7 Biomedical signal detection

Chapter: 7
Biomedical signal detection

7.1 Introduction

There are three distinct wave components in every cycle of an ECG viz. P, QRS complex and T (occasionally a Fourth Wave Component Viz U May Be Seen). For normal individuals each of these wave components has a specific wave shape corresponding to each lead, though slight but distinct variations can be noticeable. On the other hand, one can distinguish abnormal ECGs because of their difference in rhythm and/or morphologies. Thus by interpreting ECG, the manner of electrical conduction in the heart (normal or abnormal) can be understood. Hence for interpreting and ECG it becomes mandatory to detect the P, Q, R, S and T waves.

A lot of work has been done for the computerized detection of QRS complexes. The algorithms for QRS detectors can generally be divided into three categories: nonsyntactic, syntactic, and hybrid. The algorithms based on a syntactic approach are time-consuming; due to the need for grammar inference for each class of patterns hence the detectors are nonsyntactic. The syntactic methods [1], [3] use templates for identification, these methods are very tedious and show no improvement in performance. Under the non syntactic methods are the following papers [3], [4] Pan and Tompkins in their paper [1] developed a pre-processor scheme for QRS detection that is more or less considered a standard. However in the decision rule section, adaptive thresholding was used and for T wave detection, the slope criterion was used. This did not prove to be very reliable. Hamilton and Tompkins in their paper [3] used the same pre-processor technique as in [3] but modified the decision rule section; they used a two-dimensional event vector with peak signal level and elapsed time from the last fiducial mark as parameters. They also used refractory blanking and search back methods to improve the detection efficiency. However this approach was found to be slow.

In [5], authors have described a QRS complex detector based on the dyadic wavelet transform (D_y WT) which is robust to time-varying QRS complex morphology and to noise. A spline wavelet suitable for QRS detection can be designed. The scales of this wavelet are chosen based on the spectral characteristics of the electrocardiogram ECG signal.

In [1] there is developed a real time algorithm for detection of the QRS complexes of ECG signals, It reliably recognizes QRS complexes based upon digital analysis of

slope, amplitude, and width. A special digital band pass filter reduces false detection caused by the various types of interference present in ECG signals.

The detection of the QRS complex specifically, the detection of the peak of the QRS complex, or R wave in an electrocardiogram (ECG) signal is a difficult problem since it has a time-varying morphology and is subject to physiological variations due to the patient and to corruption due to noise.

The traditional approach to detecting heartbeats was to take the derivative of the ECG signal. The QRS complex is the sharp rise and fall in the waveform. Derivatives were used to detect the significant change in voltage of the ECG signal. Derivatives gave an accurate detection if the signal had been cleaned up of noise. New approaches, including neural network prediction and wavelet transforms, are able to analyze the signal with the noise intact.

This chapter includes introduction of ECG with characteristics points and features of real time detection. A brief overview of Traditional classical and transformed based methods is given. It also accounts for the work done by the researchers for detection using ANN. The MATLAB implementation of our proposed techniques is carried out.

7.2 QRS Complex

The QRS complex is the sharp rise and fall in the waveform. Derivatives were used to detect the significant change in voltage of the ECG signal. For a well-filtered ECG signal, derivatives were accurate as shown in **figure 7.1.**

