NOVEL TIME-FREQUENCY FEATURES FOR NORMAL ECG

*Nazeeh Alothmany[1], Rayan Kashkari[1], Ibrahim Shaibah[1], Ahad Masoud Ali Alarayani[1], Mehmet Akyurt[1]**

1. College of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia .

Abstract

Literature was reviewed on the current state of methods for detecting arrhythmias. A software package, Heartfelt, was developed that automatically detects arrhythmias from the lead II ECG waveforms. To this end time-frequency distributions for simulated normal and abnormal ECG waveforms were obtained. Then features representing the normal waveform were extracted and used to train a classifier. Finally the performance of the classifier in detecting arrhythmias was tested using unknown ECG waveforms.

The execution time of Heartfelt was found to be 0.92 seconds for each minute of ECG recording. It was determined that the accuracy was 96.7% for detecting normal ECG signals, 98.2% for detecting arrhythmic ECG signals suffering from arterial fibrillation, and 96.5% for detecting arrhythmic ECG signals suffering from superventricle arrhythmias.

Clinical article, Review (J Int Dent Med Res 2012; 5: (3), pp. 179-192)

Keywords: Arrhythmia, classifier, detect, ECG, software package, time-frequency distributions, features.

Received date: 19 June 2012 **Accept date:** 25 September 2012

Introduction

Electrical pulses in the heart are controlled by special groups of cells called nodes (see Fig. 1). The SA (sinoatrial) node, the pacemaker of the heart, generates an electrical signal that causes the upper heart chambers (atria) to contract; the signal then passes through the AV (atrioventricular) node, which delays the signal on its way to the lower heart chambers (ventricles), allowing them to fill up and causing them to contract.[1]

Arrhythmias are heart-rhythm problems; they occur when the electrical impulses that coordinate heartbeats are not working properly. An arrhythmia is an irregular heartbeat. The heart may beat too fast (tachycardia), or too slow (bradycardia), or ventricles may contract before atrais (premature contraction) or heart beats become irregular (fibrillation).[2]

***Corresponding author:**

Dr. Mehmet Akyurt,

College of Engineering, King Abdulaziz University,
Jeddah, Saudi Arabia

E-mail: makyurt@kau.edu.sa; akyurt99@yahoo.com

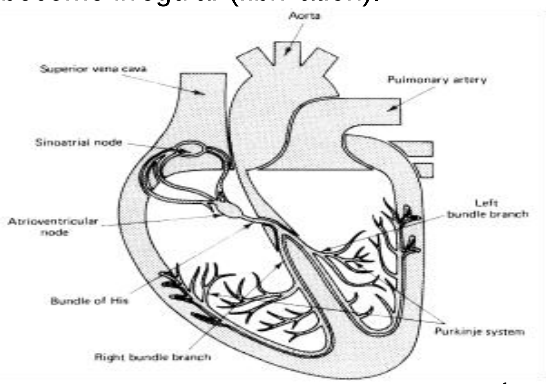


Figure 1. Electrical Model of the Heart[1]

Many heart arrhythmias are harmless. Everyone occasionally experiences irregular heartbeats, which may feel like a heart racing or fluttering. However, some arrhythmias can be harmful. Cardiac patients who suffer from coronary artery disease, heart valve defects, heart damage, and many other heart related issues, may have potentially fatal symptoms. Arrhythmia may occur at any time during normal daily routine (i.e. at home, at work, while driving, while sleeping, and while working-out), which makes life very hard for cardiac patients.

Electro-Cardiogram (ECG) is the heart's electrical signal. It can be used to monitor the functionality of the heart. Figure 2 shows a typical ECG waveform for a healthy patient.

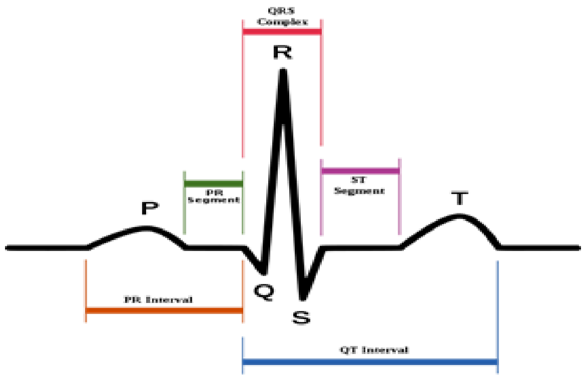


Figure 2. Typical ECG waveform with its components[1,26]

Physicians are not always available to diagnose these waveforms and to judge the functionality of the heart. In addition, a patient in the emergency room may be having chest pains for reasons not related to heart disease. Therefore, it would be more efficient for cardiologist if there is a method that can detect arrhythmias in these waveforms so that a cardiologist is called urgently.

Physicians diagnose the functionality of the heart from reading the different amplitudes and lengths of different time segments. It seems that little work has been done so far in understanding the frequency contents of the ECG signal. Since time-frequency algorithms observe signals in both time and frequency, their analysis provides a different view for signals, possibly providing additional insight over the ECG waveform to help the physicians gain a deeper view on the functioning of the heart.

II. Review of Literature

Heart diseases are caused by a multitude of reasons including abnormal propagation of pacing impulses through the specialized cardiac conduction system. Such abnormalities where cardiac rhythm deviates from normal sinus rhythm are termed as arrhythmia.[3]

Over 7 million people worldwide die annually from erratic heart rhythms (cardiac arrhythmias), and many more are disabled.[4] Citing that standard diagnostic techniques such as the electrocardiogram (ECG) provide only low-resolution projections of cardiac electrical activity on the body surface, the authors introduced a multielectrode vest that records 224 body-surface electrocardiograms; electrical potentials, electrograms and isochrones which are then reconstructed on the heart's surface using geometrical information from computed tomography and a mathematical algorithm.

Ventricular tachyarrhythmias, in particular ventricular fibrillation (VF), are the primary arrhythmic events in the majority of patients suffering from sudden cardiac death.[5] Attention has focused upon these articular rhythms as it is recognized that prompt therapy can lead to a successful outcome. Arrhythmia is one kind of disease that gives rise to death and possibly forms the immedicable danger. The most common cardiac arrhythmia is the ventricular premature beat.

Heikki and others remarked that, since the recognition of the high incidence of cardiac arrest as the mechanism of sudden cardiac death (SCD), medical scientists and clinicians have sought methods to predict and prevent these events.[6] Significant progress has already been made in the prediction and prevention of life-threatening arrhythmias during the last decade. The authors further remarked that recent studies have demonstrated that the implantable cardioverter defibrillator provides a mortality benefit compared with conventional drug therapy.

Moody and Mark observed that the MIT-BIH Arrhythmia Database was the first generally available set of standard test material for evaluation of arrhythmia detectors, and it has been used for that purpose as well as for basic research into cardiac dynamics at about 500 sites worldwide since 1980.[7] It has lived a far longer life than any of its creators ever expected. Together with the American Heart Association Database, it played an interesting role in stimulating manufacturers of arrhythmia analyzers to compete on the basis of objectively measurable performance, and much of the current appreciation of the value of common databases, both for basic research and for medical device development and evaluation, can be attributed to this experience. The authors briefly review the history of the database, describe its contents, discuss what has been learned about database design and construction, and also take a look at some of the later projects that have been stimulated by both the successes and the limitations of the MIT-BIH Arrhythmia Database.

