\beta_{i+1}$, $i=2,\ldots,M-1$, within the limits of the $l$th window, $l=1,\ldots,L$, as follows:

- class $\pi_1$ if $(\beta_{i-1} - \beta_{i-1}) \geq \varepsilon_0 \land (\beta_{i-1} - \beta_{i+1}) \geq \varepsilon_0$, where $\varepsilon_0$ is a given threshold of insensitivity to local variations in the signal;
- class $\pi_2$ if $(\beta_{i-1} - \beta_{i-1}) > \varepsilon_0 \land (\beta_{i-1} - \beta_{i+1}) > \varepsilon_0$;
- class $\pi_3$ if $(\beta_{i-1} - \beta_{i-1}) > \varepsilon_0 \lor (\beta_{i+1} - \beta_{i+1}) > \varepsilon_0$;
- class $\pi_4$ if $(\beta_{i-1} - \beta_{i-1}) > \varepsilon_0 \lor (\beta_{i+1} - \beta_{i+1}) > \varepsilon_0$;
- class $\pi_5$ if none of the above relations hold.

By appropriately choosing the threshold $\varepsilon_0$ and width $W$ of the sliding window, we can attain smooth variation in the function (13) and, applying the procedure of numerical differentiation, we can estimate the derivative $h(l)$ at the $l$th point. This will allow us to pass from the series of discrete values of sliding entropy $h(l)$ to the phase portrait on the plane $h(l), h'(l)$ and to construct its convex hull (Fig. 5).

669
Fig. 4. Classes of patterns of the upgraded permutation entropy.

Fig. 5. The phase portrait of sliding entropy (a) and its convex hull (b).

For the integral estimate of chaoticity $\beta$, it is proposed to use the area $S$ of the convex hull, coordinates $X, Y$ of the barycenter, and other parameters.

Integral parameters that characterize chaoticity of other parameters of the shape of a cyclic signal, in particular, amplitudes and durations of ECG teeth $P, Q, R, S,$ and $T$, can be estimated similarly.

Note that entropy, unlike variance, characterizes not so much scatter but the variety of values of a random variable; therefore, it allows more full analysis of the behavior of complex medical and biological systems.

For example, in processing the ECG of a sick patient, where the effect of electrical alternation of cardiac rhythm (a predictor of sudden cardiac death) is observed, the area $S$ of the convex hull of the phase portrait of permutation entropy was almost four times less than that of a healthy person, while traditional parameters (root mean square deviations of $R-R$-intervals) almost coincided in both cases [24].

The phase portrait of permutation entropy also appeared useful in the analysis of almost unobservable but important changes in ECG during drip infusion of pharmaceutical drugs and confirmed statistically significant differences in processing ECG of healthy people with different levels of adaptation capabilities.

**INTERPRETATION OF CYCLIC SIGNALS BASED ON LEVENSHEIN DISTANCE**

The dynamics of variation in the shape of cyclic signals can be analyzed within the framework of the linguistic approach based on passage from the $k$th realization $z_k(t)$ of the signal, $k=1,2,\ldots,$ to a word $S_k=\alpha_1\alpha_2\cdots\alpha_K$, which is a finite chain of symbols $\alpha_j \in \Psi$, $j=1,\ldots,K$, of some alphabet $\Psi$ of "names" of fragments. Generally, the linguistic approach assumes [25] that the set $\{S_k\}$ (not necessarily finite) of all possible words forms a formal language for which the grammar is constructed in the form of a quadruple

$$G=(\Omega_0, \Omega_T, P_G, \omega_0),$$

where $\Omega_0$ is the set of non-terminal symbols (variables), $\Omega_T$ is the set of terminal symbols (constants), $\Omega_T \cup \Omega_0 = \Psi$, $\Omega_T \cap \Omega_0 = \emptyset$, $P_G$ is the set of substitution rules, and $\omega_0 \in \Omega_0$ is the initial (root) non-terminal symbol.
TABLE 1. ECG Cycle Encoding Principle

<table>
<thead>
<tr>
<th>Value of Indicator Variable $I_i^{(b)}$</th>
<th>+1</th>
<th>-1</th>
<th>-1</th>
<th>-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value of Indicator Variable $I_i^{(RR)}$</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Symbol</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

In the majority of studies devoted to the linguistic analysis of signals, it is assumed that the alphabet $\Psi$ is known in advance and that grammars (15) adequate to the set of signals being observed are constructed by experts on the basis of their informal knowledge about the subject domain.