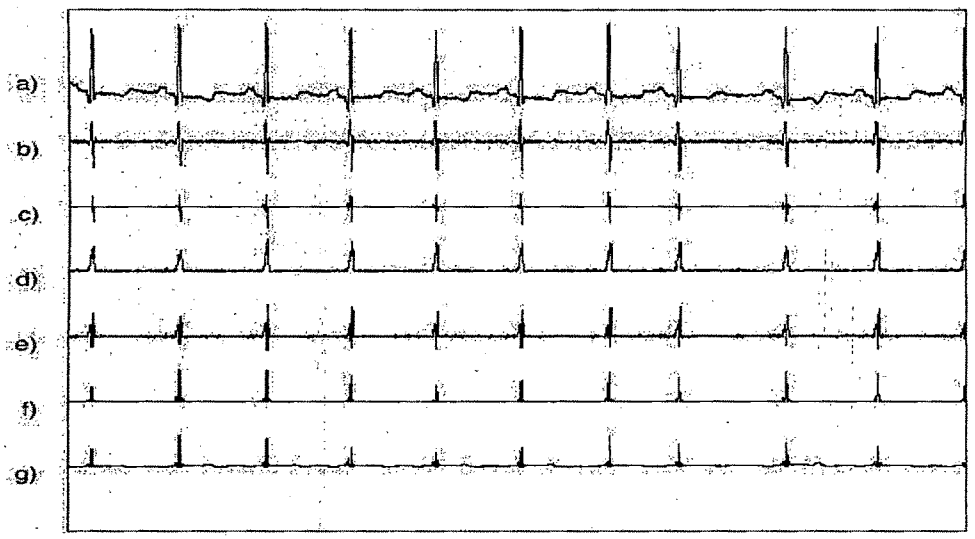


Figure 7.1 Output waveforms for the different peak detectors

The functions corresponding to the letters on the graph are listed in **table 7.1.** Letters b and c are based solely on an approximation of the first derivative while d and e include an approximation of the second derivative of the ECG signal. The second derivative was used to help reduce false positives from baseline drift, which is similar to a DC offset. x' and $\acute{y}$ are smoothed functions of the ECG signal.

Table 7.1 Description and formulae for Fig 7.2

Letter	Function	Description
a	$y(n) = x(n)$	Unfiltered ECG Signal
b	$y_1(n) = x(n+1) - x(n-1)$	First Derivative
c	$y_1(n) = x'(n) - x'(n-1)$ $X'(n) = \theta$ if $\lvert x(n)\rvert < \theta$, $\lvert x(n)\rvert$ otherwise	Smoothed First Derivative
d	$z(n) = 1.3\lvert y_1(n)\rvert + 1.1\lvert y_2(n)\rvert$ $y_2(n) = x(n+2) - 2x(n) + x(n-2)$	Linear Combination of First and Second Derivative
e	$z(n) = \acute{y}_1(n) + \lvert y_2(n)\rvert$ $\acute{y}_1(n) = \{0.25,0.5,0.25\} * \lvert y_1(n)\rvert$	Linear Combination of Smoothed First Derivative and Second Derivative
f	Results from [2].	
g	$z(n) = \Pi_k\lvert x(n-k)-x(n-k-1)\rvert$, from k= 0 to N-1	Multiplication of Backwards Difference (MOBD)

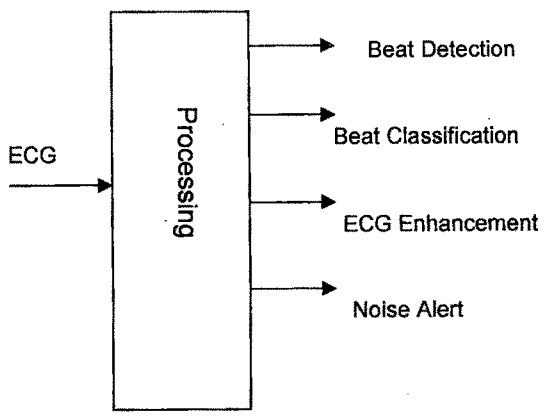


Figure 7.2 ECG Processing

7.2.1 QRS Detection

Various signal processing algorithms have been developed to process the ECG. Detecting QRS complexes in the ECG is one of the most important tasks that need to

be performed. This stage is crucial in basic ECG monitoring systems and also is important for all other ECG processing applications. Enhancement of the ECG is also important in a stress test. The stress ECG is prone to various types of noise, and it is important to reduce the noise without distorting the morphology of the ECG. Arrythmia classification is another important task in interpretive systems, which provide a diagnostic classification of the ECG. Another useful processing task is a noise alert algorithm which determines the fidelity of the ECG by indicating the level and type of noise in the signal **Figure 7.2** shows four out of the many tasks that must be performed on the ECG in different applications:

Automatic detection and classification of heart beats from the ECG waveform using biomedical signal processing techniques has become a critical aspect of clinical monitoring. These systems can, in real time, monitor the heart beat of a patient and alert clinicians when life-threatening conditions begin to surface. Many methods for automatic classification of various arrhythmias have recently been presented in literature including algorithms based on hidden Markov models, [7] self-organizing maps, [8] filter banks, [9] and neural networks. [6, 10,11] The results show that neural network classifiers are capable of the highest accuracy as local classifiers, but have not yet proved to be as successful when used as global classifiers.

