Detected Abnormal ECG Signals Utilising Wavelet Transform and Standard Deviation

Dejan Stantic and Jun Jo

Abstract—ECG contains very important clinical information about the cardiac activities of the heart. Often the ECG signal needs to be captured over a long period of time in order to identify abnormalities in certain situations. Such a signal apart from a large volume is often characterised by low quality due to the noise and other influences. In order to extract features in the ECG signal with time-varying characteristics at first need to be preprocessed with the best parameters. Also, it is useful to identify specific parts of the long lasting signal which have certain abnormalities and to direct the practitioner to those parts of the signal. In this work we present a method based on wavelet transform, standard deviation and variable threshold which achieves 100% accuracy in identifying the ECG signal peaks and heartbeat as well as identifying the standard deviation, providing a quick reference to abnormalities.

Keywords—Electrocardiogram-ECG, Arrhythmia, Signal Processing, Wavelet Transform, Standard Deviation

I. INTRODUCTION

MEDICAL technology improves with new innovations that make better diagnosis, but there is still the challenge to analyze a vast volume of low quality data. This data is most often obtained in the form of Electrocardiogram (ECG) signal.

In electrocardiography, an ECG artifact is used to indicate something that is not "heart-made." These include electrical interference by outside sources, electrical noise from elsewhere in the body, poor contact and machine malfunction. Artifacts are extremely common and knowledge of them is necessary to prevent misinterpretation of a heart’s rhythm.

An ECG captures the electrical signal within the heart. Traditional ECG signal is displayed on a graph where the X-axis is Time and the Y-axis can be either Voltage (mV) or Amplitude (dB) as can be seen in Figure 1. An ECG graph can be used for analysis of long intervals, measuring the consistency of the beats and looking for proper depolarisation, repolarisation and any irregular beats.

Fig. 1. Electrocardiogram (ECG) signal

In most people, a common abnormality comes within the heart where the heart skips a beat causing irregular heartbeat - arrhythmia. The symptoms of having a heart arrhythmia includes palpitations, dizziness, fainting, shortness of breath and chest discomfort. Palpitations have no obvious cause, but can be triggered by exercise, emotional stress, caffeine and nicotine, creating the feeling that your heart is racing, thumping or even skipping beats. An occasional palpitations that does not affect your general health is not usually something to worry about. However, if frequent or consistent palpitations occur, these signs could result in a serious arrhythmia. Arrhythmias can be diagnosed from the ECG signal, therefore, patients are provided with instruments which record their heartbeat for a long period of time. This results in a vast volume of data which often due to movements and other interactions has noise as well as needs preprocessing before any useful information can be obtained. Additionally, due to the very long duration of the signal, automatic methods to identify parts of specific interest are required.

Simply looking at a ECG graph, particularly when the signal can be hours long, is not enough to quickly and automatically make a diagnosis and also requires years of experience. To enhance the making of a diagnosis, different automatic methods have been proposed in the literature. Researchers addressed different aspects in relation to the ECG signal, proposing different filtering algorithms to improve the signal to noise value [6]. In order to achieve denoising of noisy ECG signal the wavelet theory was also utilised [4]. Different parameters have been considered, the wavelet function [10], threshold calculus [7], and level decomposition [1]. Additionally, the performance criteria to measure the quality of a wavelet, based on the principle of maximisation of variance [10], and the multi-resolution wavelet transform [14]. Apart of utilising wavelet techniques, there has been mentioned as well the usage of fast fourier transformations, neural networks [21] and other methods for improving ECG signal analysis. The wavelet found to be more precise over conventional FFT in finding the small abnormalities in the ECG signal.

Despite the significant attention in the literature which was directed toward ECG processing and automatic analysis, medical practitioners have not accepted the current findings as the definitive answers in making quick diagnosis towards arrhythmia patients, which clearly indicates that additional work needs to be done to provide more accurate diagnosis. In this work we present a method which can aid the practitioners by indicating which parts of the ECG signal can be suspicious and also provide guidance what could be symptoms.

The proposed method can be particularly useful when the ECG signal is recorded during a long period of time for patients with personal ECG interpretive equipment. At first
the signal is preprocessed to filter and eliminate any unwanted artifacts. Smoothing of the signal is achieved by utilising the Wavelet transform. In order to find optimum smoothing the parameters are tuned by analysing the ratio of integrals of the actual and smoothed signal over the window period.

Once the signal is preprocessed the heartbeat and the locations of the P, Q, R, S, T peaks can be identified. In the process of identification of the heartbeat and peaks in contrast to other works and in order to identify 100% accuracy, we introduced a variable threshold level, which is changing every window based on the actual level of the signal. To locate an abnormality in the ECG signal we calculate the standard deviation of the heartbeat and our system then closely analyses that part of the ECG signal and provides the practitioner means to view in detail those parts. Results obtained from our experiments, where we specifically looked into signals from the MIT-BIH database, which previously proposed methods struggled to achieve 100% identification, shows that our concept guarantees 100% identification but also is able to analyse long lasting ECG signal and emphasise parts of possible arrhythmias.

