Method to reduce the effect of miagrafic and sensory noise with isolating the isoline on ECG signal

Evgeny Semenishchev[1,*], *Dmitry* Chernyshov[1], *Ilya* Svirin[2]

[1]Don state Technical University, 344000, Rostov-on-Don, Russian Federation
[2] CJSC Nordavind, 117545, Moscow, Russian Federation

Abstract: This paper deals with an approach to the analysis of ECG data, which allows to remove the noise component, to preserve the peaks characterizing the work of the heart, and to identificate floating isoline.

1 Introduction

Processes automation is one of the priorities of the modern world. Constant modernization of equipment requires large investments, as well as retraining of personnel, installation works and maintenance of new systems. In this connection, the most promising direction is the automation of data processing, which allows to reduce the number of personnel, and to improve the overall efficiency of the control system. One of the leading areas of data analysis is the development of the digital signals processing theory, based on the analysis of initial data.

The use of digital signal processing methods has found wide application [1]:
- in automation and control systems
- in modern antenna systems
- in the study of biomechanical parameters, biometric data collection systems located directly on the object under study
- in modern systems of computer vision and automatic processing of two-dimensional signals
- in economics and sociology in the study of trends;
- in information-measuring systems;
- in computer technology to increase accuracy of analog-digital transformation

Data, received with ECG, in addition to the useful information on the propagation of electrical impulses in the cardiac muscle, contains the noise component, which has a different source. Analog-digital convertion also adds random distortions to signal [2].

The main task in the signal processing is the separation of the useful component and attenuation of the noise component. In practice, a the mean-square error minimization criterion or the criterion of mean-absolute deviation is used to determine the processing quality. Each of this methods has advantages and limitations of use depending on the task and a priori information on the components of the input signal. In this connection, digital signal processing based on the objective function of the combined criteria becomes relevant objective. Of particular interest is the use of multicriterial methods for processing digital signals, that are represented by a single implementation with a limited amount of a priori information about the useful component function and statistical noise characteristics [3].

Figure 1 contains an algorithm for obtaining estimates using multi-criteria signal smoothing methods, based on minimizing the objective functions:

$$\varphi(\bar{s}_1, \bar{s}_2, ..., \bar{s}_n) = \alpha \sum_{k=1}^{n} (\bar{s}_k - Y_k)^2 + \sum_{k=1}^{n-2} (\bar{s}_k - 2 \cdot \bar{s}_{k+1} - \bar{s}_{k+2})^2$$

when α – adjusting factor.

When implementing the considered smoothing methods using machine modeling, in the form of programs, regions of values of the coefficients were obtained $0,01 \leq \alpha \leq 4.44$. ECG data processing requires real time analysis. It is done by finding estimates of the objective function in window k, and applying sliding window l for all values of the input signal. Here k is size of the window and l is step of sliding window.

The process of obtaining estimates in a sliding window is performed by parallel processing of the initial values of the multicriteria objective function in the window k, with various processing parameters α. The transition between estimates obtained with different parameters α is accomplished by the condition:

$$\bar{s}_k = \begin{cases} \bar{s}_k(\alpha_1) & (\bar{s}_k(\alpha_1) - \bar{s}_k(\alpha_2))^2 \leq p \\ \bar{s}_k(\alpha_2) & (\bar{s}_k(\alpha_1) - \bar{s}_k(\alpha_2))^2 > p \end{cases}$$

Here $\bar{s}_k(\alpha_1)$, $\bar{s}_k(\alpha_2)$ are input realization estimates, received with parameters