Senhadji et al. also achieved a high accuracy (98%) but only used 25 beats for training and 28 beats for testing in their DWT / linear discriminates classifier. [12]. In contrast, de Chazal's group tested their classifier over the entire MIT / BIH database of 48 files but achieved an accuracy of only 89%. Their method of classification involved combining heart beat morphology with timing interval features for the training of their network. [6]

These results suggest that the combination of waveform shape and timing interval features is necessary for robustness while wavelet decomposition provides higher accuracy.

In [9] an algorithm for the determination of the limits of P and T waves is proposed, and its foundations are mathematically analyzed. The algorithm performs an adaptive filtering so that the searched point corresponds to a minimum. Immunity to base line drifts and full adaptation to any cardiological criteria are also taken care in the algorithm. A series of tests are made involving real registers with different morphologies for P and T-waves.

The base line of a typical ECG serves as an algorithm reference, setting the boundaries of ECG waves based on amplitude criteria. Base line drifts lead to deficient algorithm performance. This is of particular concern in exercise registers, where ECG recordings are greatly affected by the electromyogram.

The difficulty arises from the different criteria used by cardiologists in establishing the limits of the waves. For example: some set the end of T at the inflection point of the final part of the wave, while others consider the intersection between the base line and the tangent with maximum slope in this final area.

The implementation of an algorithm taking into account these often contradictory criteria is very complex. The algorithm developed in [9] overcomes both inconveniences: it is immune to base line drifts and is able to adapt to cardiologist criterion in determining the boundaries of P and T waves. It is also immune to Gaussian noise addition. The algorithm has been applied to obtain time-series in 5-min registers from healthy individuals. These series are extensively used in clinical research [10].

7.3 QRS Detection Techniques

The traditional approach to detecting heartbeats was to take the derivative of the ECG signal. The QRS complex is the sharp rise and fall in the waveform. Derivatives were used to detect the significant change in voltage of the ECG signal. Derivatives gave an accurate detection if the signal had been cleaned up of noise. Some classical techniques are described in this section.

7.3.1 Application of Detector

Most of the current QRS detectors are divided into two stages: a preprocessor stage to emphasize the QRS complex and a decision stage to threshold the QRS enhanced signal. Typically, the preprocessor stage consists of both linear and nonlinear filtering of the ECG. The ECG signal is first bandpass filtered to reduce noise and differentiated to emphasize the large slope of the R wave; it is then squared to further exploit the high-frequency content of the QRS complex. A short-time energy estimate is obtained by smoothing the resulting signal with moving window integration. The selection of the bandwidth of the bandpass filter and the duration of the sliding analysis window is not always straightforward. The bandwidth of the bandpass filter must be chosen to reflect the tradeoff between noise reduction and loss of high-frequency details; if the bandwidth is too large, noise reduction suffers. If the bandwidth is too narrow, high-frequency QRS characteristics are lost.

A detector is presented in [11] which finds changes in the repolarization phase (ST-T complex) of the cardiac cycle. It operates by applying a detection algorithm to the filtered root mean square (RMS) series of difference between the beat segment (ST segment or ST-T complex) and an average pattern segment The detector has been validated using the European ST-T database, which contains ST-T complex episodes manually annotated by cardiologists, resulting in sensitivity /positive predictivity of 85/865 and 85/76% for ST segment deviation and ST-T complex changes respectively.

The proposed detector includes signal preprocessing, computation of the root mean square (rms) difference series, filtering, and a decision algorithm which finds the ischemic events, as per **Figure 7.3**

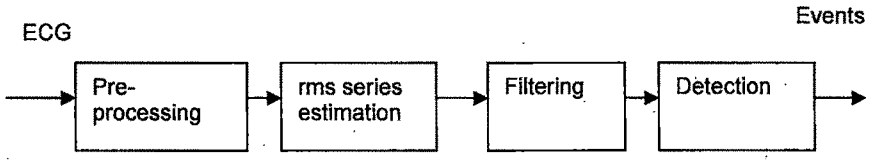


Figure 7.3 Detector Design

The preprocessing consisted of QRS detection and normal beats selection according to the arrhythmia detector and baseline wander attenuation using cubic splines, and rejection of noisy beats. In order to avoid the influence of high frequency noise in the rms difference series (e.g., 50/60 Hz noise), the ECG is low-pass filtered using a linear phase FIR filter (cutoff frequency 25 Hz). Beat segmentation is done by selecting intervals of 50 and 300 ms for the ST segment and ST-T complex, respectively, beginning at a distance from the QRS fiducial point dependent on the RR interval. These intervals definitions, related to the QRS fiducial point, avoid the always problematic estimation of the J point to define the repolarization windows, although consider the heart rate effects.