The reminder of the paper is organised as follows: In the next section we provide the necessary background information about ECG. Heart Arrhythmias, Wavelet Transform and Standard Deviation which in term will make it easier to follow this paper. In section 3 we will outline the related works that have been done in this field. In section 4 we present the core of our methodology and approach to our research, and provide the main findings. And finally, in section 5 we conclude the paper.

II. BACKGROUND

An ECG is a test that records the electrical activity of the heart. ECG is used to measure the rate and regularity of heartbeats as well as the size and position of the chambers, the presence of any damage to the heart, and the effects of drugs or devices (such as a pacemaker) used to regulate the heart. The word is derived from the Greek electro, because it is related to electrical activity, kardio, Greek for heart, and graph, a Greek root meaning "to write". In English speaking countries, medical professionals often use "EKG" (the abbreviation for the German word Elektrokardiogramm) in order to avoid audible confusion with "EEG," in emergency situations where background noise is high. Most EKGs are performed for medical purposes, but may also be performed on animals, usually for research.

ECG was first mentioned in 1872 when Alexander Muirhead reported to have attached wires to the patient’s wrist to obtain a record of the patient’s heartbeat. He recorded and visualised this activity by using a Lippmann capillary electrometer. However, the first systematic approach to the heart from an electrical point-of-view was Augustus Waller, while working in St Mary’s Hospital in London. His electrocardiograph machine consisted of a Lippmann capillary electrometer fixed to a projector. The trace from the heartbeat was projected onto a photographic plate, which allowed a heartbeat to be recorded in real time. However, there were little clinical applications for his work.

A. Arrhythmia

The heart is a vital organ. It is a muscle that pumps blood to all parts of the body. The blood pumped by the heart provides the body with the oxygen and nutrients it needs to function. Normally, this pumping (heartbeat) is controlled by the heart's electrical system. Sometimes the heart's electrical system may malfunction because of coronary heart disease, chemicals in your blood (including some medicines) or sometimes for no known reason. Changes in your heart's electrical system can cause abnormal heart rhythm, called arrhythmia. There are many different kinds of arrhythmias. Some may cause the heart to skip or add a beat now and again, but have no effect on the general health or ability to lead a normal life. While other arrhythmias are more serious and life-threatening. Untreated, they can affect the heart's pumping action, which can lead to dizzy spells, shortness of breath, faintness or serious heart problems and even be life-threatening medical emergency that can result in cardiac arrest.

B. Wavelet Transform

A wavelet is a wave-like oscillation with an amplitude that starts out at zero, increases, and then decreases back to zero. It can typically be visualised as a "brief oscillation" like one might see recorded by a seismograph or heart monitor. Generally, wavelets are purposefully crafted to have specific properties that make them useful for signal processing. Wavelets can be combined, using a "revert, shift, multiply and sum" technique called convolution, with portions of an unknown signal to extract information from.

The wavelet transform is also designed to address the problem of non-stationary ECG signals. It is derived from a single generating function called the mother wavelet, by translation and dilation operations. The main advantage of the wavelet transform is that it has a varying window size, being broad at low frequencies and narrow at high frequencies, thus leading to an optimal time-frequency resolution in all frequency ranges. The wavelet transform of a signal is the decomposition of the signal over a set of functions obtained after dilation and translation of an analysing wavelet. The ECG signal, consisting of many data points, can be compressed into a few features by performing spectral analysis of the signal with the wavelet transform. These features characterise the behaviour of the ECG signal. Using a smaller number of features to represent the ECG signal is particularly important for recognition and diagnostic purposes.

In the present work, in line with the literature, Daubechies wavelet is chosen [3].

C. Standard Deviation

In statistics and probability theory, standard deviation (represented by the symbol σ) shows how much variation or "dispersion" exists from the average (mean, or expected value). A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data points are spread out over a large range of values. The standard deviation of a random variable,
statistical population, data set, or probability distribution is the square root of its variance. The general formula for calculating the standard deviation of numbers $a_1, a_2, \cdots, a_n$ is:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} a_i^2}{n} - \left(\frac{\sum_{i=1}^{n} a_i}{n}\right)^2}$$  \hspace{1cm} (1)

### III. RELATED WORK

Computerised analysis of ECG has attracted significant attention from researchers. The purpose of such analysis is to determine any cardiac abnormalities a patient might have.

