R-Peak detection in ECG signal compression for Heartbeat rate patients at 1KHz using High Order Statistic Algorithm

Ali Tariq Bhatti[1]
[1]Department of Electrical & Computer engineering
[1]North Carolina A&T State University
Email:[1]alitariq.researcher.engineer@gmail.com, atbhatti@aggies.ncat.edu, ali_tariq302@hotmail.com

Dr. Jung H. Kim[2]
[2]Department of Electrical & Computer engineering
[2]North Carolina A&T State University
Email: kim@ncat.edu

Abstract*—One of the major problems in the analysis of Electro-Cardiogram (ECG) signal compression is the accurate detection of R-peak. It is due to the difficulties varied by the time varying morphology of ECG, the physiological variations due to the patient and the noise contamination. However, it includes power line interface, muscle contraction noise, poor electrode contact, patient movement, and baseline wandering due to respiration. R-peak applications require accurate heart beat monitoring systems including intensive care units, operating rooms, implantable pacemakers and defibrillators. Moreover, it detects a QRS complex when ECG amplitude exceeds a threshold level. If the threshold is too high, true beats can be missed. If the threshold is too low, false detection can result during EMG artifact and external interface. So, during these artifacts, the magnitude of the noise can become larger than the signal, QRS detection based on amplitude threshold alone is not satisfactory. The problem was due to the detection threshold used, not the matched filter implementation. The R peak detection threshold was two-thirds of the average value of the output of the matched filter for the specific feature of interest. To overcome this problem, R peak detection has been proposed because it is adaptive to the nonlinear and time varying features of ECG signal. It can be trained to recognize the normal waveform and filter out the unnecessary artifacts and noises. Usually R-peak detection in QRS complex can be improved by considering multiple features, including RR interval, pulse duration and amplitude. In this research paper, is to take the difference of maximum original signal and minimum original signal to obtain the filtered R-peak ECG signal after 1st and 2nd pass to observe noise for 512 sample points at sampling frequency of 1 KHz using High Order Statistics Algorithm.Once the R-peak is detected, compute the Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR). From analytical perspective of this research, it appears to be exceedingly robust, correctly *detects R-peaks even aberrant QRS complexes in noise-corrupted ECG signal compression.

Keywords—component; ECG, R-Peak, QRS, FFT, High Order Statistics Algorithm, Threshold, Skewness, Kurtosis, Power Spectrum Density, MSE, PSNR.

1. Introduction

(a) Electro Cardiogram:

The 'Electro-Cardiogram' (ECG) is an important and commonly used diagnostic aid in cardiovascular disease. ECG is a time varying signal that measures an electrical recording of the heart, depicts the cardiac cycle. The heart rate variability (HRV) for instance provides a useful risk predictor after acute myocardial infarction and an early symptom of neuropathies [1]. The heartbeat recognition including the premature ventricular contractions (PVC) events is, therefore, a basic task in numerous ECG applications [2], [3].To process ECG: irregular distance between peaks, irregular peak form, presence of low-frequency component in ECG due to patient breathing, cardiac disease, etc. are to be determined. The frequency representation of ECG signal is required to accomplish FFT (Fast Fourier Transform) that justifies the use in quantitative electro cardiology. Efficient R wave detection is a crucial preprocessing step of most of the ECG signal analysis. Its shape is similar to the QRS complex by performing the analysis of an ECG signal it should give a considerably higher coefficient value for the R peak against very close to zero coefficients to non-QRS/R peak complex.

(b) R-Peak Detection:

The R-Peak is the most prominent part of an ECG signal analysis, which corresponds to the contraction of the ventricles during a heartbeat. A statistical based approach was used for detection of R-peaks and other time plane features of the ECG [6]. The detected R-peaks are not always accurate and can have false or missed peaks. Algorithms to increase detection sensitivity by processing the RR intervals were proposed [7, 8]. It is sensitive to the noise and tends to produce a segment that is less than a heart cycle due to multiple peaks within a heartbeat. Before the occurrence of R-wave, the slope of the signal is positive and after the R-wave, the slope is negative.