Pfeiffer, Kenner and Schaefer applied statistical methods for the analysis or interval related cardiac performance variations during cardiac arrhythmia.[8] Specifically, they studied the influence of the sequence of stimulation intervals on cardiac performance indices the relationship between properties of succeeding arrhythmic aortic pressure pulses of patients with atrial fibrillation has been analyzed by methods of correlation and regression analysis. They noted that changes in the normal rhythm of a human heart may result in different cardiac arrhythmias, which may be immediately fatal or cause irreparable damage to the heart sustained over long periods of time. They reasoned that the ability to automatically identify arrhythmias from ECG recordings is important for clinical diagnosis and treatment.

This study was followed by an analysis by Coast and associates where they describe a new approach to ECG arrhythmia analysis based on hidden Markov modeling.[9] At about the same time, Safe and Maxwell studied data stored by a transtelephonic electrocardiographic monitoring system for evaluation of episodic symptoms suggestive of cardiac arrhythmia over an 18 month period.[10] Their study showed that transtelephonic electrocardiographic monitoring is useful in documentation of infrequent or sporadic episodes directly related to symptoms when 24-hour ambulatory electrocardiograph monitoring is normal. Szi-Wen and associates suggested a sequential detection algorithm for cardiac arrhythmia classification.[11] Guvenir and associates introduced a supervised machine learning algorithm for the diagnosis of cardiac arrhythmia from standard 12 lead ECG recordings.[12]

In order to try to detect and prevent the occurrence of sudden cardiac arrhythmia and death, detection systems have been developed, and implantable stimulation devices have been used. One such system was patented by McClure et al.[13] Francis et al. defined heart rate turbulence as initial acceleration and a subsequent deceleration of sinus rhythm following a ventricular ectopic beat with a compensatory pause.[14] They also suggested a cardiac arrhythmia detection system for an implantable stimulation device.

Throne and associates obtained a patent where they describe a cardiac monitor that monitors the condition of the heart of a cardiac patient and generates signals indicating one of several conditions, such as supraventricular tachycardia, ventricular tachycardia and ventricular fibrillation.[15] In order to generate these signals, the ECG from the patient is analyzed to determine a cardiac interval and heart rate, as well as a waveform factor and a waveform factor irregularity. Chuang-Chien Chiu, Tong-Hong Lin and Ben-Yi Liau also described an algorithm for detection of arrhythmia that uses correlation coefficients in ECG waveform.[16] They suggested that their system is accurate and efficient to classify arrhythmias resulting from APC or PVC. The proposed arrhythmia detection algorithm, they suggested is therefore helpful to clinical diagnosis. Vijaya had good success with using Wavelets as basis for features to analyze the ECG.[17]

Mohammadzade and associates presented a support vector machine-based arrhythmia classification system using reduced features of the heart rate variability signal.[18] They claimed that the main advantage of their algorithm, compared to the approaches which use the ECG signal itself, is the fact that it is completely based on the HRV (R—R interval) signal which can be extracted from even a very noisy ECG signal.

Holter monitoring has been a widely publicized system for arrhythmia patients. After clinical trials, Kinlay and associates and Hardy and Fletcher concluded that cardiac event recorders yield more diagnoses and are more cost-effective than 48-hour Holter monitoring in patients with intermittent palpitations.[19,20]

Admittedly a variety of methods can be used in the outpatient evaluation of symptoms suggestive of a cardiac arrhythmia. The diagnostic yield of these technologies for identifying clinically significant but infrequent, brief, and/or intermittently symptomatic arrhythmias, however, is low. Holter monitoring for 24–48 hours is typically employed, but has a diagnostic yield of only 15–28% depending on symptoms and frequency.[21] External, patient-activated loop recorders can improve the diagnostic yield to up to 63%, but, it has been suggested that they require appropriate patient activation during the recurrence of symptoms, which can limit their usefulness.

Rothman and associates suggested that mobile cardiac outpatient telemetry (MCOT) allows patients to be monitored continuously for an extended period and has been effective in the diagnosis of clinically significant, symptomatic, and asymptomatic cardiac arrhythmias.[21] This technology seems to have the potential to reduce patient error, enhance diagnostic accuracy, decrease time to diagnosis, and improve patient care.

Computer-assisted cardiac arrhythmia recognition is critical for prompt and efficient management of cardiac disorders. To this end, Srinivasan, Wong and Krishnan proposed a new phase space analysis-based algorithm for electrocardiogram (ECG) signals to facilitate the detection of cardiac arrhythmia.[22] A phase space density plot was obtained by mapping the distribution of points in the phase space of ECG signals and the phase space density values within a predefined window were used for arrhythmia detection.

Realizing that automatic detection and classification of cardiac arrhythmias is important for diagnosis of cardiac abnormalities, Prasad and Sahambi proposed a method to accurately classify ECG arrhythmias through a combination of wavelets and artificial neural networks [23].

On the other hand, ventricular tachyarrhythmias, in particular ventricular fibrillation are considered as the primary arrhythmic events in the majority of patients suffering from sudden cardiac death. Attention has been focused upon these articular rhythms as it is recognized that prompt therapy can lead to a successful outcome. Khadra, Al-Fahoum and Binajjaj experimented with higher order spectral analysis in the classification of life threatening arrhythmias.[5]

Still other researchers tried other venues to tackle the problem. Thus Kirk and associates set to determine the predictive value of cardiac T2* magnetic resonance for heart failure and arrhythmia in thalassemia major.[24] They determined that using cardiac T2* for the early identification and treatment of patients at risk is a logical means of reducing the high burden of cardiac mortality in myocardial siderosis.

A very significant study was initiated by Khadra, Fraiwan and Shahab, which however was apparently later not pursued further.[25] The authors used the time-frequency wavelet theory in combination with neural networks for the detection of life threatening ECG arrhythmias. They achieved this through the application of the wavelet transform on the ECG data, and using the resulting data as input to a neural network, to perform the classification of the arrhythmia into one of three possible cases, namely ventricular fibrillation, ventricular tachycardia and atrial fibrillation.

The objective of the current study was to design a software package that automatically detects arrhythmias from the lead II ECG waveforms. To this end time-frequency distributions for simulated normal and abnormal ECG waveforms were obtained. Then features representing the normal waveform were extracted and used to train a classifier. Finally the performance of the classifier in detecting arrhythmias was tested using unknown ECG waveforms.

III. Design of a Detector

In the past, arrhythmia used to be discovered by looking for symptoms that indicate its existence (such as palpitation, fainting, and fatigue), or through physical exams (such as checking for swellings in the legs or feet, which could signal of an enlarged heart or heart failure). There are also some arrhythmias referred to as silent arrhythmias; these do not cause symptoms, making it difficult to detect the arrhythmia.

Nowadays, there are several devices that can be used to detect arrhythmias, as outlined in the Literature Review section. However, they only use real-time analysis, which totally ignores the frequency content of the heart's electrical signal. One of these is the Holter monitor, a device that records the heart's electrical signals continuously for 24 to 48 hours.[3] The recordings are then downloaded on a computer that analyzes the heart's electrical signals and detects arrhythmia. The Holter monitor detects arrhythmias that occur randomly. It can also detect silent arrhythmias. But the disadvantage is that arrhythmias may not occur during the test period.

