Expert System for Structural Analysis of Electrocardiograms

Kristina Krasovitskaya[1], Evgeny Cherkashin[2], and Sergey Gorunchik[3]

[1]Irkutsk National Research Technical University,
Lermontov str. 83, Irkutsk, 664074, Russian Federation kristina.kras1993@gmail.com
[2]Matrosov Institute for System Dynamics and Control Theory of
Siberian Branch of Russian Academy of Sciences, Lermontov str. 134, Irkutsk, 664033,
Russian Federation eugeneai@icc.ru
[3]Specialized Tuberculosis Hospital in Rostov region,
Orskaya str. 24, Rostov-on-Don, 344000, Russian Federation sergeygreen@mail.ru

Abstract. The paper deals with process of a cardiological expert system development. A definition of a electrocardiogram is presented. Problems of ECG characteristics determination such as ECG data digitizing is considered. A problem of QRS complex recognition and P and T waves parameters measurement is discussed. A general outline of analysis technique for ECG using wavelet transformation are proposed.

Keywords: electrocardiogram, ECG, expert system, ECG recognition, electrocardiosignal, time series recognition, image processing, continuous wavelet transformation.

1.   Introduction

Medical images are popular objects of automatic analysis. An image of an electrocardiogram (ECG) contains time series of a cardiosignals, which are processed synchronously. Manual static analysis is always a multistage process. The signal processing software delivered with a cardiological diagnosis is intended for experienced functional diagnostician and rigidly bound to the appliance model, resulting to impossibility to make a comparative analysis of ECGs of the same patient data obtained in different medical centers and time moments. This, in turn, implies additional examination to be carried on the patient to obtain the time series in a required file and data format, increasing the cost of the patient residence in the diagnostic center. Also, Russia has regions where the specialized professional cardiological treatment is not possible, this can lead to fatal complications due to time delays.

In order to improve the quality and accessibility of medical cardiology provision for the population, Ministry of Health of Russian Federation started to elaborate of a law project aimed at a legal basis construction of IT infrastructure for public health protection. The infrastructure will be based on remote interaction of medical

professionals and patients. At present the interaction (such as remote consultations) is out of professional moderation and provided via forums, specialized internet sites and Skype. The cardiological data is supplied as attached photographic as scanned images. The final conclusions made in the environment are of question.

The automatic processing and analysis of the ECG will reduce human factor impact on the final conclusion since automatic processing results are not affected by the degree of fatigue, observation time, and body physical condition of the physician. Moreover the results produced by human expert are usually intuitive, expressed in quality statements, and the reasoning process can be hardly represented as explicit elementary stages (as an algorithm). For example, expert estimates the width of PQ interval to be shorter as compared to one described in the literature, in addition, the subjective estimation varies among specialists [1]. Computer processing and analysis is significantly formalized and operates by qualitative values.

The computational ECG processing has already more than fifteen years. But it seems that there are no industry standards of data representation and processing pipelines of the EGC images as well as complete set of methods, data formats, IT tools and techniques for implementation of each stage of the pipeline. The ECG image processing problem is classified as artificial intelligence recognition task with the lack and data distortion.

Among the existing ECG processing software made in Russian Federation we emphasize system "Valenta" as a popular software tool for ECG data analysis obtained from its instrument. The program does not allow processing data from other sources such as scanned tape images and time series obtained from other appliances. Monitor control systems, e.g., arrhythmia monitor system "Argus", require special equipment. Professor MD V. M. Uspensky suggested new diagnostic method that allows expert to diagnose a wide range of deceases of internal organs by describing a ECG. The method does not allow one to diagnose an essential set of hearth deceases causing arrhythmia that has 8% of all the cases in last three years. The analytic systems do not support importing raster graphics of scanned images of ECG printed on a paper — still the most popular media. Digitizing such sources and extraction of time series allow us to organize automated diagnostic environments on the base of electronic medical records.