Due to these reasons, cardiac abnormalities can be detected by this method. When an R-wave is missing, two cases may arise. In first case, the ventricles do not generate the R wave. Therefore, in second case, the peak of the R-wave is not sufficient enough because of improper ventricular conduction, which is a cause of lack of oxygenated blood supply to the ventricular muscles. However, it does not contract enough forcefully to generate a normal R-wave. Therefore, ECG waveform for R peak clarifies the robustness and accuracy of image-based period detection of various cardiovascular diseases as shown in the Figure1.

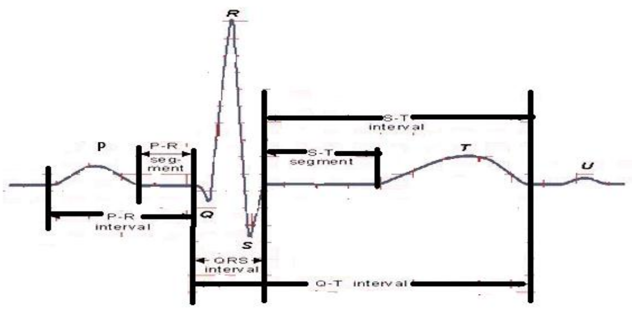

Figure 1 ECG signal and its components

From the figure 1, for each detected R peak, the first local maximum point on its left is detected as the beginning of the R wave and the first maximum point on its right side is detected as the end of the R wave. Secondly, the first local minimum from the left of the positive R wave is detected as the Q wave. If the minimum cannot be detected, the Q wave is judged to be missing. Thirdly, the first local minimum from the right of the positive R wave is detected as the S wave. Otherwise, the S wave is judged to be missing. Same time interval as for Q wave detection is set for S wave detection.

(c) Amplitude Based Technique:

Amplitude based technique is one of the techniques to be used for R-peak detection. It performs a sample of ECG points that fall beyond an amplitude threshold which can be determined to be a QRS complex candidate. The highest amplitude of the detected QRS is ascertained to be R-peak. The original ECG signal x_n from the patient body is given as $x_n = x_1,x_2,.....,x_N$. Where n=1, 2 ,…, N and N is the length of the signal.

$$(x_r,x_{r+1},x_{r+2},\ldots,x_{r+k}),\ldots,(x_i,x_{i+1},x_{i+2},\ldots,x_{N+c})$$

, which is greater than the Amplitude Threshold.

So, 1<r<1<N, x_{N-c} is the last value greater than the Amplitude Threshold.

As, R-peak=Max (QRS Complex),

RR Interval = n_r/f

Where n_r is the total number of sample between the two corresponding R-peak and 'f' is the sampling frequency of the ECG.

(d) MATLAB Implementation

Based on High Order Statistic Algorithm:

ECG bio-signal processing is based on the detection of the R peak. The detection of the R-peaks and QRS complexes using an ECG signal provide information about the heart rate, the conduction velocity, the condition of tissues within the heart, and abnormalities. It also supplies information for the diagnoses of cardiac diseases. The presence of noise and time-varying morphology makes the detection difficult. For the purpose of diagnosis, various features are planned to be preprocessed from ECG data including QRS intervals, QRS amplitudes, PR intervals, ST intervals, fetal heart rate, etc.

The detection of the R peaks is the first step of feature extraction. For example, if the patient data are broken into segments of 600 points and only one segment is analyzed at a time. The R peaks have the largest amplitudes among all the waves making them the easiest to detect and good reference points for future detections. The signal was processed using the db (decibel) 6 wavelet up to 8 levels. However for the detection of R peaks, only details up to level 2^5 were kept and all the rest removed. This procedure removed lower frequencies considering QRS waves have comparatively higher frequency than other waves. The attained data is then squared to stress the signal. A threshold equals to 30% of the maximum value is sub-sequentially applied to set a practical lower limit to help to remove the unrelated noisy peaks. At this point, the data set is ready for peak detection through a High Order Statistics algorithm that produces very accurate results.

To perform the R wave detection, the analysis of an ECG signal should give a considerably higher coefficient value for the R peak against very close to zero coefficients to R peak wave. To implement the R-peak, following steps are required for it

(1) It starts by calculating the average deviation relatively to the averaged value of all wavelet coefficients in each level.