Another device that is available is the Event Monitor. This device records the heart's electrical signals only when symptoms of arrhythmia occur. The monitor must be started by the user when he feels symptoms of arrhythmia.[3] The recordings are then downloaded on a computer that analyzes the heart's electrical signals and detects arrhythmia. The advantages of this device are that it works for short periods of time, which saves power and storage space required, and that it is small in size. The two main disadvantages of the Event Monitor are that a) it must be started manually by the user, and that b) it does not detect silent arrhythmias.

To conceive a new arrhythmia detector that combines the advantages of both of these devices, and yet does not have their disadvantages, one needs to undertake a time frequency analysis of ECG signals. To this end one needs to

a) Obtain digital representations of normal and abnormal lead II ECG signals to use them as input to the software.

b) Obtain the time frequency distributions (TFD) of the ECG signals in (a).

c) Compare the TFD of normal and abnormal ECG waveforms, and extract features representing the normal ECG signal.

d) Train a classifier to recognize the normal ECG, and finally

e) Test the performance of the classifier using unknown ECG waveforms.

It would be desirable for this automatic arrhythmia detector to accept ECG signals from a patient as well as from an ECG simulator. Such a device should have more than 90% accuracy in detecting arrhythmias. Other desirable features would be that the classifier must detect an arrhythmia rapidly, and that Heartfelt must be user friendly, executable under windows.

An automatic arrhythmia detector is a computer program, henceforth called Heartfelt, that will detect abnormalities in the ECG waveform. Features extracted from the time-frequency distribution of the lead II ECG waveform may provide additional insight over the ECG waveform to help the physicians to have a deeper look on the heart and improve the accuracy of automatic arrhythmia detection.

There are several ways arrhythmia may be detected:

a) Heart Sounds "Stethoscope": Heart sounds are the noises (sound) generated by the beating heart muscle, closure/opening of heart valves and the resulting flow of blood in/out of the heart. In cardiac auscultation, an examiner uses a stethoscope to listen to these sounds, and uses them to obtain important information on the condition of the heart and to detect arrhythmias.

b) Blood Pressure: "Sphygmomanometer": Blood pressure is a measurement of the force applied on the walls of the arteries as the heart pumps blood through the body. The pressure is determined by the force and amount of blood pumped, and the size and flexibility of the arteries. When arrhythmia occurs, less blood is pumped, thus blood pressure is disturbed and arrhythmia is detected.

c) Echocardiogram: An echocardiogram is a test that uses sound waves to create a moving picture of the heart. Echocardiography is useful in the diagnosis of fluid in the pericardium (the sac that surrounds the heart). And since the blood pumping of the heart is disturbed when arrhythmia occurs, it may serve as a way to detect arrhythmia.

d) Arrhythmia detector using time-domain analysis: The ECG signal is obtained in the time domain. So this method will only monitor the ECG waveform amplitude with respect to time, and detect any abnormalities by comparing the input ECG waveform to that of a healthy human being. In time domain, an ECG waveform is represented by the amplitude and duration of each time segment (P, P-R, QRS, and Q-T) [4]. This method ignores the frequency of the signal, which contains valuable information.

e) Arrhythmia detector using frequency-domain analysis: The ECG signal is obtained in the time domain, and then it is transformed into the frequency domain using a Fast Fourier Transform (FFT). Hence the amplitude of the signal will be displayed with respect to frequency, allowing the detection of any abnormalities in frequency content of the signal. Thus this method monitors the frequency of the signal, providing further insight information, but it ignores the timing of the signal, which is very critical and important. Also it cannot determine *when* the frequencies existed, displaying the signal from a single perspective.

f) Arrhythmia detector using time-frequency analysis: Time analysis and frequency analysis by themselves do not fully describe the nature of signals. From a spectrum an investigator knows which frequencies were present in the signal, but he cannot determine when those frequencies existed. A time frequency analysis, on the other hand, presents the ECG waveform simultaneously in time and frequency, providing an enhanced view of the ECG waveform, allowing the detection of any abnormalities in frequency as time changes.[6]

In conclusion, it is noted that time-frequency analysis takes advantage of both variables, time and frequency. This provides an additional insight over the heart's ECG signal, and allows one to fully describe it, enabling him to detect different abnormalities in the heart that cannot be detected using time analysis alone or frequency analysis alone.

The design of Heartfelt may be split into four major stages (see Figure 3). These involve the ECG simulator, the time frequency transformer, the feature extractor, and the classifier. These parts will be explained in the following:

Figure 3. Automatic arrhythmia detector design block diagram

a) The ECG Simulator generates normal and abnormal ECG waveforms. An ECG simulator will be obtained in MATLAB and the MIT-BIH database will be used for ECG recordings.[8, 27]

b) The Time Frequency Transform obtains the time frequency representation of the ECG signal generated by the ECG simulator. For this a time frequency transform function will be obtained also in MATLAB.

c) The Feature Extractor extracts features representing the ECG signal in time frequency domain, and the Classifier compares features of the input ECG signal to that of a normal ECG signal, and classifies it as normal or abnormal. Both the feature extractor and the classifier will be designed.

During validation, the selected software will be installed on a windows operating system. Then a normal ECG signal will be introduced. Next the time frequency distribution of this signal will be obtained, and the features representing the input signal will be extracted. It will be confirmed that the features representing the signal has countable values. Then the classifier will be run to confirm that the duration of detection is very short, say under 2 seconds.

Then the above procedure will be repeated for abnormal ECG recordings, and calculating each time the accuracy of detecting arrhythmias, and confirming that the accuracy of detecting arrhythmias is higher than about 90%.

Physical signals, including ECG signals, are usually analog, which means that they have a value at any given point in time, and that they are made from infinite number of points. Computers have finite memory, therefore, they cannot store or process analog signals. So instead of taking all these infinite values, values are recorded at fixed intervals of time. This is called Sampling, a technique that transforms an analog continuous signal into an analog discrete signal, which has values at certain points in time. By doing this, signals can be processed on computers.

Sampling must preserve signal properties and frequency contents. The sampling theory states that in order to preserve the essentials of an original signal, the sampling rate has to be more than twice the maximum frequency present in the signal.[7] Since in ECG signals, the maximum frequency is about 200 Hz, it is concluded that the sampling rate must be more than 400 samples per second.[4]

In some cases, ECG signals that are available may be sampled at insufficiently low rates, meaning that their sampling rate is less than twice their maximum frequencies. As a result some of the signal's frequency contents would be lost. So in order to increase the sampling rate, Interpolation is used by taking the average of each consecutive point in the signal and then adding it as a point between the two points. This also helps in unifying the sampling rate of the input signals since different databases in the MIT-BIH database have different sampling rates. For example, the normal ECG database has a sampling rate of 128 samples/sec, while the Superventricle Arrhythmia SV database has a sampling rate of 64 samples/sec.

Classification is the process of grouping objects together into classes according to their perceived likenesses or similarities.[30] Here classification is used to group ECG signals into two classes, normal and abnormal. Similarities between normal ECG signals are features that can be extracted and used to group them into one class. Classification requires the training of a classifier. The training process is basically feeding the classifier with normal ECG signals and choosing the features that characterize them.

