



ISSN: 0976-3376

Available Online at <http://www.journalajst.com>

ASIAN JOURNAL OF
SCIENCE AND TECHNOLOGY

Asian Journal of Science and Technology
Vol. 11, Issue, 01, pp.10672-10676, January, 2020

RESEARCH ARTICLE

AUTOMATIC FEATURE EXTRACTION FROM ECG SIGNAL

^{1,*}MOUROUFIE Kouassi Kouman Vincent, ²KAM Oleh and ³SANGARE Moussa

¹Centre Suisse de Recherches Scientifiques en Côte d'Ivoire (CSRS)

²Institut d' Ethno-Sociologie (IES) de l' Université Félix Houphouët Boigny

³Université Peleforo Gon Coulibaly

ARTICLE INFO

Article History:

Received 25th October, 2019

Received in revised form

09th November, 2019

Accepted 27th December, 2019

Published online 31st January, 2020

Key words:

Electrocardiograph (ECG), ECG Data,
Pan-Tompkins Algorithm (PTA), R
peak, Heart rate

ABSTRACT

ECG describes the electrical activity of the heart and represents the graphical wave through which a physiologist can identify diseases of the heart. ECG consists of five waves such as P, Q, R, S and T wave. An ECG signal contaminated with noise. The noise is removed by digital band pass filter. There are many algorithms for calculating heart rate. Pan-Tompkins algorithm is one of them which gives more accuracy to heart rate than any other algorithms. Pan-Tompkins algorithm has been used to identify the R peak. At first low pass filter has been used to remove the high frequency. Then high pass filter has been used to remove the low frequency. The derivative filter, squaring function, moving window function and threshold method have also been used to identify it. Finally R peak is detected. After that heart rate is being calculated from R-R interval emitted from the signal. Average amplitude of R-wave and total number of R-peak are being calculated from the signal.

Citation: MOUROUFIE Kouassi Kouman Vincent, KAM Oleh and SANGARE Moussa. 2020. "Automatic feature extraction from ECG signal", *Asian Journal of Science and Technology*, 11, (01), 10672-10676.

Copyright © 2020, MOUROUFIE Kouassi Kouman Vincent et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

ECG is usually a recording of an electrical signal obtains from the body surface and represents the sum of action potentials from the heart cells. Electrical activity of the heart is shown by the changes of electrical voltage on the body surface. These changes are created by summarizing of electrical activity of all heart cells. The cells fail to work synchronously for the reason that they do not have a similar potential simultaneously and currents flow on the surface of membranes and everywhere around the heart. Human body is conductive because it consists of a lot of electrical live ions. As a result, voltage generates between two locations on the body and electrocardiography registers its improvement in time. The physical activity of the heart is connected with its electrical activity. Therefore, heart beats that repeat periodically. In every heart beat we are able to identify various waves and inter-wave segments. The shape and length of these waves and inter-wave segments define cardiovascular diseases, arrhythmias, myocardial ischemia and other heart diseases. The time axis usually uses the order of milliseconds, as the potential axis generally uses the order of millivolts [Autor Prace].

Willem Einthoven (21 May 1860 – 29 September 1927) was a Dutch doctor and physiologist. He invented the first practical electrocardiogram (ECG) in 1903 and achieved the novel prize in medicine in 1924 for it [<https://ecglibrary.com/ecghist.html>]. Heart Rate Variability (HRV) can be determined from ECG signal. The ECG signal determines the number of heart beats per minute. Dr. Willem Einthoven, a Dutch physiologist inspired by the work of Waller, refined the capillary electrometer even further and was able to demonstrate five deflections which he named ABCDE (Fig. 1). Uncorrected curve is labeled ABCD. This tracing was made with processed Lippmann capillary electrometer. The other curve was mathematically solved by Einthoven to permit for inertia and friction in the capillary tube. He selected the letters PQRST for the adjusted curve dependent on mathematical history of labeling successive point on a curve. ECG means electrocardiogram. All the electrical activity of the heart representing in graph this process is called ECG. ECG signal reflects the electric activities of a heart muscles. ECG is an important part in our life. The ECG pattern consists of a repeated wave sequence of P, QRS and T-wave related to each beat. It can be used to determine arrhythmia. Arrhythmia means abnormal heart rate which is dangerous for men/women which causes death. The automatic identification of ECG waves is essential for heart disease medical diagnosis. In a clinical setting, such as intensive care units, it is essential for

*Corresponding author: MOUROUFIE Kouassi Kouman Vincent
Centre Suisse de Recherches Scientifiques en Côte d'Ivoire (CSRS)