$\alpha_1(\sigma_{ou})\,\alpha_2(\sigma_{ou}),\mathrm{p}$ is a threshold value, determined experimentally with the RMS deviation of the additive noise component $\sigma_n < 0.2$ and equals $p = 0{,}15$ [3].
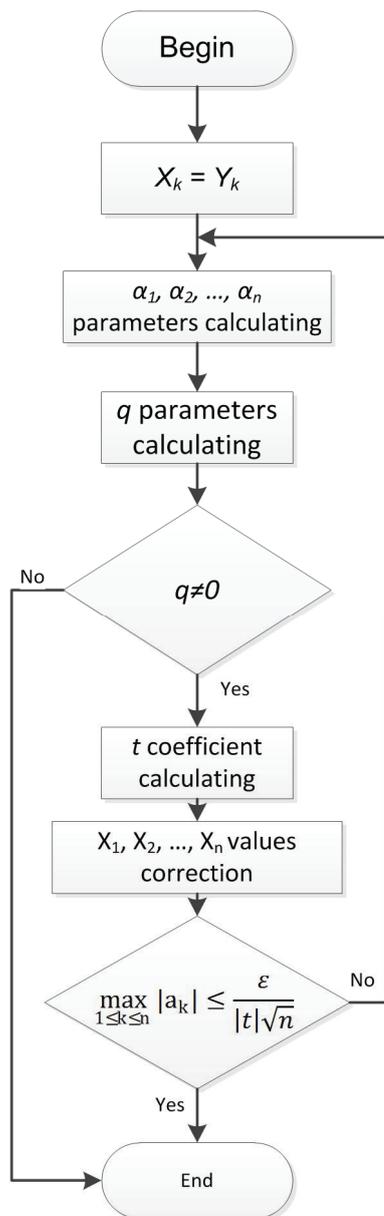
Fig. 1. Algorithm for obtaining estimates using multi-criteria signal smoothing methods.

The need to solve the isoline detection problem is due to the requirement to adjust the monitoring system to locate the report points recorded by the ECG device within the minimum error range of the analog-to-digital converter (ADC). To solve the problem of finding the midline, the algorithm presented in Figure 2 was developed.

Algorithm is shown on figure 2 and implemented as follows:

At the first step, the noise component, possibly with nonuniform distribution law is supressed. It is performed in order to exclude its influence on the middle line searching method.

In the second step, the R peak is suppressed using the R-R processing window.
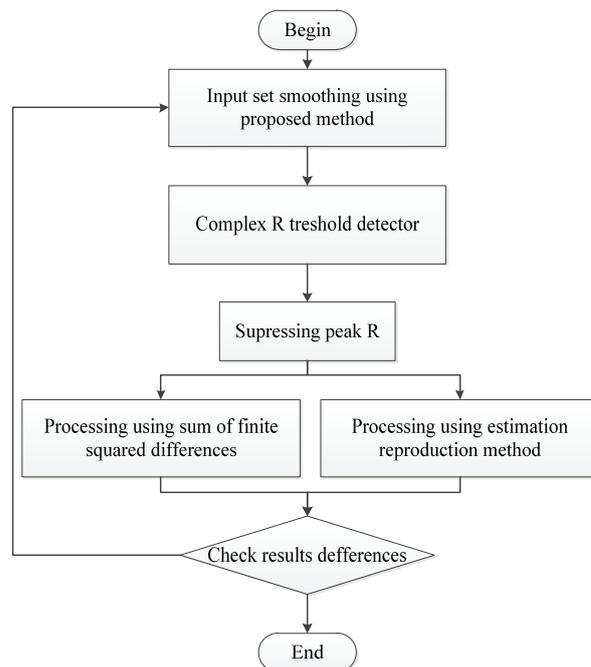
Fig. 2. Isoline searching algorithm.

At the final stage, parameters $\sigma_{ou}(\alpha)$ are fixed, and the entire set is processed. The choice of the filtering method parameter is done using the results shown in Figure 3.
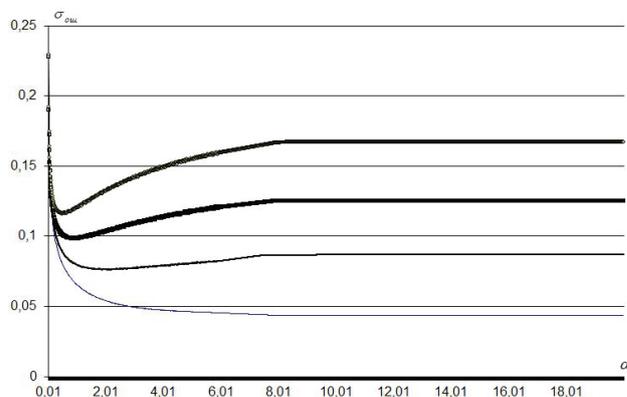
Fig. 3. A graph of the choice of the smoothing parameter for isoline selection method $\sigma_{ou}(\alpha)$.

Figure 3. - Example of the ECG isoline detection (1 - normalized ECG model, 2 - normalized ECG model with pathology, 3 - isoline).