Time frequency transform is done in MATLAB using its Spectrogram function. The function uses the Short Time Fast Fourier Transform STFFT method, which divides each and every signal into a number of time windows and calculates the Fourier transform for each window.[6] The output of this analysis is a 3D plot with time, frequency, and amplitude as its axes. The analysis requires choosing several parameters (window width, number of points in the fast Fourier transform, and whether windows overlap or not).

A classifier must classify signals that have similar features together in the same class. In some cases, a signal may get attenuated or amplified. It may also come at a slower or faster rate. A good classifier must be able to assign signals with modified amplitudes or modified rates to the same class. Therefore, time and amplitude normalization is needed before feeding a signal to the classifier. Signal normalization is a technique that unifies the shape of signals so that they can be compared to each other.

There are two types of normalization, Time Normalization, and Amplitude Normalization. The objective here is to detect arrhythmias on ECG signals. The classifier needs to be able to assign all normal ECGs to the same class. In this context, ECGs may be recorded at different rates and at different amplitudes. Ordinarily signals with different number of windows cannot be compared to each other. So in order to unify the number of windows per each signal, time normalization removes the effect of heart rate changes on the ECG signal by making the size of each window dependent on the size of each signal, thus ensuring that all signals are divided into the same number of windows. Figure 4 and Figure 5 show two signals from different heart rates.

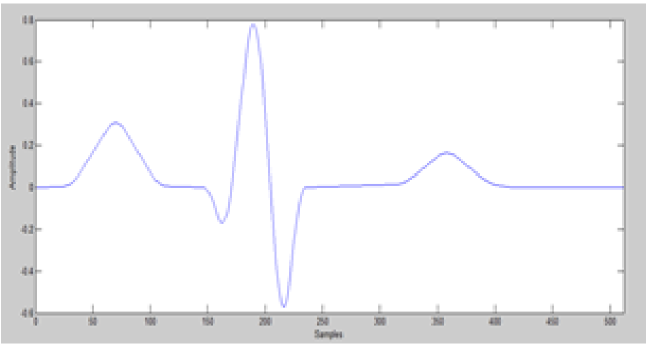


Figure 4. ECG signal with sampling rate 512 samples/sec and heart rate of 60 beats/min

Another issue in time frequency analysis of ECG signals is the scale of each signal and the distortion of the baseline (see Figure 6). The motion of electrodes due to patient movement changes the electrical conduction between the electrodes and the skin with time, which in turn affects the baseline of each signal. Also the scales of signals differ from patient to patient. So in order to unify the scales of all signals and to correct the baselines, all signals are divided by their average amplitudes, and then by their maximum amplitudes. This ensures that all signals have the same amplitude range, from 0 to 1 V.

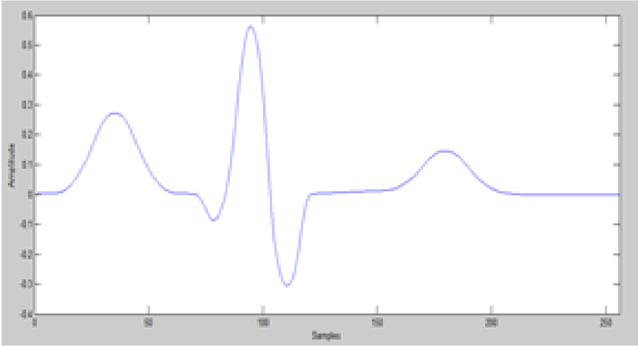


Figure 5. ECG signal with sampling rate of 512 samples/sec and heart rate of 120 beats/min

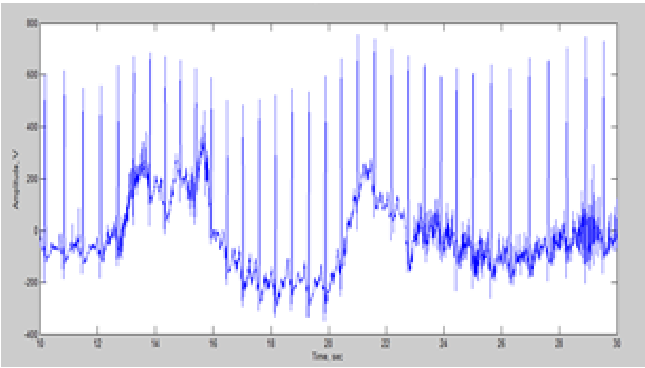


Figure 6. ECG recording with variation in the scale of each signal and baseline distortion

Feature extraction is a process that identifies certain parameters that characterize a certain class. Here the objective is to identify certain parameters that are common between all normal ECGs. Once the features are identified in the training part of the classification, then they can be used with unknown ECG signals to identify whether they are normal or not.

The sampling rate of the input signal was adjusted to 512 samples/sec by the use of MATLAB function interp (input_signal,int). In effect this step takes the number of samples of the "input_signal" and multiplies it by the integer "int". Since different ECG databases have different sampling rates, the "int" integer was made dependent on the sampling rate of the "input_signal". Moreover, the "int" has to be a real integer, so that the user may have a choice of sampling frequencies that are the power of two and that do not exceed 512 samples/sec. The reason that 512 samples/sec was chosen is because ECG databases used in Heartfelt have sampling rates that are the power of two (64, and 128) and the minimum sampling rate that is the power of two and exceeds 400 samples/sec (the minimum sampling rate required) is 512 samples/sec.

The following two normalization approaches were used to unify the shapes of all ECG signals and to correct any distortions in the baseline.

The first approach was to divide all signals by their average value, so that they are centered around the value one. Then a peak detection program was used to determine the peaks in each signal. Then all signals were divided by the average value of their peaks. This approach, however, did not work out well, simply because the signal was centered around the value one. A division by the average value of the peaks only shifted the signal up and down without adjusting its scale. As a result the baseline of the signal was distorted even more. The second approach worked better, as detailed below.

1. The signal was divided into windows of a thousand samples each. The reason behind choosing a thousand samples was that it is the minimum size at which there is always a single signal, at least one. At a heart rate of 30 beats/min, which is significantly low, and a sampling rate of 512 samples/sec, the size of each window can be calculated (check Figure 7):

$$window\ size = the\ time\ interval\ of\ a\ single\ beat \times sampling\ rate = \frac{60}{heart\ rate} \times sampling\ rate = \frac{60}{30} \times 512$$
$$= 1024 \approx 1000\ samples$$

2. The baseline was corrected by shifting each window by double the absolute value of its minimum amplitude (to ensure that the signal amplitude is always positive),

3. Divide each window by its average amplitude (so that the signal is centered around an amplitude of 1),

4. Subtracted one "1" from each window (so that the signal is centered around zero).

5. Finally, the amplitude was normalized by dividing each window by its maximum amplitude, which step rescaled each window to keep the maximum value to be always one. An example of two ECG signals at different heart rates divided into the same number of window is shown in Figure 8.