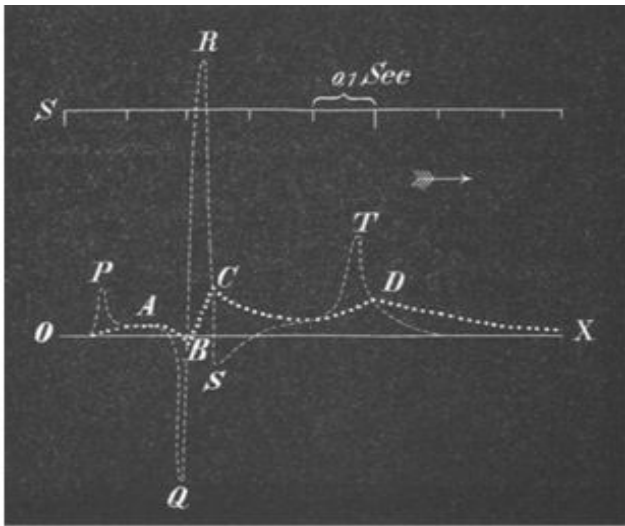


Fig. 1. Two Superimposed ECGs [3]

automated systems to accurately identify and classify ECG wave components. The accurate performance of these systems depends on several important factors, which include the quality of the ECG signal, the applied classification rule, the learning used. The first deflection, termed the P-wave has been created for the depolarization of the atria. The large QRS-complex has been created for the depolarization the ventricles [Majd AlGhatrif]. During an arrhythmia, the heart beat can too fast, too slow, or with an irregular rhythm. Myocardial ischemia is caused by a lack of sufficient blood flow to the contractile cells and many leads to myocardial information with its severe sequel of heart failure, arrhythmias and death [J Community Hosp Intern Med Perspect, 2012].

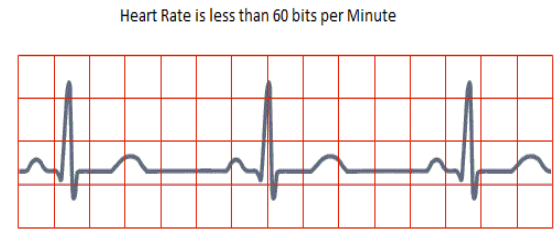


Fig. 3. Bradycardia

For adult if heart rate is more than 100 beats per minute that represent Tachycardia condition. It may be dangerous depending on the speed and type of rhythm.

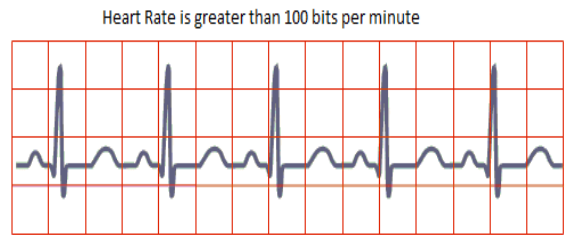


Fig. 4. Tachycardia

System Model: In order to extract information from the ECG signal, the raw ECG signal should be processed. An ECG signal contaminated with noise. The noises have been removed by digital band pass filter. We implement R-peak detection algorithm by using Pan-Tomkins algorithm. First of all there is a band-pass filter which is consisted of low pass-filter and high-pass filter. It can remove the noise. After that derivative has been used to get the slope. Then a squaring function has been used to get the positive point and then the signal is passed through the moving window integrator. After that threshold method has been done to locate the R-peaks [http://eleceng.dit.ie/dorran/matlab/ecg.txt].

Data Acquisition: A set of ECG records obtained from the data file [Mehta, 2008]. Then we have stored it in a text file. Its sampling rate 100Hz. It is contaminated with noise. It can be used for representing ECG signal. We can extract feature from the signal. Pan-Tompkins algorithm is used for higher accuracy for various beats than other traditional real-time methods. Various steps of Pan-Tompkins algorithm is shown below:

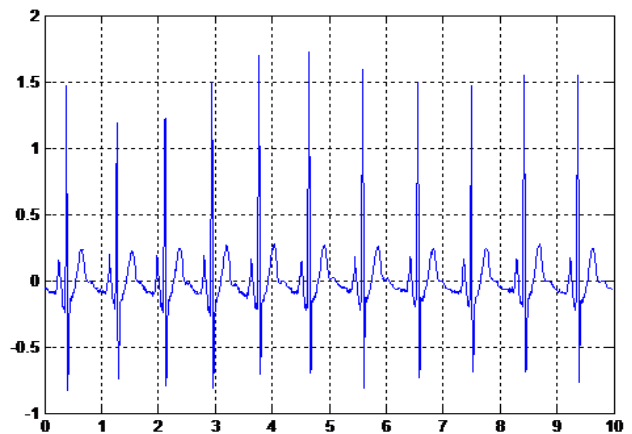


Fig. 6. Raw ECG Signal

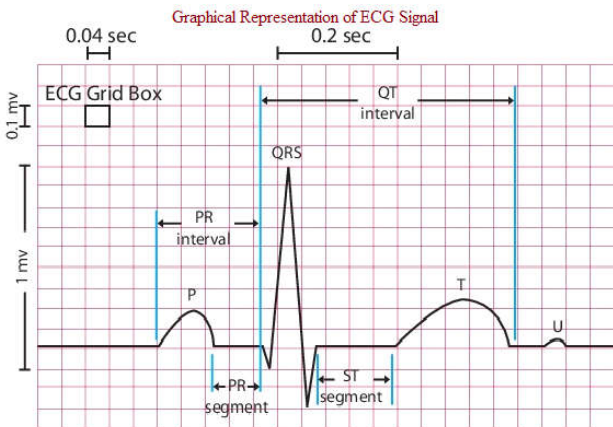


Fig. 2. Typical ECG Signal

There are two basic kinds of arrhythmias. They are

- Bradycardia
- Tachycardia

Bradycardia: Bradycardia is a slow heart rate when the heart rate is too slow that means less than 60 beats per minute. In this situation sufficient oxygen is not pumped into heart. It may sometimes cause death.

Tachycardia: Tachycardia is a fast heart rate when the heart rate is too fast that is more than 100 beats per minute [Mehta, 2008].

The signal length is 6000 samples, sampling rate is 100Hz. The following input figure (Fig. 6) for 1000 samples that have been plotted as:

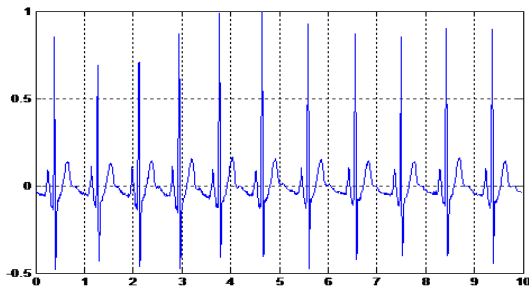


Fig. 7. After Normalization of ECG Signal.

In Fig. 7 we have got maximum value at most 1. We have found the following figure:

Band pass Filter: This band pass filter is constructed by low pass filter and high pass filter in cascade. The band pass filter can be used to reduce the baseline wander interference, muscle noise and T-wave interference [Kritika Bawa, 2014].

Low-pass Filter: A low-pass filter passes signals with a frequency lower than a particular cutoff frequency and attenuates signals with frequencies greater than the cutoff frequency. The actual frequency effect on the filter relies upon the filter design. Its cutoff frequency is about 7Hz.

The transfer for 2nd order low pass filter is given by:

$$H(x) =$$

And the functional equation for low pass filter is:

$$y[n] = 2y[n-1] - y[n-2] + x[n] - 2x[n-6] + x[n-12]$$

In Fig. 8 for low-pass filter b, a and zeros array of function have been used by

$$h = \text{filter}(b, a, [1 \text{ zeros}(1, 12)])$$

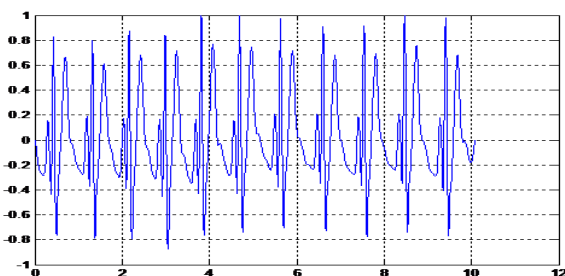


Fig. 8. Low-Pass Filter after Normalization of ECG Signal.

High pass filters remove low frequency signals and allow only high frequency signals to pass. A high-pass filter is an electronic filter which passes signals with a frequency greater than a certain cutoff frequency and attenuates signals with frequencies reduce than the cutoff frequency. Its cutoff frequency is about 5Hz. The transfer for 2nd order low pass filter is given by:

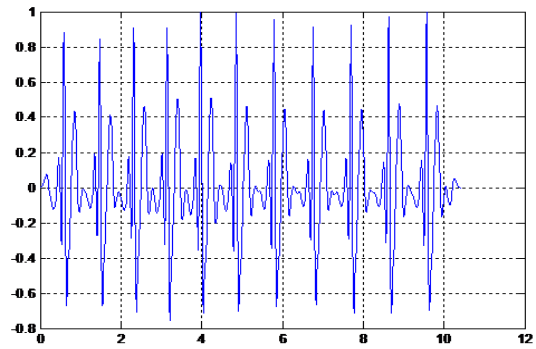


Fig. 9. High-Pass Filter after Normalization of ECG Signal.

$$H(x) =$$

In Fig. 9 for high-pass filter b, a and zeros array of function have been used by $h1 = \text{filter}(b, a, [1 \text{ zeros}(1, 32)])$.