In this investigation we develop an IT technique and software, which will process and analyze ECG raster images obtained in cardiologist's offices, medical organizations and from medical histories. The aim of the investigation is to develop software allowing cardiologists to watch the dynamics as an ECG evolution in time. To achieve so we need to digitize ECG time series, filter parasite distortions (low and high frequency), straighten the signal base, and measure signal relative entity sizes in % to the reference ones. Another problem is automatic interpretation of an ECG as a result of signal form analysis. The automatic analysis as a pedantic and formal procedure will rise the probability to mention, e.g., atrial fibrillation, and lower human factor in decision making.

2.   Structire of ECG

Electrocardiogram (ECG) is a graphical representation (a time series records) of impulse conduction on the nodes and bundles of the cardiac conduction system of heart. ECG is recorded as a number of *electrical entities* (*pikes* made of rising and drooping, and

horizontal segments) [2]. Pike entities are denoted by letters P, Q, R, S and T, whereas horizontal segments are denoted by PR segment and ST segment.

Each ECG analysis is started from verification of the registration process. The registration of hearth electric field potential, named as *lead*, is measured between two placements of the chest surface. The standard ECG consists of twelve leads (vectors): three 2-pole (3 standard leads), three 1-pole reinforced limb leads, six 1-pole chest leads.
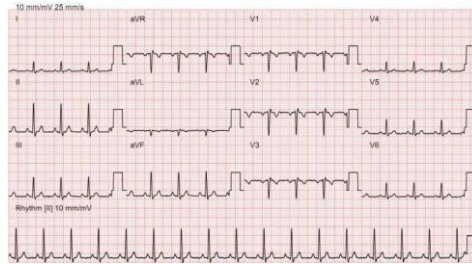

Fig.1. Standard twelve-lead ECG

3. Image processing

a. Recognition of lead time series

The first stage of the ECG image processing is a recognition of the time series if the leads. The source image scan consists of tracing paper background with square grid and lines of time series. Time series include a reference pulse at the beginning (Fig. 1). The processing of the image can be done with various techniques. We started directly from the application of edge tracing algorithms and obtained the results like ones depicted in Fig. 2.
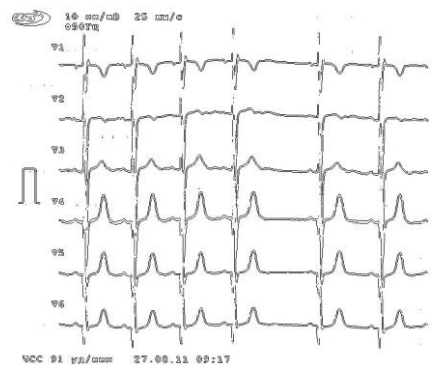

Fig.2. Sobel filter processing (inverted)

As the image consists of the combination of the background and signal time series, we decided to filter out the background [3,4]. The following schema implements the filtering.

1. A color channel of the background hue is extracted from the source image.
2. A red image copy is blurred with Gaussian filter in the red channel.
3. The digital image of the lead time series is obtained using the following formula: $s(n) = f(n) + \sigma e(n)$, where $f(n)$ is the signal, σ is the noise level, $e(n)$ is the Gaussian white noise, and $s(n)$ is the original signal. The threshold transformation is applied to the image as well.
4. The signal recording is restored in a cyclic processing of the columns of the obtained pixel array of the image.

The algorithm did not produce results on all the test given. Ten out of 156 ECGs did not completely analyzed due to threshold transformation function faults. The function parameters were adjusted by means of a machine learning technique. Hue and brightness characteristics of each column of the image were used as attributes of input train set.

Fig.3. Algorithm stages of time series recognition

b. Time series post processing

The original ECG image in a general case contains signal distortions such as baseline drift and proximity effect to energy supply network. A similar low frequency distortion could also be introduces by time series recognition algorithms. The baseline drift arises due to patient's strong movement, electrode polarization, appliance measurement error, and poor electrode contact to patient's skin. We used discrete Butterworth filter in the removal procedure of the low frequency distortions

where H is frequency response of the analog filter, k_c is a cutoff frequency (rad/s), k is the analog frequency in radiant. The energy supply network has usually 50 or 60 Hertz frequency, and its distortion is suppressed with band-stop and comb filters.