The proposed detector has a performance similar to those which have a more complicated structure. The detector has the advantage of finding both ST segment deviations and entire ST-T complex changes hereby providing a wider characterization of the potential ischemic events. A post-processing stage, based on a cross-correlation analysis for the episodes in the RMS series, is presented. With this stage sub clinical events with repetitive pattern were found in around 20% of the recordings and improved the performance to 90/85% and 89/76% for ST segment and ST-T complex changes respectively.

The detector performance needs to be evaluated by comparing the cardiologists' annotations and the detector output with regard to the following aspects:

- Detection rate
- Duration
- Magnitude of detected episodes

First two aspects are evaluated in terms of sensitivity and positive predictivity for both detection rate and duration respectively. The third aspect, related to the accuracy in the episodes magnitude estimation, is measured by comparing event-by-event the annotated amplitudes of the episodes (deviation peak as measured by the cardiologists) to the values obtained by the detector. The estimated linear correlation coefficient, between the two sets provides a measure of the detector linearity. Two kinds of statistics are commonly used for detector validation: *gross statistics*, in which the episodes of all patients are assigned equal weights, and *averaged statistics*, in which every patient is assigned equal weights.

Ischemic heart disease constitutes one of the most common fatal diseases in the western hemisphere. Myocardial ischemia is caused by a lack of sufficient blood flow to the contractile cells and may lead to myocardial infarction with its severe sequellae of heart failure, arrhythmias, and death. Ambulatory monitoring of the electrocardiographic (ECG) signal has become the noninvasive test most widely used for detecting cardiovascular diseases. Ischemic ECG changes typically precede the onset of anginal pain and, hence, these may be the only sign of "silent myocardial ischemia" [13]. Therefore, it is essential to develop methods that detect early changes in the ECG, possibly indicating the onset of an acute ischemic syndrome.

Different ECG changes related to the evolution of ischemia have been described, including T wave amplitude changes, ST deviations and even alterations in the terminal portion of the QRS complex [14]. In different situations T wave changes could precede ST segment deviations during the ischemic process [14], [15] and, therefore, should be considered in monitoring systems. The use of global representations for the ST-T complex instead of using a single point from the ST segment better characterizes ischemic patterns [16], [17], and yields better identification of an occluded artery [18]. Unfortunately, commercial equipment usually considers a fraction of the whole repolarization period, i.e., the ST60 or ST80 point. Different algorithms have been designed for analyzing the ST segment, either in the ECG signal [19], [20] or in the averaged ECG [21]–[29].

[11] describes the design and validation of a system that detects changes either in the ST segment, or in the entire ST-T complex (including the T wave), thereby providing a wider characterization of ischemic events. It also describes the different parts of the detector, the database for validation and the performance measures. A cross-correlation study between episodes is also included.

7.3.2 Application of Filter Banks

[30] uses a multi rate digital signal processing algorithm to detect heartbeats in the electrocardiogram (ECG). The algorithm incorporates a filter bank (FB) which decomposes the ECG into sub bands with uniform frequency bandwidths. The FB-based algorithm enables independent time and frequency analysis to be performed on a signal. Features computed from a set of sub bands and a heuristic detection strategy is used to fuse decision from multiple one-channel beat detection algorithm. The overall beat detection algorithm has a sensitivity of 99.59% and a positive predictivity of 99.56% against the MIT/BIH database. Furthermore this is a real-time algorithm since its beat detection latency is minimal. The FB based beat detection algorithm also inherently lends itself to a computationally efficient structure since the detection logic operates at the sub band rate.

