LATIN.SCIENCE
XVI Congresso Latino-americano
de Software Livre e Tecnologias Abertas
LATINOWARE 2019

Software System Development to Diagnose Heart Diseases

Renan Yuji Koga Ferreira*, Guilherme Camargo Fabricio de Mello†, Fabio Sakuray‡, Wesley Attrot§
Department of Computing
Londrina State University, UEL
Londrina, Brazil
*Email: *renankoga@uel.br, †guilhermecfmello@uel.br, ‡sakuray@uel.br, §wesley@uel.br*

***Abstract*—Many deaths are caused from heart diseases and several of them could be prevented with early detection. Many people do not have conditions to seek for a doctor or sometimes there are not enough physicians to attend them. In order to detect heart diseases we are developing an electrocardiogram feature extraction algorithm using wavelet transforms prioritizing a low computational cost. This algorithm will be integrated in an embedded system that is under development. This system is going to be accessible, portable and have low cost, because we intend to assist people, mostly those who live in precarious regions, that do not have a physician to attend them. To execute tests on our algorithm we will use the ECG records from MIT-BIH database and after that we will classify the heartbeats in order to detect anomalies on them.**

***Keywords*-ECG; heart diseases; feature extraction; wavelet transform.**

I. Introduction

According to World Health Organization, cardiovascular diseases are the main cause of deaths in the world. In 2016, 17.9 million people died from cardiovascular diseases, which represents 31% of all global deaths. Most of these diseases could be prevented by paying attention to risk factors such as smoking, unhealthy diet, obesity, sedentarism, harmful use of alcohol, among others [1].

In order to reduce these numbers some preventive health care should be considered. There are approaches that perform an early diagnosis in an attempt to prevent an aggravation in the future, for instance, an electrocardiogram analysis.

Electrocardiogram (ECG) is a simple recording of the electrical activity generated by the heart [2] and can be used for many purposes: heart rate, heartbeat rhythm, diagnose heart abnormalities, emotion recognition and biometric identification [3].

However, there are situations that electrocardiogram analysis are impracticable, such as long term ECG analysis or even a real time analysis. Therefore, in order to assist the physicians in that scenario, some approaches uses an automated electrocardiogram analysis.

A study conducted by professor Mario Scheffer from USP [4], with institutional support from CFM (Federal Council of Medicine) and Cremesp (Regional Council of Medicine from Sao Paulo) revealed that the physicians in Brazil have been concentrated in developed cities, capitals and coast cities.

According to World Health Organization (WHO) the ideal rate of physicians per 1000 population is 1. This rate from Brazil capitals are all higher than 1.6 reaching up to 12.27, however there are cities with low economic development where the rate is 1.28. But some cities in many states from North and Northeast Regions have a rate less than 0.50 [4], although this is not a reality only from Brazil.

The people who live away from developed cities are the most impaired, because the physicians only go there periodically, or worst, they might not have a physician in their cities forcing them to go to the nearest developed city looking for a doctor.

Therefore our objective is to develop an embedded system with low computational cost because we intend to use this to assist people who live in areas that a presence of physicians is not so frequent. Thus, if our system diagnose some anomaly (some types of arrhythmias), that person will be advised to seek a doctor for a precise diagnose. The system structure is described in Figure 1. We will develop and evaluate a system to extract the fiducial points[1] from QRS complex using wavelet transform. These extracted points will be necessary to classify the heartbeats in order to detect anomalies or a healthy heartbeat.

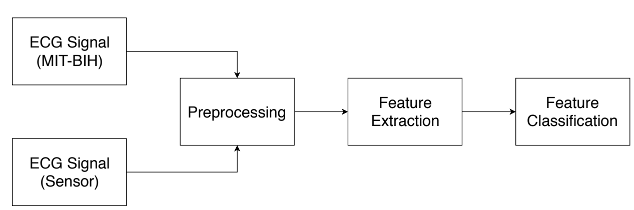

Figure 1. Our project description

The objective of this paper is to show the current status from our research. We will briefly present the techniques in the literature and how we will intend to develop our own technique based in wavelet transform.