The processing example shown in figure 3 is obtained with fixed parameters of the middle line search algorithm. The parameter of the processing window k is equal to the value of the R-R interval, the window step is fixed and is equal to half the processing window $l = k\,/\,2$, the filtering parameter $\alpha = 8$, i.e. the priority of the second summand of the objective function is eight times larger than the first.

2 Method of searching pathologies for ECG analysis

For detect pathologies by analyzing the electrocardiogram obtained by an automated mobile system and stored on a mobile device. We will develop a multi-criteria approach to the preliminary detection. Diagnosis by the means of the developed mathematical models of both a useful (ideal) ECG and possible pathologies.

As the first criteria, we will use the square measurement of the discrepancy between the developed mathematical model $s_{mat}(t_k) = s_{mat\,k}$ and the established cardiogram Y_k:

$$\left(s_{mat\,k} - Y_k\right)^2$$

As the second criteria, we will use a function representing the normalized correlation coefficient:

$$\frac{\sum(g_k - \bar{g}_k)(Y_k - \bar{Y}_k)}{\sqrt{\left(\sum(g_k - \bar{g}_k)^2\right)}\sqrt{\left(\sum(Y_k - \bar{Y}_k)^2\right)}}$$

where: g_k - model o.

Like a third criteria we will use resulting index, frequency area.

$$\sum_{i=0}^{n-1} Y_i e^{-\frac{2\pi j}{n}fft\_k*i}$$

Resulting multi-criteria function will have the form:

$$\psi(\bar{s}_1, \bar{s}_2, \dots \bar{s}_n) = \delta\left(s_{mat\,k} - Y_k\right)^2 +$$

$$+\gamma \frac{\sum\left(g_k - \sum\frac{g_n}{n}\right)\left(Y_k - \sum\frac{Y_n}{n}\right)}{\sqrt{\left(\sum\left(g_k - \sum\frac{g_n}{n}\right)^2\right)}\sqrt{\left(\sum\left(Y_k - \sum\frac{Y_n}{n}\right)^2\right)}} + \tau\left(\sum_{i=0}^{n-1} Y_i e^{-\frac{2\pi j}{n}fft\_k*i}\right)$$

here: δ, γ, τ - parameters that establish weight for each of the criteria, n - set size, fft_k - frequency index, j - imaginary unit.

Each parameters of the criteria are established during the preparation phase using the experience of specialists in ECG analysis.

Figure 4 shows the pathology search algorithm for ECG analysis.
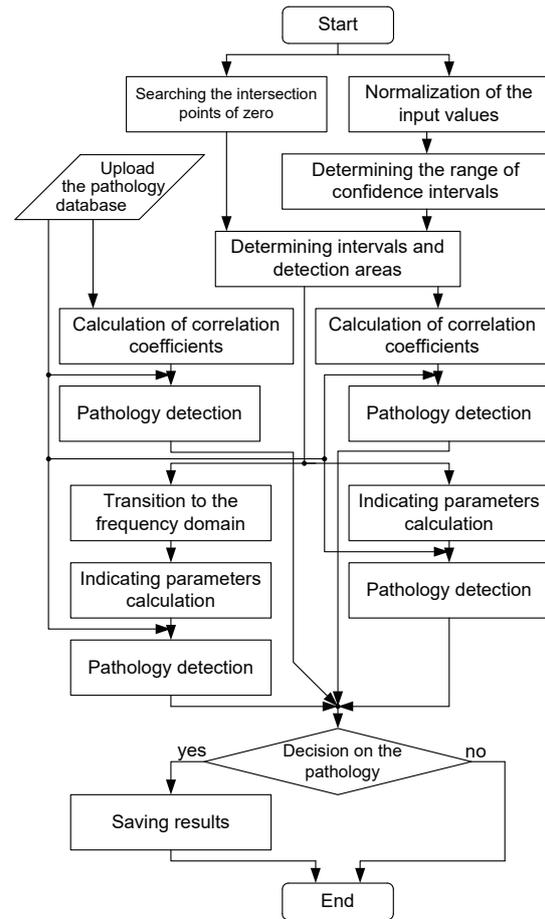
Fig. 4. Pathology search algorithm for ECG analysis.

The algorithm presented in Figure 4 is implemented as follows:

In the first step, the areas of intersection in established points of the cardiogram report with the isoline in the zones confined to confidence intervals are searched. At the same time, the input values are normalized in the range of R-R intervals.