Following [26, 27], let us pass from the signal being observed $z_k(t)$ to the word $S_k = a_1a_2\cdots a_K$, having analyzed the differences of the values of parameters on adjacent cycles. For each parameter characterizing the shape of fragments of the cyclic signal, let us introduce an indicator variable, which takes values depending on the sign of increments of the parameter of the $i$th cycle with respect to the $(i−1)$th one, $i = 2, \ldots, M$. For the sequence (11), we can present the indicator variable as

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

Other indicator variables can be introduced similarly, in particular, variable $V_i^{(RR)}$ whose values determine the signs of increments of the durations of cycles (R–R-intervals). In this case, according to (16), each cycle will be encoded by one of the alphabet symbols $\Psi = \{a, b, c, d\}$ as is shown in Table 1.

The sequence of symbols obtained according to Table 1 generates an $M$-bit word $S_k$, which uniquely encodes the $k$th signal containing $M$ cycles.

The method being proposed is based on calculating the Levenshtein distance $L(S_1, S_2)$ between two words (codegrams) $S_1$ and $S_2$, consisting of $M_1$ and $M_2$ symbols, respectively. Note that the Levenshtein distance is equal to the minimum number of editing operations such as insertion, deletion, and replacement of a symbol to transform word $S_1$ into word $S_2$ [28], and one of the well-known algorithms of calculating the distance $L(S_1, S_2)$ is as follows.

Generate $(M_1 + 1) \times (M_2 + 1)$ matrix $D$ and fill in its first row and column as follows:

$$D(i, 0) = i, \quad i = 0, \ldots, M_1,$$

$$D(0, j) = j, \quad j = 0, \ldots, M_2.$$

Sequently fill in the other elements $D(i > 0, j > 0)$ according to the rule

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

where

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

As a result, the Levenshtein distance $L(S_1, S_2)$ determines the number $D(M_1, M_2)$.

If cyclic signal $z(t)$ belongs to one of the two groups, $\Omega_1$ and $\Omega_2$, we will decide by comparing the Levenshtein distance between the codegram $S_i$ of the signal being processed and reference codegrams $S_0^{(1)}$ and $S_0^{(2)}$, constructed based on training samplings of signals in the specified groups, i.e.,

$$z(t) \in \Omega_1 \quad \text{if} \quad L(S_i, S_0^{(1)}) \leq L(S_i, S_0^{(2)}),$$

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

Algorithm of constructing reference codegrams $S_0^{(1)}$ and $S_0^{(2)}$ is as follows. Let, for example, as a result of clinical research, $\Omega_1$ and $\Omega_2$ ECG of representatives of two groups of patients be registered: $\Omega_1$ for patients and $\Omega_2$ for conventionally healthy people.
We will encode each of the $Q_1$ ECG of the first group by the words $S_q^{(1)}$, $q = 1, \ldots, Q_1$, according to Table 1. Let us use the procedures (17)–(19) and find the Levenshtein distances $L_{\mu\nu}(S_\mu^{(1)}, S_\nu^{(1)})$ between each pair $S_\mu^{(1)}$, $S_\nu^{(1)}$, $\mu = 1, \ldots, Q_1$, $\nu = 1, \ldots, Q_1$, of codegrams and generate a quadratic $Q_1 \times Q_1$ matrix of distances $L_{\mu\nu}(S_\mu^{(1)}, S_\nu^{(1)})$, $\mu = 1, \ldots, Q_1$, $\nu = 1, \ldots, Q_1$, between all the pairs of codegrams of the first group:

$$
\begin{bmatrix}
L_{11}, L_{12}, \ldots, L_{1Q_1} \\
L_{21}, L_{22}, \ldots, L_{2Q_1} \\
\vdots \\
L_{Q_11}, L_{Q_12}, \ldots, L_{Q_1Q_1}
\end{bmatrix}
$$