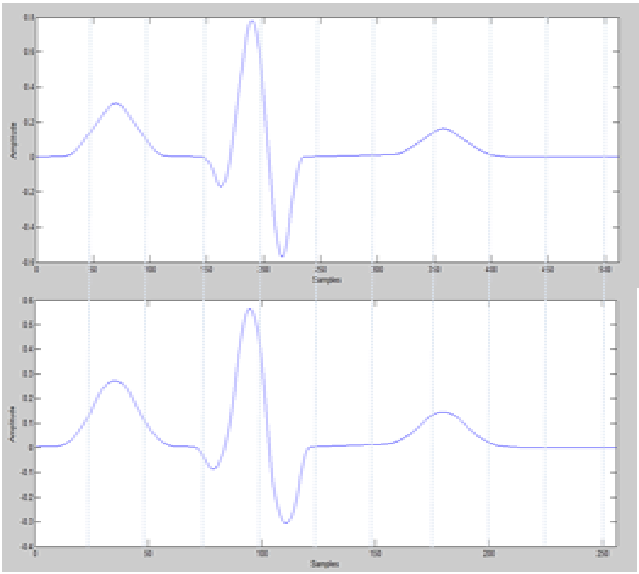
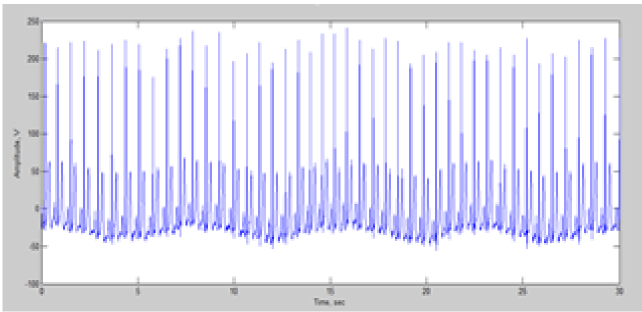
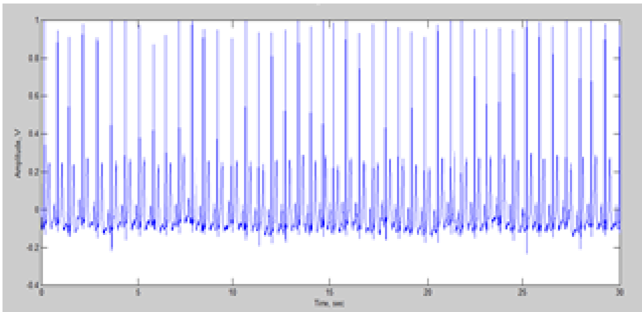


Figure 7. Two ECG signals at different heart rates divided into the same number of window



(a)



(b)

Figure 8. ECG Signal Baseline Adjustment (a) Before Normalization (b) After Normalization

After normalizing the signal, the next step is to obtain the time frequency transform of each signal. But before that can be done, each heart beat needs to be isolated from the overall signal. To this end, a threshold value needs to be defined. In order to do this, different normal and abnormal ECG signals from the MIT-BIH database were examined. It was found that the minimum value that is greater than almost all p-waves and smaller than almost all QRS-waves was 0.7 volts. The detection of peaks was achieved by:

- Running a for loop that searches for the first sample that is above a pre-defined threshold.
- Running a second *for* loop that starts recording all values that are above the threshold.
- Determining the maximum value of the recorded samples and storing it as a peak.
- Breaking out and returning to the first loop where the next first sample above threshold was determined, and the process was repeated again.

After the peaks had been detected, the "spectrogram" function was used to obtain the time frequency transform of each signal. The following parameters for the MATLAB command had to be defined:

1. The number of samples in each ECG period (between two peaks) keeps changing because the heart rate is not fixed. So in order for each signal to be divided into a fixed number of windows, the window size has to be dependent on the heart rate. This was done by dividing each period by the number of windows, which in this case was chosen to be 10. Then the function *floor* was used to round down the result. Figure 10 shows two ECG signals at different heart rates divided into the same number of windows.

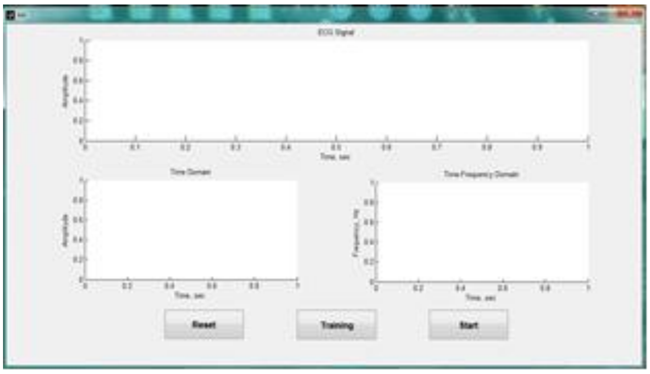


Figure 10. The GUI User Interface

2. The other parameter to be defined is the length of Fast Fourier Transform "nfft". The nfft number had to be greater than the maximum window size and in the power of two. So if the heart rate for the patient was 30 beats/min and a sampling rate of 512 samples/sec, the maximum window size becomes 1000/10 = 100 samples and the FFT number = 128.

$$at\ HR = 60/min\,, W = floor\left(\frac{512}{10}\right) = 51\ samples$$

$$at\ HR = 120/min\,, W = floor\left(\frac{256}{10}\right) = 25\ samples$$

Once the signal was transformed into the time frequency domain, an array is obtained wherein each column represents the window value in time, each row represents the frequency from zero hertz to 256 hertz and each value in the column represents the amplitude of the chosen frequency.

The next step was to extract some features from this array to represent the normal ECGs. The following features were identified:

1. The maximum amplitude of 64 frequency samples from zero hertz to 256 hertz was represented by a one column array with each row corresponding to one of the frequency samples (see Figure 10a)
2. The window locations of the maximum amplitudes of each of the frequency samples were represented by a column with each row having a value ranging from 1 to 10 corresponding to the window location of each of the frequency samples (see Figure 10c).
3. The minimum amplitude of 64 frequency samples was represented by a one column array with each row corresponding to one of the frequency samples (see Figure 10b).
4. The window locations of the minimum amplitudes of each of the frequency samples were represented by a column with each row having a value ranging from 1 to 10 corresponding to the window location of each of the frequency samples (see Figure 10d).
5. The energy of each window (see Figure 10e).

In the training part of Heartfelt, the program was trained with normal ECG data from MIT-BIH data base. Four-min recordings of normal ECGs from 10 patients were used in the training for this purpose.

1. First the recordings were checked to remove sparks, sudden spikes and silent periods.
2. Then the signal was normalized in time and amplitude

3. Then a peak detector identified R-waves (peaks) from the normalized signal, and ECG heart beats were isolated from each other.

4. Time frequency analysis was applied to each period.

5. Five features were extracted from the resulting time frequency distribution associated with each ECG period.

6. Features associated with each heart beat (one period) were stored in five arrays.

New ECG recordings were used to test the accuracy of the features obtained in the previous section. The testing part was done in the following steps:

1. The ECG associated with each heart beat in the unknown ECG is obtained using the peak detector.

2. TF distribution for each period is obtained.

3. Features for the unknown ECG are extracted and compared with the values obtained in the training part.

4. If the features fell between the range of the mean plus and minus three standard deviations of each period in the unknown ECG after the transformation part, then Heartfelt yielded an output with the number of normal ECG periods that occurred in the range, and the number of abnormal ECG periods that occurred out of the range.

In the design of the GUI, three *pushbuttons* were used on the GUI panel, each pushbutton had a specific operation. These pushbuttons trigger a *callback* when they are clicked with the mouse. Additionally three *Axis* components were provided. These components allow the displaying of graphics, such as graphs and images.

