

A Decision Support System for Ischemic Event Detection

FO Favretto, CRG Farias, LO Murta Jr

Department of Physic and Mathematic - FFCLRP, University of São Paulo, Ribeirão Preto, SP,
Brazil

Abstract

This paper describes the main aspects of a decision support system for automatic detection of ischemic events from electrocardiogram (ECG) digital data. In order to detect ischemic events, the system extracts a number of autoregressive coefficients from the ECG signal as well as a measure of the ST-segment deviation. The Quadratic Discriminant Function was used to perform the ECG classification. An ECG visualization tool and an ECG digital data to XML converter were also developed in scope of this work.

1. Introduction

Myocardial ischemia is one of the most important heart diseases. The ischemia is defined as an insufficient blood flow to maintain the normal tissue's functions [1]. The myocardial infarction is a potential result of cardiac ischemia, but its diagnostic is not always straightforward, since thoracic pain can be induced by other factors than myocardial infarction. There are also ischemia episodes that are silent, like in diabetic patients.

The electrocardiogram (ECG) is an easy and immediate diagnostic tool for assessing the probability of a cardiac event in the pre hospital phase [2]. The use of computer systems that are able to automatically process and identify abnormal events present in ECGs enables the development of new applications that improve the quality of provided health services.

This paper describes the development of a decision support system to automatically detect myocardial ischemia from ECG data. This system also provides support to ECG visualization and conversion of ECG signal to FDA/HL7 annotated ECG format [3].

The remaining of the paper is structured as follows: next section briefly introduces the FDA/HL7 annotated ECG format and explains our motivation towards its adoption; section 2 explains the ECG feature extraction and classification algorithm used in this work; section 3 provides an overview of the results achieved; and finally, section 4 describes some conclusions.

2. Methods

The electrocardiography is one of the most fundamental and frequently applied tests in medicine. The ECG acquisition and analysis are processes well established, standardized and non invasive. However, each ECG equipment manufacturer provides its own (proprietary) format to capture and store ECG data.

The Food and Drug Administration (FDA) developed the data standard *FDA-XML Data Format Requirements Specification* [3] with the purpose of avoiding problems related to communication and interoperability among ECG systems and health information systems. This standard describes the structure of an Extensible Markup Language (XML) document that contains waveform data, such as ECG data.

XML is a standard developed by the Word Wide Web Consortium (W3C) used to describe structured information using a textual (ASCII) notation [4]. The structure of XML documents can be described using an XML Document Type Definition (DTD) or an XML Schema Definition (XSD). Since XML documents are text documents, a number of Application Programming Interfaces (APIs) are currently available to support the access and manipulation of the document contents and structure.

The FDA selected XML to represent ECG data, and developed a corresponding Schema and DTD to define the structure of these documents. The FDA/HL7 annotated ECG format captures all 12 ECG leads as well as patient related information, such as name, age, sex, exam data and duration.

The decision support system was developed using the Java programming language. Our motivation to use Java includes its portability across different hardware and operating system platforms and the availability of (standard) APIs that facilitate access and manipulation of XML documents and chart generation.

Since the ECG data must be in FDA/HL7 Annotated ECG format, we have developed a format converter from ECG digital data to the specified XML format. The standard Java API used for generation and handling XML

documents was DOM (Document Object Model), which enabled the direct access to each element of the XML document. In order to produce the ECGs visualization charts, we have used the JFreeChart API [5]. We have used NetBeans IDE as our development platform.

In order to classify an ECG, we have to extract five different features from each exam: four of them refer to signal modelling, while the fifth feature refers to the ST segment deviation. The classification algorithm was based in Quadratic Discriminant Function. The features extraction phase and the classification function are explaining in the sequel.

ECG Feature Extraction

We have used Scalar Auto Regressive (AR) Modeling to extract four parameters, called AR coefficients, from the ECG. These parameters are used to estimate the signal behavior. AR is a technique widely used to model bio-signals, c.f, [6].

The fifth parameter is the ST-segment deviation, which is measured as the elevation or depression of ST-segment with respect to the ECG baseline [7]. From the medical point of view, this is the most significant ischemic change.

AR Modeling

AR models are used to analyze temporal series behavior, where each series value is described as a combination of its p previous values in the series, considering a model with order p . The model's coefficients, b_1, b_2, \dots, b_p , weight the p previous values of y_i and are denoted AR coefficients.

$$y_i = b_1 y_{i-1} + b_2 y_{i-2} + b_3 y_{i-3} + \dots + b_p y_{i-p} \quad (1)$$

Burg's algorithm [8] was used to compute the AR coefficients. To extract p coefficients from the series, a model with order p must be used. Thus, in this work we have used an AR model with order 4.

ST-Segment Deviation

The algorithm used to calculate the ST-Segment Deviation was based on the algorithm presented in [7].

Cardiac beat baseline

Given that each beat generates a baseline, in order to calculate this baseline we need to:

1. Detect the R-peaks from the signal. The R-peak detection algorithm was based in the QRS Complex detection algorithm [9]:

- i. Calculate the derivative of ECG signal;
- ii. Go through the derivative signal. When the signal value is higher than a given threshold, start searching for the point of maximum value

- in the next 0.036 seconds of the signal;
- iii. The point found in ii is marked as R-peak of the beat.