Beat detection typically incorporates a preprocessing filter which decomposes the ECG into a signal which maximizes the signal-to-noise ratio (SNR) of the QRS complex [1]. A nonlinear processing stage and moving window integrator (MWI) are used to compute a signal that emphasizes the energy of the QRS complex. Beat detection logic incorporates a history of signal peaks and noise peaks which are used to establish signal and noise levels, respectively. A threshold is then used to decide period of time corresponding to the average heartbeat interval elapses without a beat detection a "search-back" strategy is used to check the ECG again for the presence of a beat [3].

The filters used are designed to optimize the SNR of the QRS complex but the Information from other frequency components of the ECG is filtered out and cannot be incorporated into the beat detection logic. Hence the preprocessing filters are not useful to other ECG processing tasks. The search-back strategy sometimes results in a beat detection latency time of more than one heartbeats interval. This is not useful when immediate indication of the occurrence of a beat is needed. Recent and extensive work on the design and use of Fileter Banks is presented in the literature [31]–[32].

Figure 7.4 shows that a FB contains a set of analysis filters which decompose the bandwidth of the input signal into subband signals with uniform frequency bands.

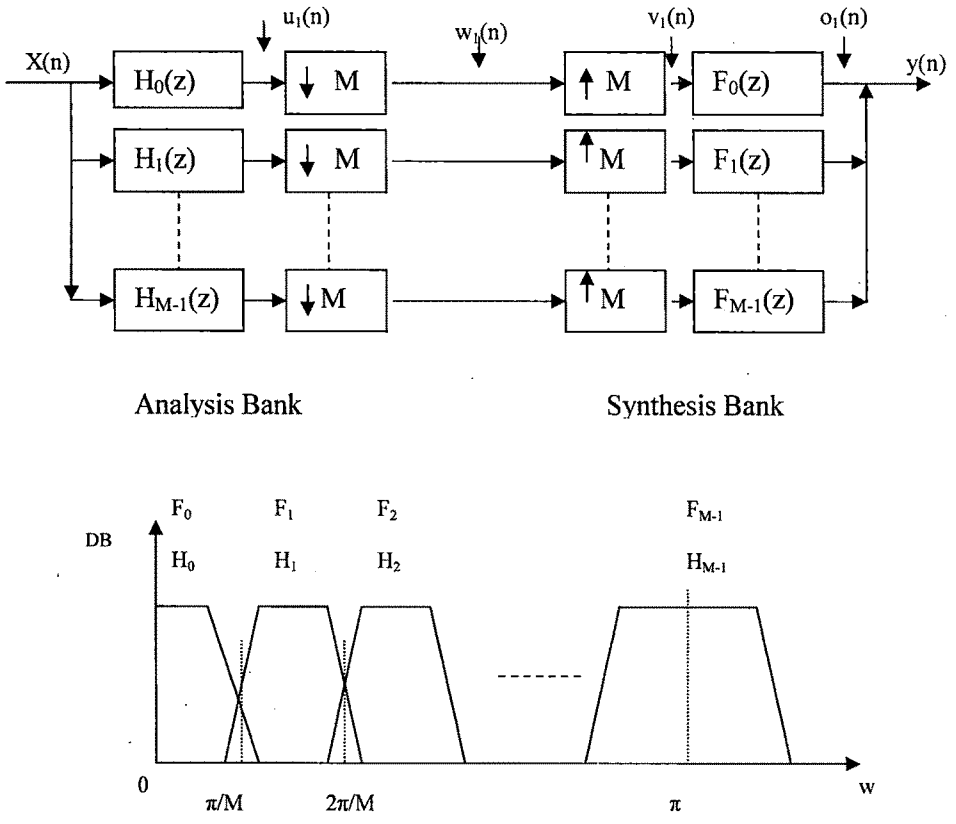


Figure 7.4 Filter Banks

The subbands can be downsampled since the subband bandwidth is much lower than that of the input signal. Processing can be performed on the subbands according to a specific application. Moreover, the subbands may be reconstructed by a set of synthesis filters which will perfectly reconstruct the input signal. **Figure 7.4** shows the ideal magnitude responses of the filters. The subbands provide information from various frequency ranges and, thus, it is possible to perform time and frequency dependent processing of the input signal. Because the subbands are downsampled, processing can occur at a lower rate than the input sampling rate. [17], [18] describe processing the subbands to reduce noise in the higher frequency subbands outside the QRS complex. The rationale in this is that there are no high frequency components of interest outside the QRS complex. This noise removal strategy is potentially useful to enhance the stress ECG. Thus, the FB allows time and frequency-dependent processing to be performed at a computationally efficient rate to analyze the ECG.