In the next step, we make the determination of 5 peaks: P, Q, R, S, T and segments ST and R-R, the QRS complex. The determination is made according to the standard ranges [4] shown in Figure 5

In the third step, we use the ranges defined in the previous step to select intervals. The values for each selected range are normalized.

At the fifth step, the correlation coefficients are calculated for each pathology and for each examination area. Processing is performed in the sliding window. At the same time, we perform the transition to the frequency domain and the calculation of the main central moments for the normalized indicating parameters. The calculations of the central points are made using the analysis of the differential coefficient.
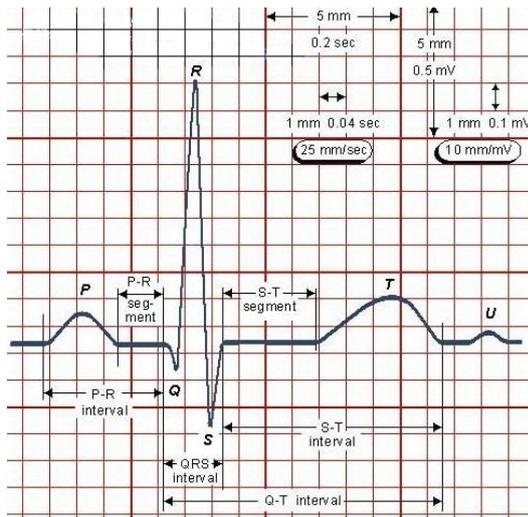
Fig. 5. The main complex of ECG.

On the sixth step the obtained frequency coefficients are compared with a mathematical model of each of the pathologies and the excess of a threshold value is detected. The calculation of the established threshold is also made by analyzing the divergence of the central moments and correlation coefficients. Pathology is considered detected if more than two indicators coincide and in case of repeated confirmation notification decision is made. In the case of coincidence of all criteria for ECG analysis and their coincidence with the investigated pathology, the patient is notified [5].

3 The implementation of the pathology search methods for the analysis of electrocardiograms

The first step is a searching for intersection of the fixed points of the cardiogram report with a baseline in that zones, which are limited by confidence intervals. We'll use a value of two quantization levels as confidence limit [6]. The device uses 10 bits analog-to-digital convertor or ADC and the range of standard QRS complex amplitudes for different areas of data sources should not exceed 22 mm and 25 mm for adults. This limit specifies the monitoring interval, which is limited in the range of 5mV. It means that the device indicates the intersection of the baseline when fixed values exceed the confidence intervals which are limited by two quantization levels and it's equal to ±0,02 mV. We will impose restrictions in the form of derivative of the change in direction for the obtained values. It will be used the criterion of exceeding the measurement sequence over the baseline for more than 3 reports as a simple pike detector. The restrictions on the excess of three reports are empirical, they are based on the minimum possible duration of restriction of the QRS complex. As a rule, this QRS complex lasts 0,08 seconds. For children under the age of 5 it lasts 0,09, and for children over 10 years – 0,10. When the QRS complex lasts more than 0,11 for adults, it is said of pathology and suggests ventricular hypertrophy or ventricular blockade. We get the minimum duration of QRS complex, which doesn't

exceed 500 reports at a sampling frequency of the device limited to 5 kHz. This article introduced a restriction of the range of fixed changes, which is equal to 100 reports. It will be made a decision about the intersection of the baseline during these reports.

We will search for intervals of consecution of the pikes R-R according to the following procedure: we will use the complex approach in which the maximum value will be determined as the main condition for detecting the R wave, as secondary and confirming conditions we will also use the relative time and sequence of intersection intervals of the baselines corresponding to P and Q and the following ST pikes, as well as their relative intervals of consecution. We can see the ECG duration model at the pic.6. Along with the R-R pikes, it is also possible to detect the remaining pikes and their ranges according to the data, which was obtained in the previous stages.
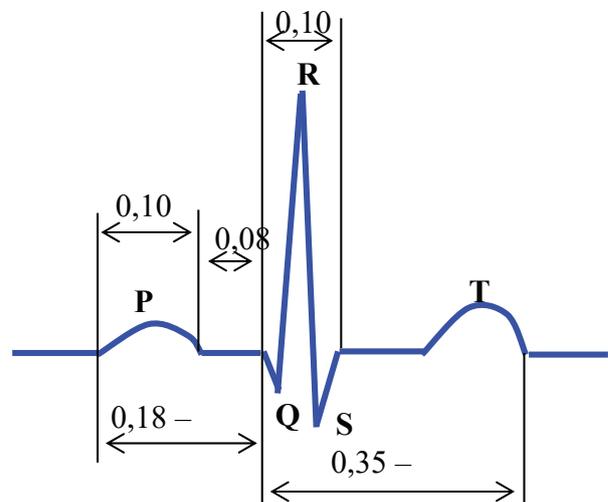
Fig. 6. The duration model ECG within normal limits.