The first pushbutton is called "Start" button. What it does is it brings a message box and asks the user to enter the ECG signal. After the user enters the ECG signal, another message box comes up and asks the user to choose the proper sample rate for the ECG signal. Then if the user chooses the sample rate, this causes MATLAB code of the pushbutton (Start) to be executed. The results will show the accuracy of the signal. Also it will show the normality and the abnormality of the ECG signal.

The three axis components located on the GUI panel display three graphs. First graph will show the "Time Domain" of the ECG signal with respect to time and amplitude. The second graph will show the "Time-Frequency Domain" of the ECG signal with respect to time, frequency and amplitude. The third graph will show the original "ECG signal" in time domain. The same graph will also show the normal and abnormal parts of the ECG signal with different colors.

The second pushbutton on the GUI is called "Training". This is for use as a callback function to feed the software with more ECG signals just to train the software. This function is meant for making Heartfelt more efficient.

The third pushbutton is called "Reset" to reset the entire ECG database. Basically when reset is pressed, a message box will come up showing a wringing message to the user that he is about to reset the entire database, and if he likes to proceed, then this pushbutton will ask the user to enter a password that is given only to permitted users. Figure 10 shows a print screen of Heartfelt showing the graphical user interface (GUI). It will be observed that the GUI meets the specification of being simple and easy to use by patients.

The program for training the classifier has the following steps:

1)-Normalization: The normalization ensures that signals with different rates, baseline and amplitudes for the same heart conditions are assigned to the same class. The normalization has to address three parts (Amplitude, Baseline and Time normalization). The testing of this part was done by feeding it with ECG signals of different amplitudes, varying heart rates, and distorted baselines.

For amplitude normalization, the design was made so that all input signals have an amplitude of 1.0 V. However, after testing, it was found that amplitudes of all input signals range from 0.7 to 1.0 V. The reason behind this is that in some cases, there is more than one signal in the minimum window size, which leads to them being divided by the maximum amplitude of the one with the maximum peak value. It is also because the threshold being used to detect the peaks is 0.7 V. The error caused by this is trivial however. In practice the number of signals with peaks lower than 0.7 V is very few, and these are removed by proper training of the classifier so it can realize that this is not an abnormality.

As for the time normalization and baseline correction, the design was made so that each signal is divided into 10 windows no matter what the heart rate is, and has a smooth baseline that is centered around zero (Figure 7). After testing 8 recordings of normal ECG signals and 18 recordings of abnormal ECG signals, it was found that each and every signal has been divided into 10 windows successfully and without a single error. As for the baseline, it was found that the smoothness of the baseline was greatly improved. But still, there were some areas where the baseline was distorted beyond fixation. Figure 8 shows an ECG signal with distorted baseline before and after the correction.

2)-Peak Detection: This part of Heartfelt stores the location for all peaks that are above a predefined threshold. After testing for peaks, it was found that there were cases where there were some fluctuations in the signal, and more than one peak was detected for the same signal (see Figure 11). Such poorly recorded signals caused a constant crash for the specgram command in MATLAB.

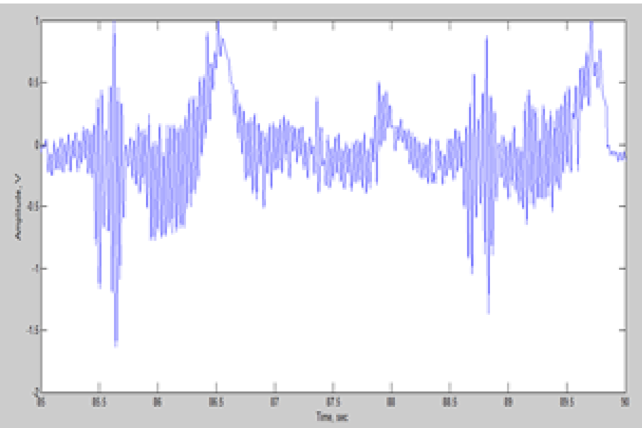


Figure 11. A Poorly Recorded ECG Waveform

This problem was fixed by adding a proofing condition that consistently checks if the duration between any two consecutive peaks detected is too short, and skips to the next peak where the condition is checked again, and so on until it finds the peak of the next heart beat.

<u>3)-Spectrogram:</u> This part of Heartfelt obtains the time frequency representation of the input ECG signal. It was designed so that all input signals are divided into 10 windows. Then a 64 point Fast Fourier Transform is applied to each window such that each window has a sampling rate of 512 samples per second.

<u>4)-Training/Testing the Classifier:</u> There are two parts in training the classifier. One is for training for the first time (only records the extracted features of the input ECG signals), and the other is for further training (first loads any saved features, then adds to it the features of the input ECG signals, and then saves the file again). Testing of the first part was carried out by applying ECG recordings, and then checking the saved features in an external ".mat" extension database. Testing of the second part was carried out by applying ECG recordings and checking that their features have been added to the already saved database.

<u>5-Feature Extraction:</u> In the feature extraction part, Heartfelt identifies features from the TF transform of the input ECG signal. Testing of this part is done by inserting ECG recordings from the MIT-BIH database and checking that the extracted features are in the expected shapes, which are (Figure 9):

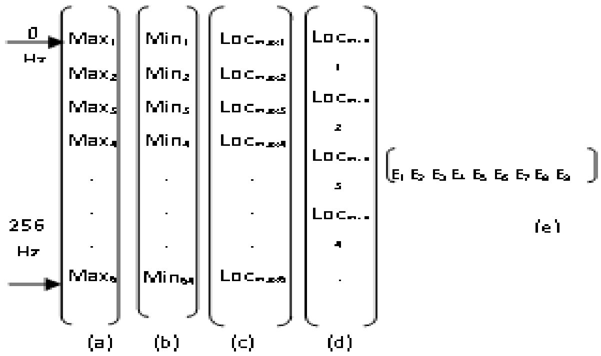


Figure 9. Novel Extracted Features for Normal ECG

- Four 65x1 arrays for the maximum and minimum amplitudes of 64 frequency samples from zero hertz to 256 hertz and their locations.

- One 1x10 array for the energy of each window.

In the classifier part of Heartfelt, the saved features from the training part were used to check that any new (unknown) ECG signal is classified as normal or abnormal. Recordings from the Normal Sinus Rhythm database were used in testing the accuracy of identifying normal ECGs. Then abnormal ECGs taken from the Arterial Fibrillation and Superventricle

Arrhythmias databases were applied to the classifier to verify the accuracy of identifying abnormal ECGs.

Testing for the accuracy of identifying normal ECGs was carried out by inserting several ECG recordings that are known to be normal, then checking to see if the percentage of the number of signals that were classified as normal was 90% minimum. Testing for the accuracy of abnormal ECGs was done in the same way, but instead of normal ECGs, abnormal ECGs were fed, and the percentage check was done on signals that were classified as abnormal.

A drop in the accuracy of detecting abnormalities was noticed throughout the classification testing. Abnormalities were included in the range, and as such were classified as normal ECGs because the range taken to define normal ECGs was constructed by taking the maximum and minimum values of the features. This problem was solved by creating a system that:

- Computes the maximum and minimum values of the features.
- Then computes the mean of the features plus and minus three standard deviations.
- Then checks which of the maximum or mean plus three standard deviations has a smaller value, and uses it as the upper limit of the range.
- And checks which of the minimum or mean minus three standard deviations has a greater value, and uses it as the lower limit of the range.

