Report on assignement 1 of Data Capture and Processinig course

Hamza Omari[1*]

[1*]faculty of science and technology, University Paris est Créteil, Créteil, 94100, France.

Corresponding author(s). E-mail(s): hamza.omari@etu.u-pec.fr;

Abstract

Ecg signal processing is very important in detremining key features for patient diagnosis. In this report we will explore various techniques for implementing methods to extract key features. from using a public ecg database to computing basic signal characterstics, exploring some methods for R peak detection and some final metrics that help in dagnosis

Keywords: Signal processing, ECG, R peak detection

1 Introduction

The electrocardiogram (ECG) is widely used for monitoring the electrical activity of the heart. providing critical features for diagnosis. It is an electrogram of the heart which is a graph of voltage versus time of the electrical activity of the heart[1]. the normal ecg signal composes of the P,Q,R,S and T waves. As show in the figure 1 It is interesting for example to detect the R peaks, that will lead us for calculating the BPM of the heart, and other factors that will explain in details in the next sections. In our work we will work with the famous MIT-BIH Arrhythmia Database from physionet [2] [3]

The database contains a lot of files of recordings, for simplicity reasons, in this report, we will plot only a portion of the signal. And we'll be working with a certain random record file. sine all the functions can be applied to the other files.

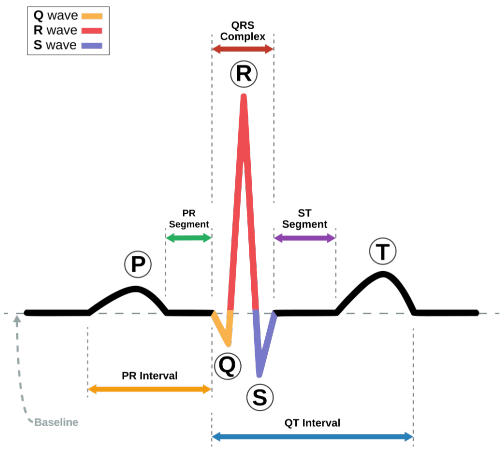

Fig. 1: Normal ECG signal [1]

2 Basic signal characteristics

In this section we will calculate some characteristics of the signal. First let us plot a portion of the signal and see how it looks

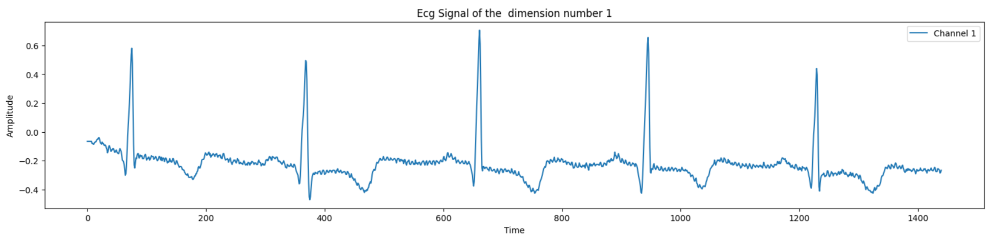

Fig. 2: Portion of the ecg signall [1]

we will be calculating the mean, energy, variance and the signal to noise ratio given these formulas respectively :

$$mean = \frac{1}{N}\sum_{i=1}^{n} x_i \tag{1}$$

$$energy = \sum_{i=1}^{n} x_i^2 \tag{2}$$

$$var = \frac{1}{N}\sum_{i=1}^{n} (x_i - \mu)^2 \tag{3}$$

$$SNR = \frac{\mu}{\sigma} \quad (4)$$

Where:

- μ is the mean of the signal.
- σ is the standard deviation of the noise.

After coding and computing these characterisitcs we got the values represented in Table 1

Table 1: Signal characteristics values.

mean	-0.3062989769230769
energy	85244.33212499997
var	0.037326063082030236
SNR	-1.5854021885079466

3 R peak detection

In this section we will discuss the methods we used for finding R peaks

3.1 Thresholding

Based on the article [4] The thresholding method fnids teh R peak. By setting a certain threshold, if the signal is above this threshold, then we should locate the R peaks from other peaks. It means we have some candidate points that can be an R peaks. The condition that satisfies: a point s r peak, if only he has values greater than it's neighberhood.