(2) In each level, High Order Statistics algorithm sweeps the coefficients, selecting the ones with absolute value greater than the level threshold value.

(3) If no coefficient is found, then the threshold is adjusted to 90% of its value. To avoid the processing of repeated coefficients, the algorithm selects the maximum absolute coefficient value.

(4) The window is set because of only one QRS complex exists. For e.g., if after the identification of relevant coefficients is at level 2^4, a search is performed in level 2^3. Therefore, search is done by the combination of these two levels.

(5) If still no relevant value is found, algorithm adjusts the threshold and repeats the search until a relevant coefficient is found, or the threshold reaches the value of the average deviation.

(6) At this point, the number of relevant coefficients selected at level 2^3 is at most equal to the number of relevant coefficients at level 2^4. The algorithm continues the R peak detection in a similar procedure at levels 2^2 and 2^1. The only difference between these two levels and the search in level 2^3 is the search time window in milliseconds.

To conclude, if any R peak on the ECG signal has been undetected, the algorithm performs a confirmation by distinguishing the R-R interval with the immediate two before and after intervals. In case of the number of the R-R intervals being lower than required, it only uses the existing ones, without exceeding two before and two after. Once R-peak detection has been implemented, the remaining task is only to locate QRS complexes onsets and offsets.

QRS Onset detection. After selecting all local extremes before the R peak, the algorithm searches for the two closest zero crossings to the R peak. If only one zero crossing is found, then the two points are relevant and the onset of the QRS is located at the time instant of the point more apart from the R peak.

QRS Offset detection. It starts by selecting the local extremes after R peak. As the first value is due to the R peak, then it's performed a search the first value after a zero crossing. So, this point is known as the QRS complex offset.

2. Flow-chart of R-peak Detection

The flow-chart explains that we load our mat file for heartbeat rate patients and initialize the Sampling frequency rate at 1KHz and 512 sample points.

Detection of R peak wave in ECG based on the High order statistics, such as Skewness and Kurtosis, in order to formulate it with high accuracy. It could be feasible in a real-time context and applied in routine ambulatory or clinical heat rate screening due to its simplicity.

The flow-chart represents ECG signal X(k) takes place where amplitude normalization and DC extraction of the N-sampled are performed.