IV. Results

The complete Heartfelt system underwent 5 stages of testing:

1- First, 8 recordings (10 minutes each) of ECG signals from the MIT-BIH Normal Sinus Rhythm ECG database, which were not used in the training of Heartfelt, were entered as inputs. The following results were found (see Table 1):

2- Then, 8 recordings (10 minutes each) of ECG signals from the MIT-BIH Arterial Fibrillation ECG database were entered as inputs. The following results were found (see Table 2).

3- After that, 10 recordings (10 minutes each) of ECG signals from the MIT-BIH Superventricle Arrhythmia ECG database were entered as inputs. The following results were found (see Table 3)

4- The fourth test was on how much time it takes to execute Heartfelt from the beginning all the way to the end. So a function in MATLAB was used to trace Heartfelt and compute the execution time of processing 10 random ECG signals and for various recording times. Table 4 shows the results obtained.

5- The final test was for the size of Heartfelt. The size of Heartfelt, with a database that contains 3395 reference signals, was found to be 3.71 MB.

# Normal	# Abnormal	Heart Rate (beats/min)	Class	Accuracy (%)
761	18	78	Normal	97.7
799	1	80	Normal	99.9
842	79	92.1	Normal	91.4
687	1	68.8	Normal	99.9
822	44	86.6	Normal	94.9
738	4	74.2	Normal	99.5
808	30	83.8	Normal	96.4
878	56	93.4	Normal	94
Overall Accuracy of Detecting Normal ECG Signals				96.7

Table 1. Results from testing Heartfelt with normal ECG signals.

# Normal	# Abnormal	Heart Rate (beats/min)	Class	Accuracy (%)
152	980	113.2	Abnormal	86.6
0	377	37.7	Abnormal	100
0	636	63.6	Abnormal	100
0	760	76	Abnormal	100
0	660	66	Abnormal	100
0	570	57	Abnormal	100
0	553	55.3	Abnormal	100
7	985	99.2	Abnormal	99.3
Overall Accuracy of Detecting AF ECG Signals				98.2

Table 2. Results from testing Heartfelt with AF ECG signals.

# Normal	# Abnormal	Heart Rate (beats/min)	Classification	Accuracy (%)
1	328	32.8	Abnormal	99.7
0	317	31.7	Abnormal	100
0	460	46	Abnormal	100
122	341	34.1	Abnormal	75
0	423	42.3	Abnormal	100
0	334	33.4	Abnormal	100
54	488	48.8	Abnormal	90.1
0	381	38.1	Abnormal	100
0	326	32.6	Abnormal	100
0	487	48.7	Abnormal	100
Overall Accuracy of Detecting SV ECG Signals				96.5

Table 3. Results from testing Heartfelt SV ECG signals.

Execution Time (sec)			Ave Exec Time for a 1 min Recording (sec/min)
1 min Rec	5 min Rec	10 min Rec	
0.86	4.43	11.11	0.95
0.63	3.96	10.51	0.82
0.70	4.21	10.85	0.88
0.87	5.93	13.04	1.12
0.40	2.57	8.05	0.57
0.37	2.23	6.01	0.47
0.50	5.99	22.22	1.31
1.01	6.24	14.67	1.24
0.62	3.71	9.14	0.76
0.58	5.02	16.23	1.07
Overall Average Execution Time of 1 min Recording			0.92

Table 4. Results from testing execution time of Heartfelt

V. Discussion

The first step in the validation process was to install Heartfelt on a windows operating system that does not have MATLAB loaded on it. An executable windows stand-alone version of Heartfelt was generated by using MATLAB's built-in compiler, and successful installation of Heartfelt was confirmed. It must be stressed that running of Heartfelt does not require additionally the installation of the software MATLAB Compiler Runtime MCR; this latter software is already provided with Heartfelt.

The next set of steps was to obtain the time frequency distribution of a normal ECG signal, to extract the features representing it, and to confirm that the features representing it has countable values. Figure 12 shows a normal ECG signal. Fig. 12 shows the time frequency distribution of the normal signal. It also shows that all input signals have been normalized to meet the specification for having an amplitude range from 0.0 V to 1.0 V.

The time-frequency distribution is represented by arrays S, F, and T where S represents the amplitudes of a range of frequencies for each window, F the frequency samples, and T the window intervals.

Five features (Figure 9) were extracted from the S array, all five were confirmed having countable values. The features are

a) the maximum amplitude at each frequency sample,

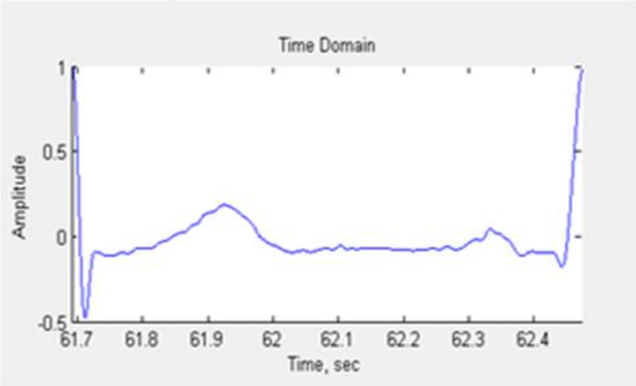
b) window location of the maximum amplitude,

c) the minimum amplitude at each frequency sample,

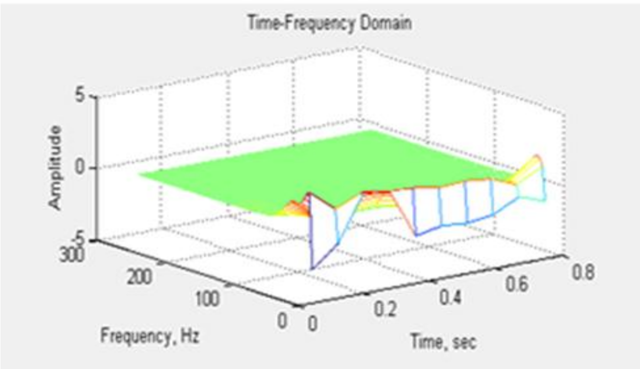
d) the window location of the minimum amplitude, and

e) the window energy.

The execution time was found to be 0.92 seconds for each minute of ECG recording. It was observed that the achieved accuracy was 96.7% for detecting normal ECG signals, 98.2% for detecting arrhythmic ECG signals suffering from arterial fibrillation, and 96.5% for detecting arrhythmic ECG signals suffering from super-ventricle arrhythmias.



(a) Normal ECG



(b) Spectogram for the Normal ECG

Figure 12. Spectogram of a Normal ECG. (a) Normal ECG (b) Spectogram of ECG

VI. Conclusions

Heartfelt is a computer program, available on a CD or a USB. All a patient needs to do is to obtain an ECG interface, and record his ECG. Then he loads Heartfelt on a computer, feeds the ECG when Heartfelt asks for it. Then Heartfelt rapidly analyzes his heart condition.