In the case when many copies of ECG is accessible, one can take advantage of synchronous and ensemble averaging techniques. The noise is a stationary stochastic process, and the ECG signal itself is (quasi-) periodical or cyclically-stationary. Sliding mean filer in time space is applicable to a statistically stationary low frequency signals. This technique provides filtering capabilities in real time mode.

If the real time mode is not required we use filtering in the frequency space. If a power spectral density and autocorrelation function parameters are known, one can make use of optimal Wiener filter. If the noise is not correlated to the signal or the distortion has nonstationary (or even stochastic) properties and there are no additional data, but we have a record from a secondary source, then adaptive filtering is used.

$$|H(k)|^2 = \frac{1}{1 + (k/k_c)^{2N}}$$

4. ECG entities recognition

c. QRS complex detection

The QRS complex detection is usually based on locating R pike with the following removal of the whole QRS complex from the ECG time series. The R location can be obtained by means of various methods, e.g., based on the derivative [5], sliding mean filer, weighted and squared operators of the first derivative [6], detection algorithm of PanTompkins [7]. For PQRST complex detection, there are discrete wavelet transformation based algorithms [8,9], and duration transformation [10].

In this project we used discrete wavelet transformation method as it can determine pikes in the signals with base line drift distortion. The previously recognized time series are loaded from a file, and 4th degree wavelet decomposition is computed over the loaded data (Fig. 4), the source signal is decomposed with Daubechies wavelet.

Fig.4. Low frequencies decomposition

At the next stage the obtained time series analyzed visually by operator (user). Operator chooses the variant that is similar to the original time series. The chosen variant contains no distortion. Then the R pikes are determined as maximal values among time moments, where signal amplitude is greater than 60% of the maximal digital value on the interval.
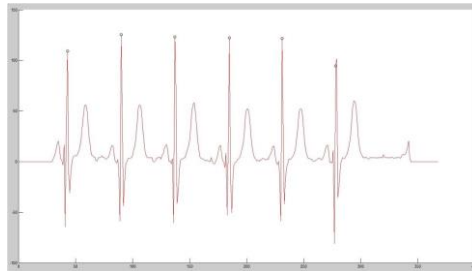

Fig.5. R pike determination

The determined maximal values positions are also filtered according to pike average distance criterion. The measured pike amplitude is corrected with respect to the basis of the original signal, then the locations of the rest of the pikes are determined (Fig.6 and 7).
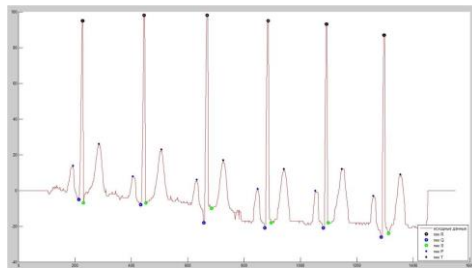

Fig.6. R pike determination (base line drift)

5.   Usage the digitized time series of leads

A comparative dynamics analysis of EGC of abnormal hearth rhythms in form of various arrhythmia and conduction disorder has been carried on during patient observations of primary manifestations of undesirable side effects on the months-long anti-tuberculosis chemotherapy. The abnormalities are as follows (with the number of observations): the rhythm of the right atrium (18 cases), atrioventricular block I st. (7 cases), full blockade of the left His' leg bundle (2 cases), Brugad syndrome (7 cases), sick sinus syndrome (1 case), pacemaker migration on the atria (2 cases), pacemaker migration on the atria (10 cases), lengthening the interval Q–T (1 case), shortening of P–Q interval — syndrome premature ventricular heart lerk-Levi-Cristesco (CLC) (12 cases), atal syndrome of

premature ventricular Wolff-Parkinson-White (WPW), type "B" (1 case), acute myocardial infarction with the formation of pathological wave the Q (3 cases), acute myocardial in-farction without pike Q (-) but with the rise of segment S-T (3 cases), sinus tachycardia (4 cases), isolated syndrome early ventricular repolarization (4 cases).
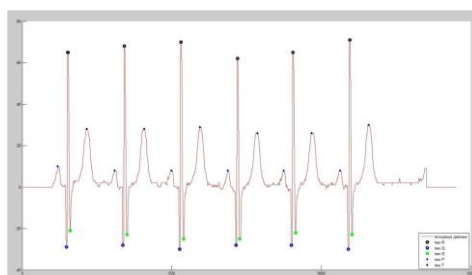