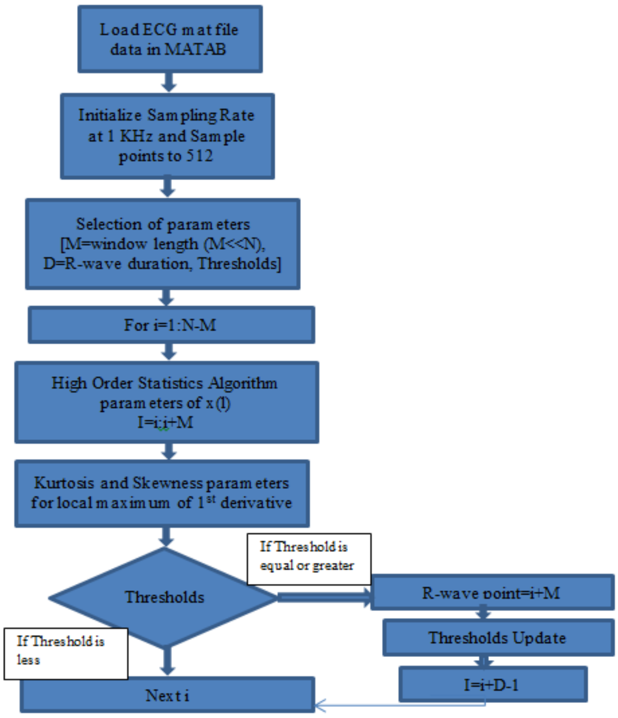

Figure 2 Flow-chart of R-Peak ECG signal compression at 1KHz

Then, the length (M<<N) of a sliding window along with the R-wave length 'D' are set in accordance to the sampling frequency 'f_s'. However, initial values of the threshold used for High Order Statistics parameters are selected. Next, X(k) is windowed with a sliding window of M samples. At each window $\hat{\gamma}_3$ and $\hat{\gamma}_4$ are calculated using equation 4 and equation 5 and its values are located at the end of the window. Then the local maximum derivative of $\hat{\gamma}_3$ and $\hat{\gamma}_4$ are calculated based on the location of R peak wave. Therefore, the values of the $\hat{\gamma}_3$ and $\hat{\gamma}_4$ at possible location compared with the two thresholds; in case they are smaller, so slide the window by one sample. If the value exceeds the thresholds, the two locations are compared with each other. The thresholds are updated using the mean value of the last five maximum values of $\hat{\gamma}_3$ and $\hat{\gamma}_4$. After the estimation of the first R-wave, the window skips 'D' samples to the right and then it proceeds with the next one until the end of the input data is reached.

If,

$$\{x(k)\},\ k = 0, \pm 1, \pm 2\ldots.. \quad (1)$$

It is a real stationary random process and its moments up to 'n' order exixts. Then,

$$m^x_n(\tau_1, \tau_2, \ldots, \tau_{n-1}) = E\{X(k), X(k+\tau_1)\ldots X(k+\tau_{n-1})\} \quad (2)$$

Due to stationary, it depends only on the time differences i.e.

$$\tau_1, \tau_2, \ldots \tau_{n-1},\ \tau_i = 0, \pm 1, \ldots \text{ for all } i; \quad (3)$$

Where E{.} denotes the Expected Value. For n=3, 4 and $\tau_1 = \tau_2 = \tau_3 = 0$. Then, the Skewness γ^x_3 and

and Kurtosis γ^x_4 can be given by equation 4 and 5.

$$\gamma^x_3 = E\{X^3(k)\} \quad (4)$$

and

$$\gamma^x_4 = E\{X^4(k)\} - 3[E\{X^2(k)]^2 \quad (5)$$

For a real ECG signal, R-peak can be calculated by Skewness γ^x_3 and Kurtosis γ^x_4 as:

$$\widehat{\gamma_3} = \frac{\sum_{i=1}^{N}(x(i)-\hat{m})^3}{(N-1)\hat{\sigma}^3} \quad (6)$$

$$\widehat{\gamma_4} = \frac{\sum_{i=1}^{N}(x(i)-\hat{m})^4}{(N-1)\hat{\sigma}^4} - 3 \quad (7)$$

Where $\hat{m}$ and $\hat{\sigma}$ are the estimated mean value and standard deviation of X(k).

3. MATLAB Implementation Procedure:

In this research, the sampling rate frequency of 1000 Hz at 512 points is being use from our MATLAB mat file for extracting the R-peak detection of ECG signal compression analysis. Use the Fast Fourier Transform for ECG signal to create the Original ECG signal compression by doing the Inverse Fast Fourier Transform to obtain Filterd R-peak ECG signal as in figure 3 and figure 4.

Then, use ECG R-peak Window for filtering the R peak detection. The basic purpose of this window is that after filtering the R-peak detection after two passes, it also samples the data points, checks the noise at which point and tends to produce heart beats on various R-peak points. Now, filtering the window for first pass of first R-peak by using the threshold filter as a default window size and also be calculating QR distance. Next, filtering the window for second pass of second R-peak by using the threshold filter, which has a optimized window size.

So, finally what this MATLAB code in this research paper is taking the difference of maximum original signal and minimum original signal to obtain the filtered compressed R-peak ECG signal. However, observe noise on different sample points of R-peak.