After applying the thresholding method we get these candidate points, in fgure 3

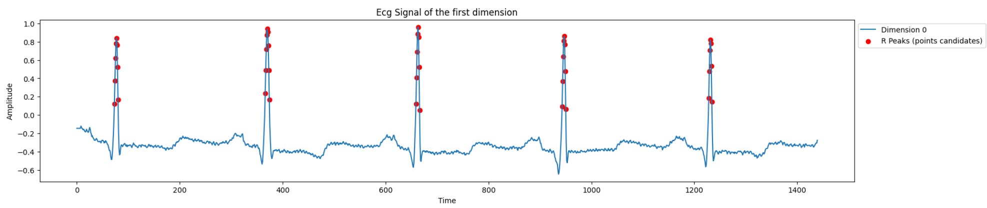

Fig. 3: Candidate R peaks after thresholdiing

As we can see, we have multiple red points(candidate points) but we need to filter them to keep only the real R peaks. as show in figure 4

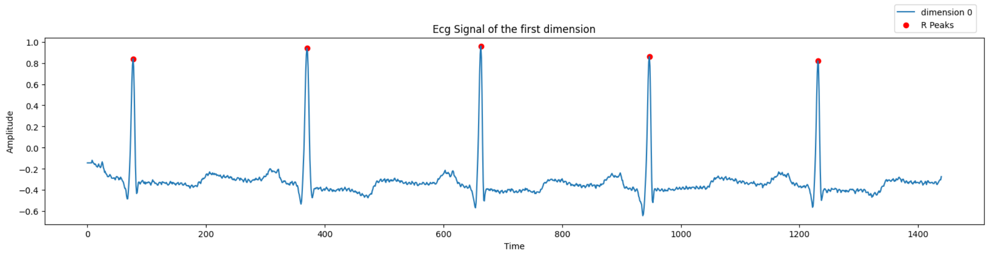

Fig. 4: R peaks

3.2 Pan Tompkin's algorithm

Based on this medium article [5] the algorithm follows a lot of steps to fnid the R peaks as displayed in figure 5

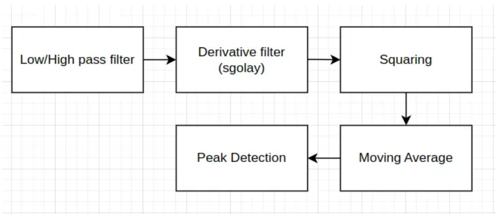

Fig. 5: Pan Tompkin's algorithm

first we would like to keep only a certain range of frequencis in our signal, to remove the noise and smoothen it. let it be between 5 and 15 Hz this first step gives us the following signal in figure 6

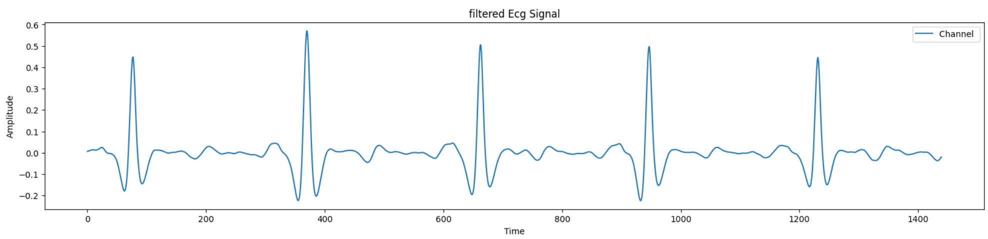

Fig. 6: Ecg signal after low/high pass filtering

After the low/high pass filtering we will also apply a derivative filter, to capture the high sudden variations, that typically occurs in the R peaks like. as show in the figure 7