[1]The standard fiducial points in ECG are: P, Q, R, S and T

II. RELATED WORK

A. Electrocardiogram

In the electrocardiogram, it is possible to detect a cardiac cycle through the P-QRS-T complex. This complex is related to muscular contraction, which is associated with electrical changes known as depolarization [5]. The P wave is caused by the atria contractions, and the ventricles contractions produce the QRS complex. The T wave is the outcome of ventricles returning to a rest state, known as repolarization. Figure 2 represents the P-QRS-T complex in an electrocardiogram. The waves, segments[2] and intervals[3] of this complex are described bellow:

- P wave – low amplitude wave, because the atria muscle is smaller than ventricular;
- PR interval – the interval between P wave and the beginning of QRS complex. After the P wave, the ECG returns to the same electric level that there were before it. The PR interval is the junction of the P wave and the PR segment;
- QRS complex – the Q, R and S waves together;
- RR interval – interval between two R peaks;
- ST segment – region between the ending of the S wave and the beginning of the T wave;
- T wave – represents the main ventricular repolarization;
- QT interval – interval between the beginning of the QRS complex and the ending of the T wave;
- PP interval – interval between the beginning of two P waves;
- TP segment – region between T and P wave;
- U wave – low amplitude wave at the end of T wave and this is seen occasionally;
- J point – the junction between the QRS complex and the ST segment.

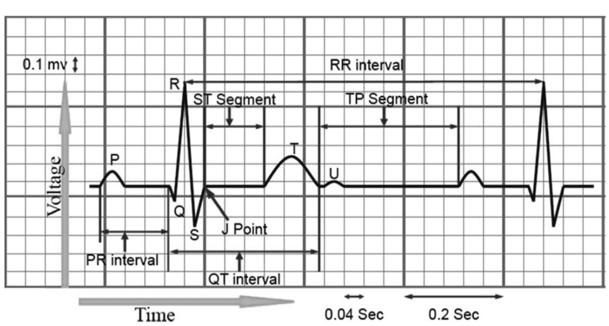

Figure 2. P-QRS-T complex in a ECG as shown in [3]

B. MIT-BIH Database

The MIT-BIH is a database developed by MIT and Boston's Beth Israel Hospital (now the Beth Israel Deaconess Medical Center) laboratories to perform research about arrhythmia analysis and correlated subjects [6].

The electrocardiograms in this database were obtained from the Beth Israel Hospital Arrhythmia Laboratory on the 70's. There are 48 records from a mixed population of inpatients and outpatients. Thus, these records contains electrocardiograms from healthy people and also from people who have arrhythmias [7].

Two cardiologists worked independently adding subtitles on the heartbeats and removing false detections when it was necessary. The physicians were not able to classify some heartbeats, most of the time, the reason was the lack of information or technical issues.

This database is still used today for many researchers specially for those who want to develop an arrhythmia detection technique. The main reasons for using this database, are the representative and well-characterized collections of ECGs [8]. Some works such as Jen et al [9] and Zhao and Zhang [10], developed approaches using the MIT-BIH.

C. Feature Extraction

In order to detect heart diseases from ECGs, it is necessary to analyze the features that can be extracted from a electrocardiogram signal. These features can be used individually or combined with others.

The P-QRS-T complex is extracted from an ECG signal, and basically corresponds to the local, amplitudes, waveform or wave deviation [11], [12]. They are very important to analyze an electrocardiogram.

Due to fiducial points extracting importance, many algorithms have been developed over the literature, although the most used is Pan-Tompkins [13]. This algorithm consists in digital analyses on waves slopes, amplitudes and width. A bandpass filter reduces the false detections caused by noises, and can increase the sensitivity using lower thresholds.

There are other feature extraction algorithms, and according to Berkaya et al. [3] they can be classified in derivative, digital filters, wavelet transform and neural network.

The main problem of QRS complex detection is the physiological variations and the several kind of noises. These noises can be caused by muscle artifact, electrode movement, high frequency T wave and so on. In order to reduce these noises and improve the fiducial points detection, the electrocardiogram is preprocessed by the following steps: filtering, resampling, digitization, noise removing, and normalization.

D. Wavelet Transform

Wavelet transform (WT) are used to analyze periodic, noisy, intermittent and transient signals. It can be used for climate analysis, image compression, cardiac monitoring, among others ([5], [14]).

Transforming a signal using wavelets can highlight some features on it, this is known as wavelet transform. Mathemat-

[2]Line that connects two waves without including either of them
[3]Portion that includes a segment and one or more waves

ically speaking, the WT can be interpreted as convolution of the signal with the wavelet function.