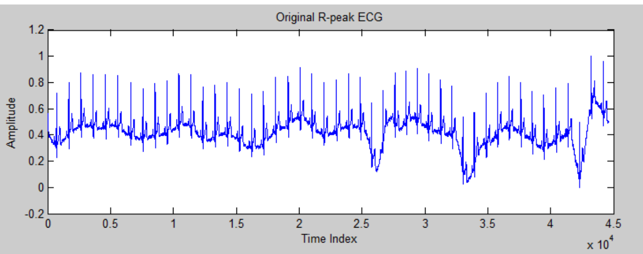

Figure 3 Original R-peak ECG signals compression

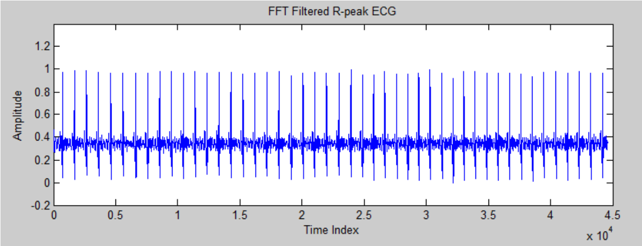

Figure 4 FFT Filtered R-peak ECG signals compression

For first pass, the difference of maximum filtered-1 R-peak and the minimum filtered-1 R-peak to obtain the 1st pass of R peak. Once the 1st pass is detected for Filter1 from figure 5, so peak1 is calculated in figure 6.

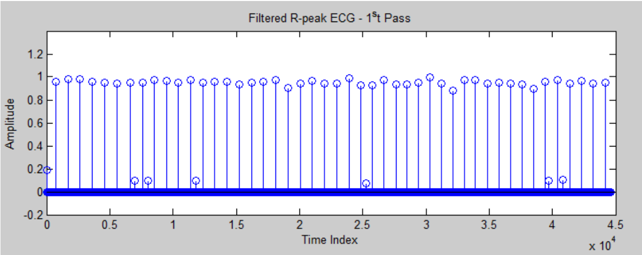

Figure 5 Filtered R-peak ECG for 1st Pass

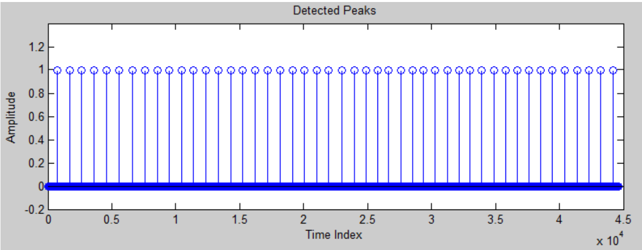

Figure 6 Detected Peaks for 1st Pass

For second pass, the difference of maximum filtered-2 R-peak and the minimum filtered-2 R-peak to obtain the 2nd pass of R-peak. Then the 2nd pass is detected for Filter2 from figure 7, so peak2 is calculated in figure 8.

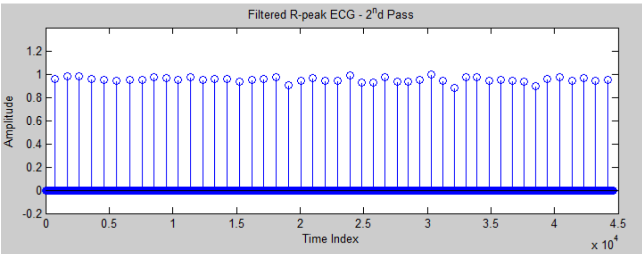

Figure 7 Filtered R-peak ECG for 2nd Pass

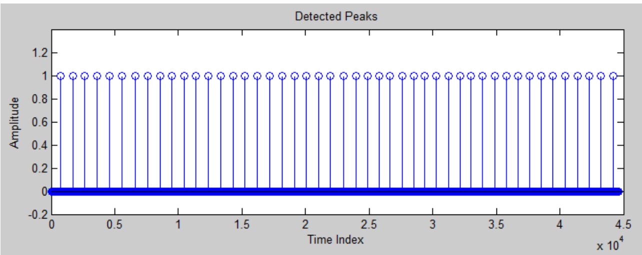

Figure 8 Detected peaks for 2nd Pass

The Final R-PEAK detection after the 2nd Pass in figure 9, which is the final result.