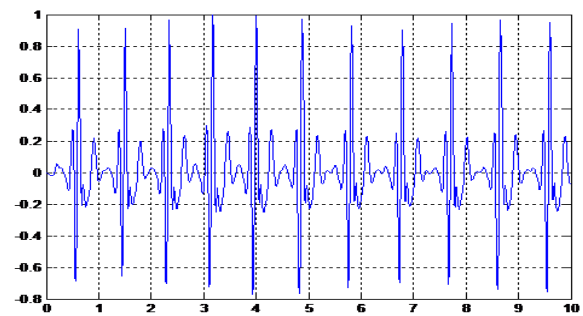


Fig. 10. Differentiator after Normalization of ECG Signal

Derivative Filter: After filtering, the filtered signal is derivative i.e. differentiated is used to get the slope information of the input signal. In this we have been used five -point derivative with the transfer function

$$H(x) = (-x^{-2} - 2x^{-1} + 2x^1 + x^2)$$

In Fig. 10 for differentiator, differential parameter has used h2. We have used convolution function with h2 and the output of normalization of high-pass filter (Fig. 9). In this process delay are 2 samples. Calculate the absolute value of the convolution result and have found the maximum value from convolution result.

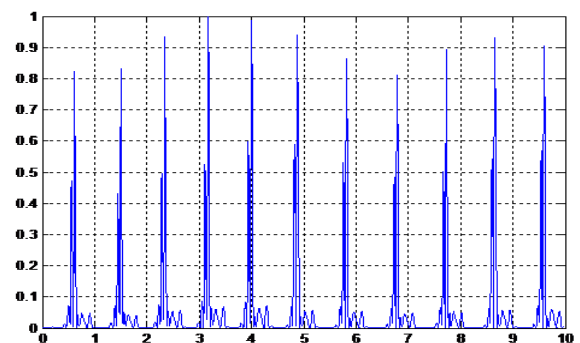


Fig. 11. Squaring Function after Normalization of ECG Signal.

Squaring Function: Squaring function can be used to get positive values from the negative or positive values. It also used to fraction value into smaller value and integer value into bigger. After differentiation, point by point squaring is done on the signal. The equation of this operation is

$$Y(nT)=$$

It highlights the higher frequencies in the signal which are mainly due to the QRS complex. In Fig. 11 for squaring function, square the output of the normalization from the differentiator (Fig. 10).

Moving Window Integration: The moving filter smoothing the data by replacing each data point with the given number of neighbors. The objective of moving window integration is to obtain wave form feature information in addition to the slope of the R wave. It can be calculated from

$$Y(nT)=$$

Where N is the number of samples in the width of the integration window. The number of samples N in the moving window is essential. Normally, the wide of the window should be approximately the same as the widest possible QRS complex. The QRS complex corresponds to the rising edge of the integration waveform. In this delay are 15 samples.

In Fig. 12 for moving window integrator ones array of function h3 is used. We have used convolution function with the h3 and output of normalization of squaring function (Fig. 11).

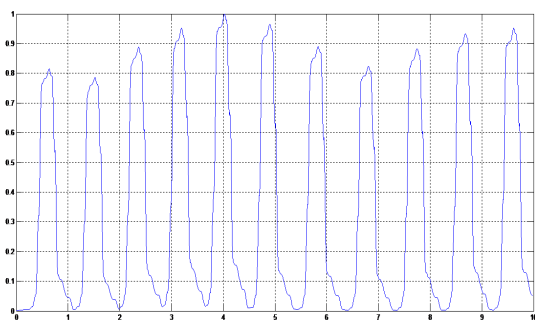


Fig. 12. Moving Window Integrator after Normalization of ECG Signal

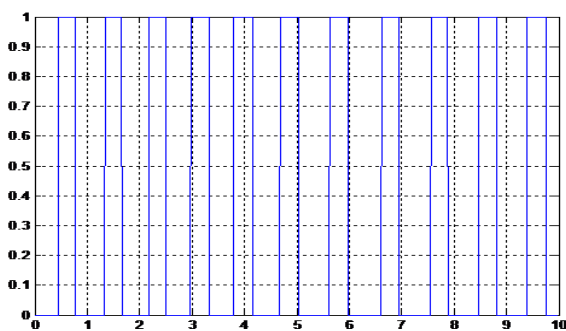


Fig. 13. Output of Pulse Detection.