2. Find the middle point (mp) of each RR interval (interval between two R-peaks);

3. Calculate the mean of the signal in a small interval [mp, mp+RR/20];

4. The mean value found in 3 is taken as the baseline value for the current beat.

Figure 1 illustrates the small interval used to calculate the mean of the signal. The value found will be the baseline of the beat.

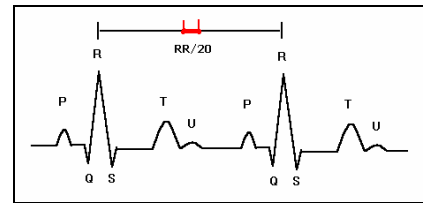


Figure 1. Interval used to calculate the mean of the signal.

ST-Segment Deviation

The point of ST-segment that will be used to calculate the deviation of ST-segment with respect to the baseline depends on the heart rate of the patient [7]. Table 1 shows the relation between heart rates and the ST-segment deviation measuring points.

Table 1. Measuring point of the ST-segment deviation

Heart rate	ST-segment deviation measuring point
<100	R+120ms
100~110	R+112ms
110~120	R+104ms
>120	R+100ms

For each beat, we have calculated the difference between the signal value on the ST-segment and the baseline value of the beat. After calculating this difference for all beats, we can calculate the mean of these values. The value found is the fifth parameter of the ischemia detection algorithm.

Classification Algorithm

The classification algorithm is based on the Quadratic Discriminant Function (QDF). The matricial form of QDF is given by Equation 2:

$$y_i = X_i \beta + \varepsilon_i \quad (2)$$

where y_i is a scalar that represents the classification result for a given exam i , been -1 for ischemic ECGs and 1 for normal ECGs; X_i is the QDF feature vector, a line vector that carry the information of the 5 extracted

parameters, been (x_1, x_2, \dots, x_d) the features extracted from the signal, when $d = 5$, the structure of QDF feature vector is given by $X_i = [1, x_1, x_2, \dots, x_d, x_1^2, x_2^2, \dots, x_d^2, 2x_1x_2, 2x_1x_3, \dots, 2x_1x_d, 2x_2x_3, 2x_2x_4, \dots, 2x_2x_d, \dots, 2x_{d-1}x_d]$; β is the QDF estimator, a column vector which structure is given by $\beta = [\beta_0, \beta_1, \beta_2, \dots, \beta_d, \beta_{11}, \beta_{22}, \beta_{dd}, \beta_{12}, \beta_{13}, \dots, \beta_{1d}, \beta_{23}, \beta_{24}, \dots, \beta_{2d}, \beta_{(d-1)d}]^T$; and ε is the classification error.

The vector β can be found by mathematics calculus from D ECG signals already classified, i.e., diagnosed by an expert physician. Been:

$$\tilde{Y} = A\beta + E \quad (3)$$

where $\tilde{Y} = [y_1, y_2, y_3, \dots, y_D]^T$ is a column vector D -dimensional that contain the classification values ('-1' or '1') of D ECG signals; A is a matrix that contains the features vectors of D ECG signals, $A = [X_1, X_2, X_3, \dots, X_D]^T$; and E is the error vector, $E = [\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_D]^T$.

Considering that the set of D signals was correctly classified by the expert, the error vector E , will be null, so we can find β using Equation 4:

$$\beta = (A^T A)^{-1} A^T \tilde{Y} \quad (4)$$

The classification of a new ECG data i can be done multiplying its feature vector X_i by the QDF estimator, β , according to Equation 5:

$$y_i = X_i \beta \quad (5)$$

In case the resulting scalar y_i is lower than a threshold, the classification result indicates that the ECG does not have ischemic events. Otherwise, the classification indicates the occurrence of ischemic events. The threshold value was computed by ROC Curve.

3. Results

The ECG data used in the tests was obtained from the MIT PhysioBank database. The database, extracted from Holter exams, consists of both normal and ischemic data. Each exam corresponds to a 48 hour ECG. However, we have extracted normal and ischemic segments of 10 seconds from channel 1, according to the expert annotations. Overall, we have extracted 50 normal ECG segments and 22 ischemic ECG segments.

Each segment was pre-processed using two digital filters: Notch Filter of 50 Hz, to eliminate 50 Hz noises generated by electric influence of the equipment; and High-Pass Filter, to remove low frequency oscillations and put the signal on its baseline.

A percentage of the data was used to create the QDF estimator β , and the remainder used to test the classifier. Three data configuration were established:

- i. 30% for create β and 70 % for test;
- ii. 40% for create β and 60% for test;

- iii. 50% for create β and 50% for test.

For each configuration (30-70, 40-60 and 50-50), we have performed three tests. The data of QDF estimator set and classifier test set were randomly selected in each one of the three tests. Several values of thresholds, that discriminates the group of Normal and Ischemic exams, were used and a ROC Curve was generated for each test. We have observed that by using only 30% of the data to generate the classifier and the remainder 70% to test the classifier, it was already possible to obtain 100% of sensibility and specificity (see Figure 2).