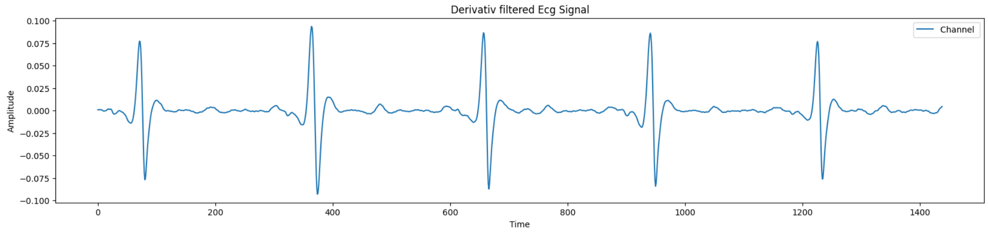

Fig. 7: signal after derivative filtering

The next step involves in squaring the signal, this will make our signal positive, the high variations we got in the previous step in the negative way, will be flipped. and make it more tight. As shown in figure 8

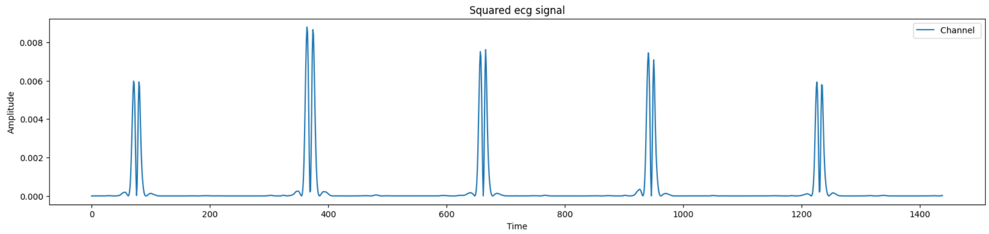

Fig. 8: signal after squaring

The last step in the algorithm will be applying a moving average window, this will merge the two consecutive peaks, in a one smooth peak, that will make it easier for us to find the corresponding R peaks. Result in figure 9

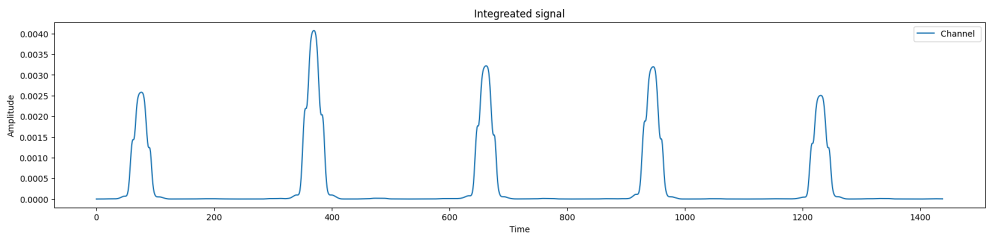

Fig. 9: signal after moving average

the last step involves finding the peaks in the last output signal. And they should correspond to the R peaks in the origianal signal. The result in 10 shows that actually we didn't got the real R peaks, meaning that those retuned indexes in the last step don't actually coincide with the real R peaks indexes. For this we will need to add another step to make a simple verification and locally search for finding the real peaks in our original signal.

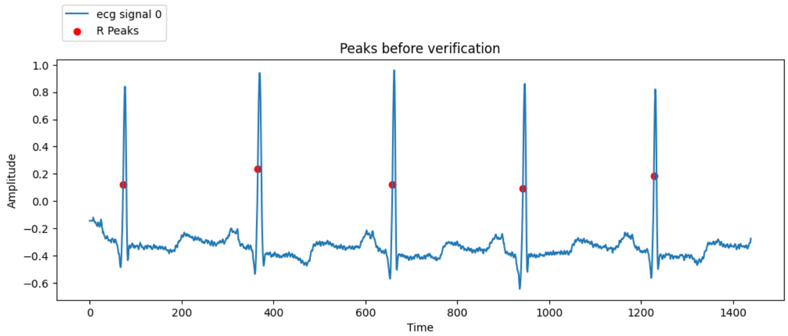

Fig. 10: R peaks before verification

After the verification and local search, we will get the desired R peaks, as shown in figure 11