Threshold Method: The minimum amount of energy required for detection of a stimulus is called threshold. It can be used to pass a value that treated as peak. After moving window integration, threshold method is used to find the R-peak. A maximum level is set which helps in detecting R-peak. QRS complex is also called fiducial marks. In Fig. 13 we have calculated the threshold first we have calculated maximum value from the normalization of the moving window integrator (Fig. 12).

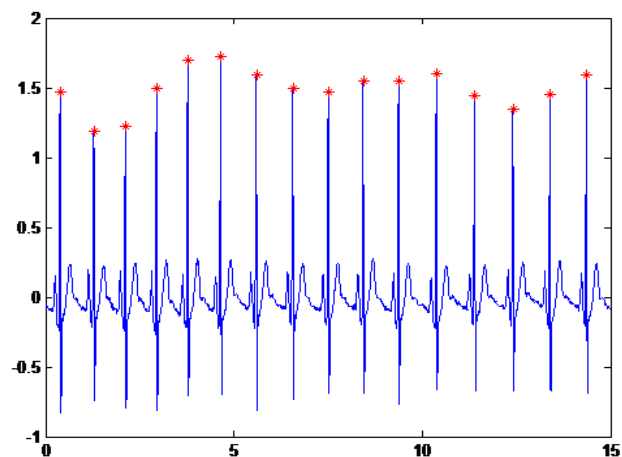


Fig. 14. R-peak Detection.

Heart Rate can be calculated from the above formula: In Fig. 14 R-peak has been detected. Then we have calculated R-R interval. R-R interval has been calculated between two R-peaks and average R-R interval has calculated from the total number of R-peak.

RESULTS

For the samples of 1000: We have got

Total Number of R- Peak:11.

Average R Wave Amplitude:1.4950.

Heart Rate is:64.5161 bits/min.

Decision: The Subject is suffering from Bradycardia.

Conclusion and Future Work: ECG is an important tool in the diagnosis and analysis of the heart activity. Heart rate can be calculated from the ECG signal. Physiologist can measure functional status of the heart from ECG signal that depends on different waves and intervals. In previous work has shown that using of one or more filtering methods results in more accuracy in R-peak detection. We have proposed Pan-Tomkins algorithm for developing a system for feature extraction from ECG signal. In this research paper, to classify the ECG heart rate from the data file using Pan-Tomkins algorithm. The Pan-Tomkins algorithm is signal dependent in the sense of fiducial points detection.

The threshold used for R-peak detection. After R-peak detection, average R-R interval has calculated and heart rate has been calculated. After calculating heart rate we can identify heart arrhythmias. The R-peak identified essential for classification, analysis and arrhythmia detection including Tachycardia and Bradycardia. The Pan-Tomkins algorithm gives more accuracy value than any other algorithms. Pan-Tomkins algorithm achieves accurate R peak detection and could be simply implemented with respect to other methodologies, but the squaring function might enlarge the noise. In this thesis paper, we have been used only one ECG data file. In future we will use multiple ECG data files. We have been calculated average amplitude of R-wave, total number of R-peak, R-R interval and heart rate. We will try to calculate P-wave, P-R interval, ST segment.

REFERENCES

- Autor Prace, Vedouci Prace and Doktorska Prace, "Methods For Detection And Classification In Ecg Analysis".
<https://ecglibrary.com/ecghist.html>.
- Majd Al Ghatri, Joseph Lindsay, "Methods For Detection And Classification In Ecg Analysis", J Community Hosp Intern Med Perspect. 2012; 2(1): 10.3402/jchimp.v2i1.14383,2012 Apr 30.
- Mehta, S. S. 2008. IAENG Member and N. S. Lingayat, IAENG Member, "Detection of P and T-waves in Electrocardiogram", Proceedings of the World Congress on Engineering and Computer Science, ISBN: 978-988-98671-0-2.
- <http://eleceng.dit.ie/dorran/matlab/ecg.txt>.
- Kritika Bawa, Pooja Sabherwal, 2014. "R-Peak Detection by Modified Pan-Tompkins Algorithm "International Journal of Advancements in Research & Technology, Volume 3, Issue5 , May-30 ISSN 2278-7763.
- Jiapu Pan and Willis J. 1985. Tompkins, a real-time QRS detection algorithm, IEEE transactions on biomedical engineering, vol. bme-32, no. 3, march.