Significant attention in the literature was devoted to accurate analysis of ECG signal that would lead to cardiac abnormality diagnosis [8], [16]. In order to improve the quality of the ECG signal different filtering algorithms which improve the Signal to Noise values have been proposed [6]. Different types of noises can also prevent accurate analysis of the ECG signal and therefore extraction of useful information:

- Electromyogram (EMG) signal: this is the signal at high frequency and it is related to muscle activity;
- Baseline Wandering (BLW): it is a low frequency signal caused mainly by the breathing action;
- The electrode motion: it is usually represented by a sharp variation of the baseline.

One of the prominent methods for denoising of noisy ECG signals is the wavelet theory [4]. Specifically, several wavelet denoising algorithms have been proposed and each explore a particular parameter:

- the wavelet function [10]
- threshold calculus [7], and
- level decomposition [1].

Also, the wavelet packet method, which is a generalisation of wavelet decomposition, offers a rich range of possibilities for signal analysis. For example, the multi-resolution framework gives wavelets a very powerful compression and filtering tool [10], and characteristics of the time and frequency localisation of wavelets ensures the feature extractions [11].

Apart of filtering and preprocessing of the ECG signal there are works on detection of ECG using Fast Fourier Transformations and wavelet. For example, Karel proposed the performance criteria to measure the quality of a wavelet, based on the principle of maximisation of variance [10], [12]. Mahmoodabadi developed and evaluated an ECG feature extraction method based on the multi-resolution wavelet transform [14]. Frau presented a method to reduce the baseline wandering of an ECG signal [6]. Romero described a new ECG spectrum feature that allows the classification of dangerous arrhythmia by using the statistical methods [18]. Shantha discussed the design of good wavelet for cardiac signal from the perspective of orthogonal filter banks [19]. Nikolaev and Gotchev proposed a two-stage algorithm for electrocardiographic (EGG) signal denoising with Wiener filtering in the translation-invariant wavelet domain [17].

Most of the works mentioned in the literature focused on abnormalities with respect to extreme noisy signals using conventional Fast Fourier Transformations and wavelet methods. Most of the clinically useful information in the ECG is found in the intervals and amplitudes defined by its features (characteristic wave peaks, frequency components and time duration).

Attention in the literature was also given to the online beat segmentation and classification of electrocardiograms. The Hidden Markov Model (HMM) framework has been utilised because of its ability of beat detection, segmentation and classification, and has been mentioned as highly suitable to the ECG problem. By using HMM authors were able to look into topics which has not had much attention before, such as waveform modeling, multichannel beat segmentation and classification, and unsupervised adaptation to the patients ECG. The performance was evaluated on the two-channel QT database in terms of waveform segmentation precision, beat detection and classification. Results from the study indicates that the framework with HMM performs better to other systems mentioned in the literature with regard to waveform segmentation. They also obtained high beat detection performance with sensitivity of 99.79% and a positive predictively of 99.96%, using a test set of 59 recordings. Additionally, premature ventricular contraction beats were detected using an original classification strategy [2].

Before the actual analysis, preprocessing stages of beat detection and feature extraction need to be completed. Therefore there is a lot of attention in the literature in relation to the preparation and extraction of important and relevant information from the patients ECG signal. Given that the overall ECG process is dependent on the preprocessing stage therefore the preprocessing is still a very important area of the research. ECG classification systems have the potential to benefit from the inclusion of the automated measurement capabilities. The first stage in computerised processing of the ECG is beat detection. The accuracy of the beat detector is very important for the overall system performance hence there is a benefit in improving the accuracy. It has been shown that the concept of Discrete Wavelet Transform which is suitable for the non-stationary ECG signals. Baseline wander and different types of noise elimination are considered as a

![Fig. 2. Types of Daubechies Wavelet Transforms](image-url)
classical problem in ECG analysis, Joshi proposed a wavelet based search algorithm in different scales for denoising and subtractive procedure to isolate the baseline wander from the noisy ECG signal. This algorithm is tested using the data record from the MIT-BIH database and excellent results are obtained [9].

Detection of ECG characteristic points has also attracted the attention from researchers. As a result an algorithm based on wavelet transforms has been proposed which with the multiscale feature of wavelet transforms, the QRS complex can be distinguished from high P or T waves, noise, baseline drift, and artifacts. The relation between the characteristic points of the ECG signal and those of modulus maximum pairs of its wavelet transforms is illustrated. Results indicates that by using this method, the detection rate of QRS complexes is above 99.8 percent for the MIT/BIH database, and also the P and T waves can be detected, even with significant baseline drift and noise [13].