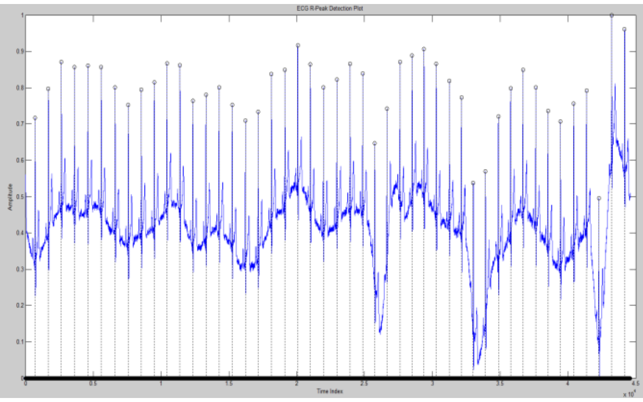

Figure 9 ECG R-peak Detection

4. Results

The sampling rate in this research paper implementation is 1 KHz. From figure 10 and figure 11, the Power Spectrum Density (PSD) of R peak ECG signal at 1000 Hz corresponds to the low frequency component of the heart variability. Due to the small number of data points, equal to 512, each R-peak ECG's was zero-padded to the same 1024 sample points' length before the PSD was determined. In figure 10, taking the power spectrum of Input Original R-Peak ECG waveform at 512 points and 1000 Hz.

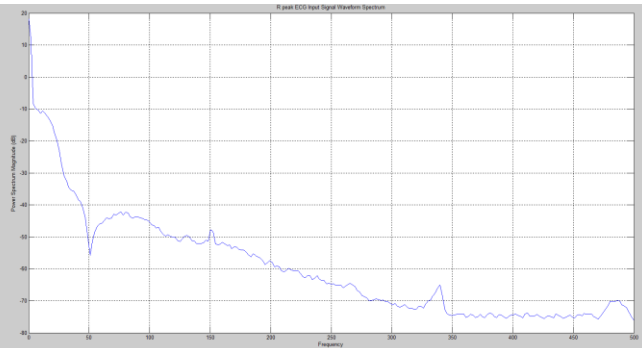

Figure 10 R-peak ECG Input Signal Waveform Spectrum

Therefore, in figure 11, taking the power spectrum of Output filtered R-Peak ECG Waveform.

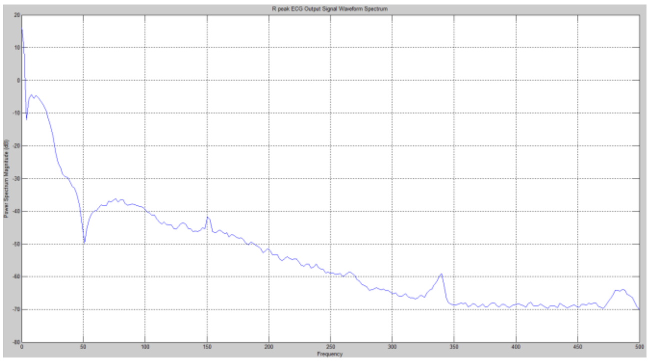

Figure 11 R-peak ECG Output Signal Waveform Spectrum

So from figure 10 and 11, R-peak is at 150 Hz, the other one is approximately be 340 Hz and vice versa.

If we change it to the frequency to 10,000 Hz, then observe noise in R peak, but not in the case of 1000 Hz. As the figure 10 signals are little lower than the figure 11, but having the same peaks. It is due to this reason that once filtering the Original R-peak ECG waveform will change the signals where is that noise flowing. From figure 4, FFT filtered ECG have noise at different R-peaks, especially if observing in 2 millisecond time index. Therefore, in figure 5 and figure 7 for Filter1 first pass and Filter2 second pass, most of the R-peaks are very down for filter1 first pass especially observe in 2.5 millisecond time index. But for Filter2 2nd pass having better results than Filter1. After detection of first pass and second pass of Filter1 and Filter2, results looks same. However, finally observe from figure 9 that the highest R-peak detected with the noise is at 4.4 millisecond time index and the lowest R-peak detected with the noise is at 4.3 millisecond time index.

Each of the preprocessed ECG signals are as shown in the figure 3. It means that only 60 seconds of the 300 seconds of each ECG signal is shown so that individual heart beats are easily discernable. Since, re-sampling introduces aliasing errors and the comparison of these ECG signals before and after re-sampling and low-pass filtering as shown from figure 3 through figure 8. There were only small amplitude changes in both ECG signals after re-sampling of filter1 and filter2 passes. Therefore ST segment noise did not increase significantly though the re-sampled ECG transition point of the S-wave. S-wave show more noise than in the original signal, but not enough to require a different re-sampling frequency.