The WT is computed in several places and for diverse scales among the signal. This computation can be continuous in the whole signal and is known as Continuous Wavelet Transform (CWT), or in discrete steps known as Discrete Wavelet Transform (DWT) ([5], [15]).

In this paper we will focus on the DWT because it allows the fast computation of the wavelet transform and its inverse, using infinite summations of discrete coefficients rather than continuous integrals used in CWT [5]. The DWT performs a highpass and lowpass filtering, decomposing the signal in two parts in each level: approximation coefficients and details coefficients. This filtering process is shown in Figure 3, where S is the signal, A_i and D_i are the approximation and details coefficient, respectively, of i-th level.

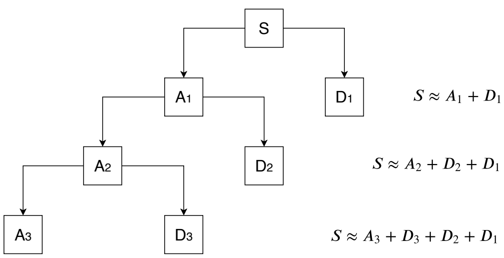

Figure 3. Discrete Wavelet Transform Filtering

There are many discrete wavelets, but we will focus on Daubechies family. The simplest known wavelet from Daubechies is the Haar wavelet, therefore this is very used for testing discrete wavelet transform algorithm. This wavelet is also known as D2 because it has only two scaling coefficients. These coefficients are used in the convolution process in DWT represented as c_k in (1) and the b_k in (2) is $b_k = (-1)^k c_{N_k - 1 - k}$. The S and T are the approximation and details coefficients respectively, where $S_{0,n}$ is the original signal, n is the n-th discrete point from signal and N_k is the total of scaling coefficients.

$$S_{m+1,n} = \frac{1}{\sqrt{2}} \sum_{k=0}^{N_{k-1}} c_k S_{m,2n+k} \tag{1}$$

$$T_{m+1,n} = \frac{1}{\sqrt{2}} \sum_{k=0}^{N_{k-1}} b_k S_{m,2n+k} \tag{2}$$

The total coefficients from each daubechies wavelet is defined in J as DJ. There are other daubechies wavelets such as D4, D6, D8, D10, among others.

III. The Good Heart Project

We developed an application in Qt Creator [16] to read and process the electrocardiograms (ECGs) from the MIT-BIH. In this application we show the ECG from a selected record and its annotations, as shown in Figure 4. There are some windows containing subtitles to assist the user to understand each annotation in the ECG.

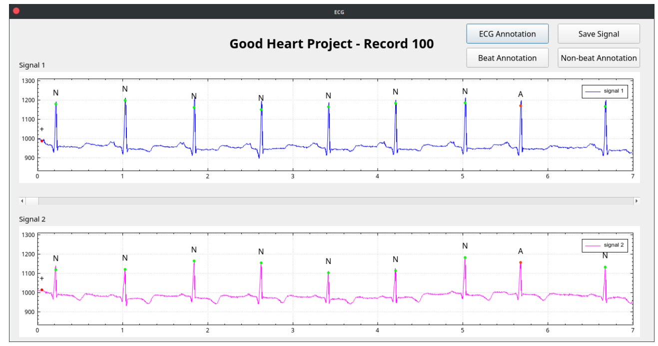

Figure 4. Screen from the application showing the ECG with annotations

Qt Creator was chosen because we needed to develop a cross platform and user interface (UI) software and it has a friendly interface to manage the items to build an UI application.

We are developing a QRS complex extraction method using wavelet transform, focusing on Daubechie family. The Daubechies wavelets are very similar to a cardiac cycle, hence we expect to have a great result choosing it as a mother wavelet. We are running tests using D4 and D6 wavelets because their waveform is similar to a heartbeat and Mahmoodabadi et al. [4] achieved a good result using these daubechies with a sensitivity of a approximately 99.18% and a positive predictivity of approximately 98.00%.

Using wavelet transform we can detect the fiducial points and when they occurred. Therefore we apply the DWT on the signal to decompose it in approximations and details, repeting this process until we get the frequency range of fiducial point R. After detecting it, we can detect the Q and S peaks delimiting the standard interval between Q and R waves and R and S waves [17].

Another activity of this project that is under development is the real time ECG reading from a person using an Arduino [18] pulse sensor as shown in Figure 1. After reading the ECG the record will be used in our software to process it and diagnose whether have an abnormality or not. Also we can convert the reading to the MIT-BIH format.