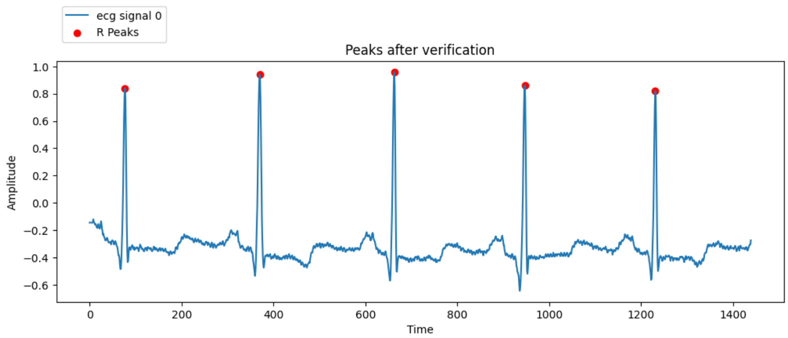

Fig. 11: R peaks after verfification and local search

4 Cardiovascular Parameter Calculation

In this section we will calculate some basic metrics, depending on the R peak detection result we get earlier. For Heart rate per Minute (BPM), RR Intervals and Heart Rate variability (HRV).

Heart rate per minute (BPM)
giving the sampling frequency fs the BPM will be counting the number of R peaks in $fs * 60$ period of time

RR intervals
After investingating the R peaks locations, it appears that the RR interval is not a fixed value between all two consecutives R peaks. Hence we will calculate the average time interval, between all the RR intervals in one minute.

It will be just the mean of the derivative of the R peaks. divided by the sampling frequency f_s as this formula below

$$RR = \frac{1}{f_s} \cdot \text{mean} * (\text{diff}\,(\text{array}\,(peaks)))$$

Heart rate variablility (HRV)
to calculate the HRV calculate the standard deviation of the first derivative of the peaks

$$\sigma = \text{std}\,(\text{diff}\,(\text{array}\,(peaks)))$$

Summary Results of the above metrics the table 2 shows the obtained results for the studied signal.

Table 2: Cardiovascular Parameters

BPM	73
average RR time	0.8235725308641975 s
HRV	36.35205674127652

based on [6] the results shows that the patient is in a healthy state, especially that BPM = 73 since it's between 60 - 100.

5 Code and implementation

All of the methods presented in this paper, are implemented in a jupyter notebook using python programming language, and libraries like numpy and scipy for numerical array manipulation and matplotlib for visualizations.

Notebook is accessible in this github repository: LINK

References

[1] contributors, W.: Electrocardiography. https://en.wikipedia.org/wiki/Electrocardiography. Last Accessed: 07/10/2024 (2024)

[2] Moody, G.B., Mark, R.G.: The impact of the mit-bih arrhythmia database. IEEE Eng in Med and Biol **20**(3), 45–50 (2001)

[3] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P.C., Mark, R., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation **101**(23), 215–220 (2000). [Online]

[4] Bittharia, K., Tiwari, P., Saxena, S.: R–peak detection of ecg signal using thresholding method. International Research Journal of Engineering and Technology (IRJET) **6**(4), 1–5 (2024). Last Accessed: 07/10/2024

[5] Rank, C.: Electrocardiogram QRS Detection Using Pan-Tompkins Algorithm and R Programming Language. Medium. Last Accessed: 07/10/2024 (2024). https://medium.com/@claucioranl/electrocardiogram-qrs-detection-using-pan-tompkins-algorithm-and-r-programming-language-c65c916e6a

[6] Edward R. Laskowski, M.D.: What's a normal resting heart rate? Mayo Clinic. Last Accessed: 9 October 2024 (2024). https://www.mayoclinic.org/healthy-lifestyle/fitness/expert-answers/heart-rate/faq-20057979