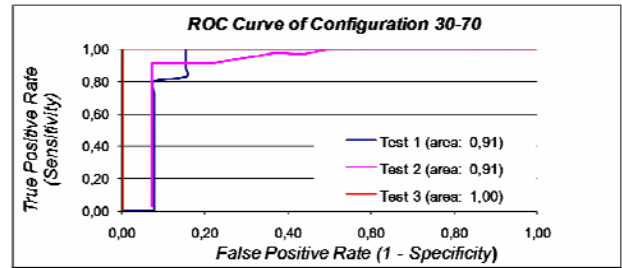


Figure 2. ROC Curves for classifier test with the configuration 30-70.

The other configurations also provided good results. Table 2 resumes the tests performed using the three configurations. The tests shows that the methodology used to construct the Ischemia Classifier seems to be adequate once satisfactory results were obtained.

Table 2. ROC curve areas for the tests

Configuration	Area above the ROC Curve			
	Test 1	Test 2	Test 3	Mean
30-70	0.91	0.91	1.00	0.94
40-60	0.82	1.00	0.88	0.90
50-50	0.82	0.99	0.88	0.90

Another important result of this work was the ECG visualization tool developed. This tool uses the ECG digital data available in XML document and generates a graphic with the 12 leads signals. It is also possible for users to visualize each lead separately. In case we need to analyze a piece of signal in detail, it is possible to select the stretch and the tool will zoom it accordingly. The Figure 3 illustrates the visualization tool module.

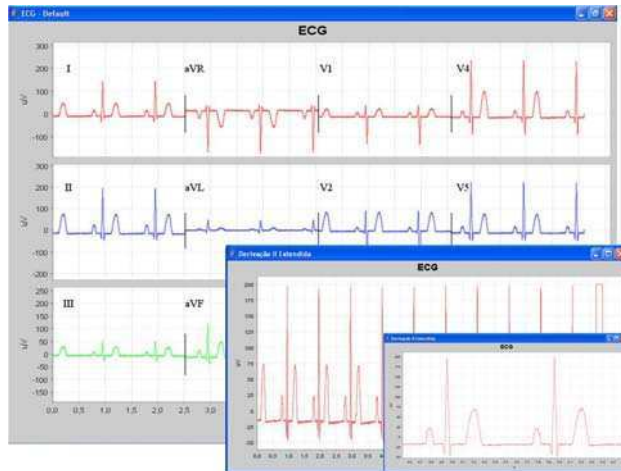


Figure 3. ECG visualization tool.

4. Discussion and conclusions

The achieved system has potential to assist a screening process in ischemia detection. The accuracy tests showed a good agreement to manually annotated ischemic events. Possible applications to the system ranges from a situation with the absence of specialty cardiologist to when a second opinion is needed.

This system provides agility in patient care, ischemia diagnostic at screening time, and can be an improvement in cardiologic treatments. Furthermore, the purposed system has potential applications in telemedicine.

Acknowledgements

This work has been supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under project number 50.6284/2004-2.

The authors are also thankful to São Paulo State foundation (FAPESP) for financial support.

References

- [1] D'alecy LG, Zelenock, GB. Pathophysiology of ischemia and hypoxia: The CV Mosby company. 1990;147-158.
- [2] Rubel, P, Fayb, J, Nollo G, Assanelli, D, Li B, Restier L, Adami S, Arod S, Atousi H, Ohlsson M, Simon-Chautemps L, Telisson D, Malossi C, Ziliani G, Galassi A, Edenbrandt L, Chevalier P. Toward personal eHealth in cardiology. Results from the EPI-MEDICS telemedicine project. J Electrocardiol 2005; 38: 100-106.
- [3] Brown B, Kohls M, Stockbridge N. FDA XML Data Format Design Specification. 2002; <http://www.fda.gov>
- [4] W3C. Extensible Markup Language (XML) 1.0. 2006; <http://www.w3.org/TR/2006/REC-xml-20060816/>
- [5] JFreeChart;: <http://www.jfree.org/jfreechart/>
- [6] Zhang ZG, Jiang HZ, Ge DF, Xiang XJ. Pattern Recognition of Cardiac Arrhythmias Using Scalar Autoregressive Modeling. IEEE-WCICA 2004; 6: 5545-5548.
- [7] Pang L, Tchoudovski I, Braecklein M, Egorouchkina K, Kellermann W, Bolz A. Real time heart ischemia detection in the smart home care system. IEEE- EMBS 2005; 3703-3706.
- [8] Shiavi R. Introduction to Applied Statistical Signal Analysis. 1999; Academic Press,.
- [9] HH So. Development of QRS detection method for realtimeambulatory cardiac monitor. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society 1997; 282-292.

Address for correspondence

Name: Luiz Otávio Murta Júnior

Postal address: Department of Physics and Mathematics, University of São Paulo, Av. Bandeirantes 3900, 14040-901, Ribeirão Preto (SP), Brazil

Email address: murta@ffclrp.usp.br