IV. Conclusion

In the future we intend moving to a Raspberry Pi [19] because it has more computational power than an Arduino and integrate our feature extraction algorithm on it. Hence we need to adjust our feature extraction algorithm prioritizing a low computational cost. These adjusts can decrease the prediction of the algorithm however we will not allow this compromise the system, maintaining a balance between prediction rate and computational cost. So our goal is to achieve at least 95% of accuracy in feature extraction. In order to improve the ECG detection we intend to use several electrodes sensors to monitor the heartbeats.

After feature extracting we need to classify these beats and we are considering the possibility to classify them using support vector machine (SVM), artificial neural network (ANN), linear discriminant analysis (LDA), k nearest neighbor (kNN), decision tree (DT) or bayesian [3]. We will execute tests and choose one of them.

The validation will be performed comparing the system results with the physician's diagnoses. In another phase of the project we will restrict the set of diseases focusing on the most common anomalies in Brazilian's reality, such as Chagas disease caused by *Trypanosoma cruzi*.

Our system is being built to assist people from precarious regions that do not have a physician to attend all people. Therefore it will be user friendly and accessible to people and organizations.

Another benefit of our system is that it is portable, thus we can reach outlying cities and assist their population without much effort. We do not intend to replace the physicians, that is, in a diagnose of some anomaly, the system will advise the person to seek for a doctor. Our purpose is to save lives preventing (or minimizing) the consequences of untreated heart diseases.

References

[1] World Health Organization, "Cardiovascular diseases (cvds)," https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), May 2017.

[2] A. Alberdi, A. Aztiria, A. Basarab, "Towards an automatic early recognition system for office environments based on multimodas measurements: A review," *Journal of Biomedical Informatics*, 2016.

[3] Selcan Kaplan Berkaya, Alper Kursat Uysal, Efnan Sora Gunal, Semih Ergin, Serkan Gunal, Mehmet Bilginer Gulmezoglu, "A survey on ecg analysis," *Biomedical Signal Processing and Control*, 2018.

[4] C. C. F. de Medicina, "Demografia mdica: Publicao j est disponvel na internet," https://portal.cfm.org.br/index.php?option=com_content&view=article&id=27509:2018-03-21-19-29-36&catid=3.

[5] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*. CRC Press, 2016.

[6] Physionet, "Mit-bih arrhythmia database directory," https://physionet.org/physiobank/database/mitdb/, May 1997.

[7] Physionet, "Mit-bih arrhythmia database directory," https://physionet.org/physiobank/database/html/mitdbdir/intro.htm, May 1997.

[8] George B. Moody, Roger G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, 2001.

[9] Kuo-Kuang Jen, Yean-Ren Hwang, "Ecg feature extraction and classification using cepstrum and neural networks," *Journal of Medical and Biological Engineering*, 2006.

[10] Qibin Zhao, Liqing Zhan, "Ecg feature extraction and classification using wavelet transform and support vector machines," *International Conference on Neural Networks and Brain*, 2005.

[11] Yun-Chi Yeh, Wen-June Wang, Che Wun Chiou, "Cardiac arrhythmia diagnosis method using linear discriminant analysis on ecg signals," *Measurement*, 2009.

[12] Yun-Chi Yeh, Wen-June Wang, Che Wun Chiou, "Feature selection algorithm for ecg signals using range-overlaps method," *Expert Systems with Applications*, 2010.

[13] Jiapu Pan, Willis J. Tompkins, "A real-time qrs detection algorithm," *IEEE Transactions on Biomedical Engineering*, 1985.

[14] D. L. Fugal, *Conceptual Wavelets in Digital Signal Processing*. Space & Signals Technologies LLC, 2009.

[15] S. Mallat, *A Wavelet Tour of Signal Processing (3. ed.)*. Academic Press, 1999.

[16] T. Q. Company, "Qt - cross-platform software development for embedded & desktop," https://www.qt.io/.

[17] Gari D. Clifford, Francisco Azuaje, Patrick E. McSharry, *Advanced Methods and Tools for ECG Data Analysis*. Artech House, 2006.

[18] Arduino, "Arduino - home," https://www.arduino.cc/.

[19] R. P. Foundation, "Raspberry pi," https://www.raspberrypi.org/.