The input values are normalized in the range of P-P intervals in the next step. The normalization process consists of fixing the maximum values at a given interval and dividing all values by a given value. Normalization of the values allows not to consider the scaling factor for the subsequent analysis and detection of pathologies. At the third stage, we produce the allocation of pike and segment intervals. We produce normalization of the values in the area of each obtained intervals. We define the group of central moments, such as mathematical expectation:

$$v_k = M(Y_k) = \sum Y_k \Big/ k$$

dispersion

$$D(Y_k) = M\left[(Y_k - M(Y_k))^2\right]$$

skewness ratio

$$\gamma_a(Y_k) = \frac{M\left[(Y_k - M(Y_k))^3\right]}{\left(\sqrt{D(Y_k)}\right)^3}$$

index of kurtosis

$$\gamma_a(Y_k) = \frac{M\left[(Y_k - M(Y_k))^4\right]}{\left(\sqrt{D(Y_k)}\right)^4} - 3$$

The calculation is made for each segment defined at this stage and the pathology specified in the mathematical model. After that, a comparison is made for a match in the parameters in the range defined by the rule 3σ. This calculation helps to define solution of the first criterion of the function.

We should set a research window $n = \frac{1}{3}k_d$ for realization of the second criterion of the function. The search for correlation coefficients and their comparison is performed in a sliding window in the range of the interval P-P.

$$cor_{pat} = \frac{\sum\left(g_k - \sum\frac{g_n}{n}\right)\left(Y_k - \sum\frac{Y_n}{n}\right)}{\sqrt{\left(\sum\left(g_k - \sum\frac{g_n}{n}\right)^2\right)}\sqrt{\left(\sum\left(Y_k - \sum\frac{Y_n}{n}\right)^2\right)}}$$

when k - elements of observations in the range of the processing window, cor_{pat} - correlation coefficient for each pathology.

When the threshold value is exceeded, a decision is made to detect the pathology, if it is confirmed for more than 30% of the searches in the processing window.

We'll make the transition to the frequency area using discrete Fourier transform for realization of the third criterion of the function.

$$X_{fft\_k} = \sum_{i=0}^{n-1} Y_i e^{-\frac{2\pi j}{n}fft\_k*i}$$

Comparison in the frequency area will be made according to the shape of the envelope of the frequency coefficients and the central moments calculated for them.

The envelope will be constructed using the method of least squares, taking into account the criterion $\sum(Y_n - f(x))^2 \to \min$. The function $f(x)$ is polynomial with the size of a polynomial not exceeding the 4th degree. The degree of the polynomial is bounded by the fourth order because of the computational complexity of the calculations of these indicators. In the case of coincidence of polynomial functions, or their discrepancy with an inaccuracy that does not exceed the threshold value, it is decided to confirm the diagnosis. Confirmation of the diagnosis in the frequency area is made in case of coincidence of the group of central moments with the central moments of the mathematical model of the supposed pathologies.

The diagnosis is made after the analysis of the results of each the components of the objective function.

4 Conclusion

As a result of the conducted researches the approach to the analysis of electrocardiograms is received. The proposed approach makes it possible to eliminate the noise component while retaining the signal pikes. According to the analysis of the received data, an algorithm for finding the diagnosis is proposed.

References

1. Chung, S. H., and R. A. Kennedy, Journal of neuroscience methods, 71-86 (1991)
2. Veeneman, D., and S. BeMent, IEEE transactions on acoustics, speech, and signal processing, no.2, 369-377 (1985)
3. Semenishchev, Evgeny, et al. *In Proc. EWDTS,* 444-449 (2016)
4. Stadler, Robert, et al. Patent No. 6115628 A, United States, A61B5/0452
5. Sameni, Reza, et al. Computers in Cardiology (2005)
6. Vorobyov, Sergiy, and Andrzej Cichocki, Biological Cybernetics. No. 4, 293-303 (2002)