Fig.7. R pike determination (nondecomposed signal)

6. Conclusion

It this paper we briefly described an IT technique for ECG image scans processing and their time series recognition. The recognition result consists of P, Q, R, S and T section values expressed in qualitative terms, e.g., size, width, and amplitude. The results will be used as fact data for expert system of structural analysis and interpretation in medical cardiological domains (ES ECG "SERG"). We analyzed 156 ECG images obtained by scanning ECG tapes. The tapes are taken from archives of a hospital, ambulance brigades, and from patients' personal archives.

For some diagrams obtained in an emergency room and in a department of functional diagnostics; a comparative qualitative analysis of ECG time dynamics and accompanying patients supporting documentation has been carried out. Functionalist now is able to submit the objective values of amplitudes and time intervals of P, Q, R, S and T pikes and the segments of clinically significant QRS complexes in millimeters and milliseconds and in percent relations to a cardiologist. This resulted in rising the objectivity of dynamics evaluation with the domain criteria and coefficients, mentioned in manuals and guidelines in cardiology, but still rarely used by practitioners due to a lot of mathematical calculation and analysis to be done previously.

Cardiologist carried on an analysis of amplitude dynamics of the amplitudes and the time intervals, and segments parameters, resulting in a safe rational load correction of medicaition dosage for only 2–3 days instead regular 5–6 ones.

The further development of the technique, its software and their implementation in practice to ECG diagnosis expert systems will significantly improve the quality of cardiological monitoring of the patients with a complex comorbid disorders on an interdisciplinary level.

Kristina Krasovitskaya, Evgeny Cherkashin, and Sergey Gorunchik

Acknowledgment

References

1. Rangaraj, M.R.: Biomedical Signal Analysis. A Case-Study Approach. IEEE Press and Wiley, New York, NY. (2002). ISBN 0-471-20811-6

2. Electrocardiography – Wikipedia, the free encyclopedia, [Online]. Available:https://en.wikipedia.org/wiki/Electrocardiogram (current March 2016)

3. Gonsalez, R.C., Woods, R.E.: Digital Image Processing. MA: Addison-Wesley. (1992)

4. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital Image Processing Using MATLAB, 2nd ed. Gatesmark Publishing. (2009)

5. Ahlstrom, M.L., Tomhins, W.J.: Digital filters for real-time ECG signal processing using microprocessors. IEEE Trans. Biomed. Eng. vol. 32. pp. 708–713 (1985)

6. Murthy, I.S.N., Rangaraj, M.R.: New concepts for PVC detection. IEEE Trans. Biomed. Eng. vol. 26, 7 pp. 409–416 (1979)

7. Pan, J., Tomhins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. vol. 32. pp. 230-236 (1985)

8. Khaled, D., Ibrahim, N.A., Abdel-Rahman, Al-Q.: QRS Complex Detection Based on Symmetry Wavelet Function. In: 5th International MultiConference on Systems, Signals and Devices (2008)

9. Dubrovin, V.I., Tverdohleb, J.V., Kharchenko, V.V.: R-peaks detection using wavelet Technology. Radioelectronka, informatika, upravlinnya. 2(29) (2013), [Online]. Available:http://cyberleninka.ru/article/n/ r-peaks-detection-using-wavelet-technology (current March 2016)

10. Gritzali, F., Frangakis, G., Papakonstantinou G.: Detection of the P and T waves in an ECG //Comput. Biomed. Res. vol. 9. pp. 125–132 (1976)