Heartbeat signal analysis is widely used in the medicine and medical research area, such as energy expenditure measurement, autonomic nervous system assessment, sports research, etc. Physical activities are commonly recognised to greatly affect the changes of the heartbeat. However, the direct relationship between the heartbeat and physical activities is hard to describe. Therefore, the prediction of heartbeat as a factor of physical activity also attracted attention from researchers. For such prediction neural network has been utilised. Experiments were conducted based on the real-life signals from a healthy male. The mean absolute error of the predicted heartbeat was relatively small [5],[21].

We have used wavelet transform to decompose the signal as it has been shown that the wavelet transform performs good in signal processing [20]. We investigated which particular Wavelet would achieve the best performance with regard to the smoothing and baseline drift removal. We considered the family of Daubechies wavelets, and in Figure 2 we show several Daubechies Wavelets. In Figure 4 we show results that we obtained by calculating the Integrals of the signal after the smoothing and baseline drift removal. It is important to mention that the Integral value for the Actual signal was 4906.6. We have taken into consideration Daubechies Wavelets 1 to 20, and we also varied the decomposition level from 2 to 12.

As it can be seen in Figure 4 Daubechies wavelets 'db14' performed the best and achieved the lowest integral number, which means it smoothed and removed the baseline drift of the signal the most.

We also investigated which specific span would achieve the lowest Integral. In figure 5 we show the data obtained after calculation of the integral for the different Daubechies Wavelets (1 - 20) and for the different spans. We considered spans of 30, 50, 60, 80, 100. We previously identified that the decomposition level of 12 was the best in this test. It can be seen that the span of 60 performs the best and has the lowest Integral. Based on the experiment conducted we concluded that the best performance can be achieved with
Daubechies Wavelets 'db14' (Figure 7) with a decomposition level of 12 and with the span of 60. All other experiments in the reminder of this paper we performed with the above mentioned parameters. If Figures 4 and 6 we show to what extent the baseline drift was removed and how much the signal was denoised.

Every particular window we calculate the mean value of the maximum signals, which ensures that the identification of the peaks in the ECG is correct. Based on careful preprocessing of the signal, with the best wavelet transform and optimal parameters, and by introducing the variable threshold we obtained a 100% identification of the peaks in the signal where other state-of-the-art methods were unsuccessful.

Once we correctly identified all the peaks in the ECG signal we could automatically analyse actual information it holds. This is particularly important in case when the signal has a long duration such as the one obtained from equipment attached to the person for 24 hours. It is obvious that the medical practitioner would waste time to go through such a long signal. To automatically identify abnormalities we utilised the standard deviation of the heartbeat, which could be a good indication of any problem. In Figure 10 we show a graph where it can be clearly seen parts of the signal where there are significant deviation in the heartbeat. A Medical practitioner would then be directed to the part where the standard deviation of the heartbeat is bigger than the chosen threshold. Those parts of the signal then can be viewed in detail, as shown in Figure 11.
Electrocardiograms are very widely used as an inexpensive and noninvasive means of observing the physiology of the heart and it contains very important clinical information about the cardiac activities of the heart. Sometimes to make a correct diagnosis it is required to capture and analyse an ECG signal over a long period of time with personal ECG interpretive equipment. Analyse of such a signal is a challenge not only due to the vast volume of data but also the long lasting signal often captured in different conditions contain a significant amount of artifacts. In this work we addressed a problem of the field:

• We introduced the variable threshold which ensures that the signal is smoothed and the baseline is removed the most.
• We also concluded that the span of 60 achieves the lowest level of Integral, which means that the signal is smoothed and the baseline is removed the most.
• We introduced the variable threshold which ensures that the method is applicable for different situations as it is adjusting to the activity of the person.
• We introduced the standard deviation of the heartbeat, which provides quick reference to abnormalities.

V. CONCLUSION

We evaluated different Daubechies Wavelets and identified that the 'db14' perform the best with regard to the baseline drift removal and smoothing.
• We identified that the decomposition level of 12 has the best performance.
• We also concluded that the span of 60 achieves the lowest level of Integral, which means that the signal is smoothed and the baseline is removed the most.
• We introduced the variable threshold which ensures that the method is applicable for different situations as it is adjusting to the activity of the person.
• We introduced the standard deviation of the heartbeat, which provides quick reference to abnormalities.

REFERENCES

Dejan Stantic completed Bachelor of Multimedia with first class Honours. He previously graduated Popular Music and Digital Video Production where he got interested in sound and signal processing. His research interest focuses on signal formant analysis. Currently he is a PhD student where he extended his work on signal analysis to the health domain and electrocardiogram signals.

Dr. Jun Jo was awarded his PhD degree from the University of Sydney in 1994. He has conducted research in various areas including Computer Vision, Robotics and e-learning and has published over 100 refereed publications. Dr. Jo is currently taking the positions of the Director General of International Robot Olympiad Committee (IROC) and the President of Australian Robotics Association (ARA).