Many elderly people cannot go to the hospital each time they feel heart pain, and wait for the doctor to measure their ECGs to analyze if his pain is normal or abnormal. This is a long procedure that may be difficult for most elderly people. Using Heartfelt, the patient does not need to take an appointment with the heart specialty doctor or even to go to the hospital. All he has to do is to enter his ECG to Heartfelt, and to run it to see if his heart pain is normal or abnormal. There is no medical intervention. He will know the true condition of his heart in less than 10 seconds. Heartfelt may thus help save lives.

For further development, it may be desirable to augment Heartfelt with a low-cost ECG interface such that a patient does not need to buy anything other than Heartfelt to determine his heart condition. Likewise Heartfelt may be bolstered with another program to detect the type of arrhythmia that the patient has.

References

1. http://www.health.com/health/library/mdp/0,,sts14215,00.html, SA Node and AV Node (downloaded 26 Feb. 2012).
2. http://www.medicalnewstoday.com/articles/8887.php, What Is Arrhythmia? What Causes Arrhythmia? (downloaded 26 Feb. 2012).
3. Mahesh S. Khadtare and J. S. Sahambi, ECG Arrhythmia Analysis by Multicategory Support Vector Machine, AACC 2004: 100-107.
4. Charulatha Ramanathan, Raja N Ghanem, Ping Jia, Kyungmoo Ryu and Yoram Rudy, Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia, Medical and Biological Engineering and Computing, Volume 42, Number 3, 288-293(2004).
5. Khadra, L.; Al-Fahoum, A.S.; Binajjaj, S.; ,A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques , Biomedical Engineering, IEEE Transactions on Issue Date: Nov. 2005 Volume: 52 Issue:11 On page(s): 1840 – 1845.
6. Heikki V. Huikuri, Timo H. Mäkikallio, M. J. Pekka Raatikainen, Juha Perkiömäki, Agustin Castellanos, ;Robert J. Myerburg, Prediction of Sudden Cardiac Death Appraisal of the Studies and Methods Assessing the Risk of Sudden Arrhythmic Death, Circulation. 2003; 108: 110-115.
7. Moody, G.B.; Mark, R.G.; The impact of the MIT-BIH Arrhythmia Database , : Engineering in Medicine and Biology Magazine, IEEE Issue Date: May-June 2001 Volume: 20 Issue:3 On page(s): 45 – 50.
8. K P Pfeiffer, T Kenner, J SchaeferApplication of statistical methods for the analysis or interval related cardiac performance variations during cardiac arrhythmia in man Cardiovasc Res (1984) 18 (2): 80-98
9. Coast, D.A.; Stern, R.M.; Cano, G.G.; Briller, S.A.; An approach to cardiac arrhythmia analysis using hidden Markov models, Biomedical Engineering, IEEE Transactions on Issue Date: Sept. 1990 Volume: 37 Issue:9 On page(s): 826 - 836 .
10. A. F. Safe, R. T. Maxwell, Transtelephonic electrocardiographic monitoring for detection and treatment of cardiac arrhythmia. Postgrad Med J 1990; 66:110-112 doi:10.1136/pgmj.66.772.110.
11. Szi-Wen Chen; Clarkson, P.M.; Qi Fan; A robust sequential detection algorithm for cardiac arrhythmia classification, Biomedical Engineering, IEEE Transactions on Issue Date: Nov. 1996 Volume: 43 Issue:11 On page(s): 1120 – 1124.
12. Guvenir, H.A.; Acar, B.; Demiroz, G.; Cekin, A.; A supervised machine learning algorithm for arrhythmia analysis, Computers in Cardiology 1997 Issue Date: 7-10 Sep 1997 On page(s): 433 – 436.
13. Kelly H. McClure et al, Cardiac arrhythmia detection system for an implantable stimulation device, Patent number: 5685315, Nov 11, 1997.
14. Johnson Francis, Mari A. Watanabe, Georg Schmidt, Heart Rate Turbulence: A New Predictor for Risk of Sudden Cardiac Death, Annals of Noninvasive Electrocardiology Volume 10, Issue 1, pages 102–109, January 2005.
15. Robert D. Throne, Janice M. Jenkins, Lorenzo A. DiCarlo, US Patent number: 5000189, 1991.
16. Chuang-Chien Chiu, Tong-Hong Lin and Ben-Yi Liau, Using correlation coefficient in ECG waveform for arrhythmia detection, Biomedical Engineering Applications Basis and Communications (2005), Volume: 17, Issue: 03, Pages: 147.
17. V. Vijaya, K. Kishan Rao, V. Rama, Arrhythmia Detection through ECG Feature Extraction using Wavelet Analysis, European Journal of Scientific Research, ISSN 1450-216X Vol.66 No.3 (2011), pp. 441-448.
18. Babak Mohammadzadeh Asl , Seyed Kamaledin Setarehdan , Maryam Mohebbi, Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. Artificial Intelligence in Medicine 44(1): 51-64 (2008).
19. Scott Kinlay, James W. Leitch, Amanda Neil, Barry L. Chapman, David B. Hardy, Peter J. Fletcher Cardiac Event Recorders Yield More Diagnoses and Are More Cost-effective than 48-Hour Holter Monitoring in Patients with Palpitations: A Controlled Clinical Trial, Ann Intern Med January 1, 1996 vol. 124 no. 1 Part 1 16-20.
20. James W. Leitch,Amanda Neil,Barry L. Chapman, David B. Hardy, Peter J. Fletcher, Ann Intern Med January 1, 1996 vol. 124 no. 1 Part 1 16-20.
21. Steven A. Rothman, James C. Laughlin , Jonathan Seltzer , Jasjit S. Walia , Rakesh I. Baman , Samer Y. Siouffi , Robert M. Sangrigoli , Peter R. Kowey, The Diagnosis of Cardiac Arrhythmias: A Prospective Multi-Center Randomized Study Comparing Mobile Cardiac Outpatient Telemetry Versus Standard Loop Event Monitoring, Journal of Cardiovascular Electrophysiology, Volume 18, Issue 3, pages 241–247, March 2007.
22. Srinivasan, N.; Wong, M.T.; Krishnan, S.M.; A new phase space analysis algorithm for cardiac arrhythmia detection, Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE Issue Date: 17-21 Sept. 2003 On page(s): 82 - 85 Vol.1.
23. Prasad, G.K.; Sahambi, J.S.; Classification of ECG arrhythmias using multi-resolution analysis and neural networks, TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region, Issue Date: 15-17 Oct. 2003 On page(s): 227 - 231 Vol.1 .
24. P. Kirk, MB,; M. Roughton,; J.B. Porter, J.M. Walker, M.A. Tanner, J. Patel, D. Wu, J. Taylor, M.A. Westwood, L.J. Anderson, D.J. Pennell, Cardiac T2* Magnetic Resonance for Prediction of Cardiac Complications in Thalassemia Major, *Circulation* 2009, 120:1961-1968: originally published online October 2, 2009.
25. L. Khadra, A. Fraiwan, W. Shahab, Neural-wavelet analysis of cardiac arrhythmias, hua83081, 3rd. WSES Int. Conf. on Neural Networks and Applications. www.wseas.us/e-library/conferences /switzerland 2002/
26. Wikipedia, Electrocardiograph. Last visited, March 2012. http://en.wikipedia.org/wiki/Electrocardiography
27. MIT-BIH Arrhythmia database. Last visited, March 2012 http://www.physionet.org/physiobank/database/mitdb/