It is evident that R-peak detection of ECG signal analysis is able to pick up the R peak from the noisy waveform, even when the noise level is high. This method exhibits stronger to noise than the derivative-based and other filter based approaches.

The R-peak detection threshold was two-thirds of the average value of the output of the matched filter for the specific feature of interest. Good R-wave peak detection is necessary in order to determine the heart rate variability. In base-line wander, amplitude changes or noise which can cause poor detection and both of these R-peak ECG signals from figure 3 and 4 had more noise and extraneous peaks than the other two filters passes. Only one heartbeat was not detected and incorrectly detected beats (false positives) were much lower than with R-wave peak detection as observed in figure 9.

The Cardiac Arrest ECG had a different detection problem. There were areas within the ECG signal where the ECG signal was lost, noisy or chaotic due to cardiac arrest. The preceding R-wave peaks before the arrest section were low in voltage level in relation to the rest of the signal, so the matched filter threshold High Order Statistics algorithm had a difficult time distinguishing the R wave peaks.

The use of the matched filter to detect the R-R intervals made this possible. Once the R-wave peaks were detected, the R-R intervals were then calculated by using Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) in MATLAB equations as

squaredErrorImage = (max(ecg) - min(ecg)) .^ 2; Eq(a)

squaredErrorImage = (max(ecg) - min(Original)) .^ 2; Eq(b)

squaredErrorImage = (max(ecg) - min(filtered2)) .^ 2; Eq(c)

MSE = sum(sum(squaredErrorImage)) / (512000); Eq(a1)

%512 are sample points and 1KHz frequency which are both then multiplied

PSNR = 10 * log10(256^2 / MSE); Eq(a2)

From Table 1 analysis:

By using equation (a), we get less MSE of 6.24 from eq(a1) and more PSNR of 40.21 from eq(a2) as compare to equation(b) and equation(c) because once they filtered, they produce more noise than the Original R-Peak ECG signal.

ECG R-Peak	MSE	PSNR	Sample points	Frequency (Hz)
MSE and PSNR b/w maximum and minimum ECG data for Original R-peak	6.24	40.21	512	1000
MSE and PSNR b/w maximum ECG data and FFT filtered Original R-peak ECG	24.64	34.25	512	1000
MSE and PSNR b/w maximum ECG data and Filtered Detected R-peaks ECG at or after 2nd Pass	20.52	35.04	512	1000

Table 1: Performance metrics for R-Peak ECG signals

5. Discussion of the Result and Performance Analysis:

(a) Result:

A lot of information on the normal and pathological physiology of heart can be obtained from ECG. However, the ECG signals being non-stationary in nature, it is very difficult to visually analyze them. Thus the need is there for computer based methods for ECG signal analysis [4]. R- peak detectors are very useful tools in analyzing ECG features thus form the basis of ECG feature extraction [5]. The detection of R-peak in the Original ECG waveform was little hard, but it was not hard due to the implementation of the matched filter. It was happen because of this reason that R peak detection method was the difficult part to extract the feature of interest in each ECG signal. As a result, observed that ST signals were similar, but their respective detection thresholds were not as affected by extraneous peaks as the Arrhythmia and Cardiac Arrest ECG signals. The other thing observed from the result was that the matched filter was sensitive of extraneous or multiple R peaks and had difficulty extracting the signal of interest, when its duration was short. When the matched filter signal had short, sharp R-peaks to detect, then the detection was not as good as with a larger and less sharp as a QRS complex. Therefore, R peak varies from heart beat to heart beat and is highly susceptible to noise, which makes slope detection difficult and requires multiple R-peak detection passes.

(b) Performance Analysis

To analyze the detection of R-peak, filtered ECG signal passes through the moving window integrator. The window width of the filter correlates to the heart rate, improving the detection quality. However, the window output signal goes to the local peak detector to check whether the samples differs from the previous value or not. The QRS-detector itself gets the filtered ECG signal and a binary output of the R-peak detector. Therefore, it is only detected, if a peak in the ECG signal is detected.

On the other hand, undetected R-peaks always result in the loss of information [9]. If no R-peak is detected for a certain time, then the High Order Statistics Algorithm is used, so then the peaks with lower amplitude can also be detected due to this reason. Moreover, double heartbeats can be detected too, if the amplitude of the T-wave is too high.