(22)

Then the reference codegram $S_0^{(1)}$ of the first group can be defined by the row of the matrix (22) whose sum of elements is minimum, i.e.,

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

(23)

The reference word $S_0^{(2)}$ of the second group can be found similarly, based on the analysis of the sum of elements of the matrix of Levenshtein distances $L_{\mu\nu}(S_\mu^{(2)}, S_\nu^{(2)})$, $\mu = 1, \ldots, Q_2$, $\nu = 1, \ldots, Q_2$, constructed for all the pairs of codegrams of the second group, i.e.,

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

(24)

Analysis of the diagnostic capabilities of the proposed method was based on the database of real ECG registered in the department of ischemic heart disease at the National Scientific Center M.D. Strazhesko Institute of Cardiology NAMS of Ukraine (Kyiv) and four German clinics: Essen University Hospital (Essen), Katholische Hospital "Philippusstift" (Essen), Heart and Diabetes Center of North Rhein-Westphalia (Bud-Einhauzen), and German Heart Center (Berlin).

The clinical material included $Q_1 = 100$ ECG entries of patients with the chronic form of ischemic heart disease (IHD), whose diagnosis was preliminary and based on the results of coronaryography (group $\Omega_1$), and $Q_2 = 100$ ECG entries of healthy volunteers included in the control group (group $\Omega_2$).

Reference codegrams (23) and (24) constructed for the above groups of patients had the form

$S_0^{(1)} = adcbadbacdabcabdbdcadah$,

$S_0^{(2)} = cbdcabdcabdbdcaadca$.

Based on the processing of the available data it has been established that decision-making by rules (20), (21) ensures the sensitivity $S_F = 72\%$ and specificity $C_F = 79\%$. Note that no traditional signs of myocardial ischemia (elevation or depression of ST-segment) were detected on the ECG of verified patients. And this means that according to traditional electrocardiography, all the ECG being processed (including, ECG of patients with diseases) can be referred to the group of healthy people. At the same time, on the basis of comparison of the Levenshtein distances $L(S_0^{(2)}, S_0^{(1)})$ and $L(S_1^{(2)}, S_0^{(1)})$, it became possible to carry out a comprehensible classification of representatives of groups $\Omega_1$ and $\Omega_2$ on such “diagnostically difficult” clinical material.

An analysis of the Levenshtein distances between codegrams also appeared useful for evident illustration of the dynamics of variations in ECG in case of medicamentous treatment of cardiovascular pathologies and for solution of other problems.

The proposed approach can be naturally generalized to the cases where not only durations of $R$–$R$-intervals and values of $\beta_T$ are used to generate ECG codegrams but also other informative parameters, in particular, amplitudes of teeth $R$ and $T$. This approach can also be applied for the analysis and interpretation of other cyclic signals.
CONCLUSIONS

We have considered original methods of computer processing of cyclic signals of complex shape, intended to retrieve a localized diagnostic information under perturbations. We have shown that passage from a scalar signal to its mapping on the phase plane increases the accuracy of determining the averaged cycle.

The phase portrait of ECG has allowed the first-time implementation of the procedure of reliable determination of an additional diagnostic indication: a parameter that characterizes symmetry of a repolarization fragment that ensures increase in sensitivity and specificity of ECG-diagnoses even when the traditional analysis appears non-informative.

We have also proposed original methods to estimate the dynamics of the parameters that characterize the shape of informative fragments of the signal, which are based on creating convex hull of the phase portrait of the permutation entropy and Levenshtein distance. Clinical trials have confirmed the value of these methods for diagnostics in detecting the effect of electrical alternation of cardiac rhythm, drip infusion of pharmaceutical drugs, and a number of other medical applications.

REFERENCES


673