If the input ECG signal compression is not stable, then the R-R interval is irregualar and R-peak amplitude is too low. So, the “Low Signal Quality” flag is activated.

Determining the slopes on both sides of zero crossing between peak of the lowest slope and the highest slope is chosen as the R wave peak location. If R-to-R interval turns out to be longer than the previous interval, then ECG search is repeated with a lower threshold to detect a possibly missed heartbeat. The inverse of the interval between two consecutive R wave peaks gives the instantaneous heart rate that shows how heart rate varies.

In this research paper, it was tried to find out the R-peak in QRS complex with missing peak and false peak as well and therefore result come accurate after the 1st and the 2nd pass of R-peak detection for filter1 and filter2 by observing in MATLAB figures 3 through figure 8 and finally R-peak detection in figure 9. But the correct decision can be made by the physician and clinician. From this research paper, we analyzed that the problem appears to be exceedingly robust, correctly detects R-peaks even aberrant QRS complexes in noise-corrupted ECG signal compression.

References:

[1] Task Force Eur. Soc. Cardiol., N. Am. Soc. Pacing and Electrophysiology, “Heart rate variability. Standards of measurement, physiological interpretation, and clinical

[2] I. Christov, I. Jekova, and G. Bortolan, “Premature ventricular contraction classification by theKth nearest-neighbours rule,” *Physiol. Meas.*,vol. 26, pp. 123–130, 2005.

[3] J. Zhou, “Automatic detection of premature ventricular contraction using quantum neural networks,” in *Proc. 3rd IEEE Symp. Bioinformatics and Bioengineering*, 2003, pp. 169–173.

[4] B.U. Kohler,C. Hennig, and R. Orglmeister. The principles of software QRS detection. IEEE Eng. Med. Biol. Mag, vol. 21, no. 1, pp. 42-57, 2002.

[5] Benitez, P.A. Gaydecki, A. Zaidi, and A.P. Fitzpatrick, “The use of the Hilbert transform in ECG signal analysis,” Computers in Biology and Medicine, vol. 31, Sep. 2001, pp. 399-406.

[6] Chatterjee H.K, Gupta R, Mitra, A statistical approach for determination of time plane features from digitized ECG, Comp. in Bio and Med. 2011, vol 41; 278–284.

[7] Widjaja D, Vandeput S, Taelman J, Bracker M, Otte R.A. et. al., Accurate R peak detection and advanced preprocessing of normal ECG for heart rate variability analysis, Computers in cardiology. 2011, 533-536.

[8] Adnane M, Jiang Z, Choi S, Development of QRS detection algorithm designed for wearable cardio-respiratory system, Comp. Methods in Biomed. 2009, vol. 93, 20-31.

[9] Berntson G, Bigger Jr J, Eckberg D, Grossman P, Kaufmann P, Malik M, Nagaraja H, Porges S, Saul J, Stone P, et al. Heart rate variability: origins, methods, and interpretive caveats. Psychophysiology 1997;34(6):623–648.

BIOGRAPHIES

Ali Tariq Bhatti received his Associate degree in Information System Security (Highest Honors) from Rockingham Community College, NC USA, B.Sc. in Software engineering (Honors) from UET Taxila, Pakistan, M.Sc in Electrical engineering (Honors) from North Carolina A&T State University, NC USA, and currently pursuing PhD in Electrical engineering from North Carolina A&T State University. Working as a researcher in campus and working off-campus too. His area of interests and current research includes Coding Algorithm, Networking Security, Mobile Telecommunication, Biosensors, Genetic Algorithm, Swarm Algorithm, Health, Bioinformatics, Systems Biology, Control system, Power, Software development, Software Quality Assurance, Communication, and Signal Processing. For more information, contact **Ali Tariq Bhatti** alitariq.researcher.engineer@gmail.com.

Dr. Jung H. Kim is a professor in Electrical & Computer engineering department from North Carolina A&T State University. His research interests includes Signal Processing, Image Analysis and Processing, Pattern Recognition, Computer Vision, Digital and Data Communications, Video Transmission and Wireless Communications.