7.3.3 Real time Detection Algorithm

[1] describes a real time algorithm for detection of the QRS complexes of ECG signals, based upon digital analyses of slope, amplitude, and width. A special digital band pass filter reduces false detection caused by the various types of interference present in ECG signals. This filtering permits use of low threshold, thereby increasing detection sensitivity. The algorithm automatically adjusts thresholds and parameters periodically to adapt to such ECG changes as QRS morphology and heart rate. For the standard 24 h MIT/BIH arrhythmia database, this algorithm correctly detects 99.3 percent of the QRS complexes.

Software QRS detectors typically include one or more of three different types of processing steps. Linear digital filtering nonlinear transformation and decision rule algorithm. [1] use all three types. Linear process includes a bandpass filter, a derivative and a moving window integrator. The nonlinear transformation that is used is signal amplitude squaring. Adaptive threshold and T-wave discrimination techniques provide part of the decision rule algorithm.

The slope of the R wave is a popular signal feature used to locate the QRS complex in many QRS detector [33]. An analog circuit or a real time derivative algorithm that provides slope transformation is straightforward to implement. However, by its very nature, a derivative amplifies the undesirable higher frequency noise components Also, many abnormal QRS complexes with large amplitudes and long durations are missed in a purely derivative approach because of their relatively low R-wave slopes. Thus R-wave slope alone is insufficient for proper QRS detection. To achieve reliable performance, we must extract other parameters from the signal such as amplitude, width and QRS energy [34], [35].

7.3.4 An automated approach

An algorithm has been proposed in [36] to detect the QRS complexes and the T waves in an ECG signal. The algorithm incorporates a parameter viz SDA for identification of QRS complexes and T waves. This method has been mainly developed to detect the wave components when there are morphological changes rather than any rhythm changes.

7.4 Wavelet based detection methods

Wavelet transforms decompose the signal with respect to a wavelet family (mother and her many daughters). Daughter wavelets are scaled and translations of the mother wavelet. For QRS detection, typical wavelets include Harr, spline (quadratic, cubic, and higher powers), and Daubechies.

A short-time energy detector may be developed using a sliding analysis window. The choice of the duration of the sliding window results in a tradeoff between false and missed detections. A long window allows a large energy accumulation which easily exceeds a threshold, whereas narrow window duration allows too little energy to accumulate. In short, in the frequency domain, the fixed bandwidth of the bandpass filter cannot adapt to changes in the bandwidth of the QRS complex, and in the time domain, the fixed length of the moving window cannot adapt to changes in the duration of the QRS complex. A prefixed bandpass filter/short-time energy technique does not accurately account for the inherent time-varying morphology of the QRS complex i.e. they do not adapt very well to changes in QRS morphology. To overcome the limitations imposed by fixed duration windowing techniques in detecting time-varying transients, a general, adaptive technique that captures the spectral/temporal variations in QRS morphology is needed [36].

[36] uses the dyadic wavelet transform (Dy WT). A chosen "mother wavelet" has a fixed shape; however, the wavelet functions derived from it by changing scales, referred to as "daughter" wavelets, have different bandwidths and time supports. At any particular scale, the Dy WT is the convolution of the signal and a dyadically time-scaled daughter wavelet. Scaling the mother wavelet is the mechanism by which the Dy WT adapts to the spectral and temporal changes in the signal being analyzed. The Dy WT inherently has a multi resolution capability. For small scale values, it exhibits high temporal and low spectral resolution whereas for large scale values, it exhibits low temporal and high spectral resolution. The Dy WT was applied for edge detection and image compression, while in [37] it was used to extract pitch period information from voiced speech segments. The Dy WT has also been previously applied to ECG analysis in the context of 1) detecting ventricular late potentials (VLP's) [38], and 2) separating the various waves (P, R, and T) in the ECG [38]

A specific spline wavelet, suitable for the analysis of QRS complexes, is designed and the scales are chosen adaptively based on the signal. The Dy WT is computed at two consecutive scales and at one additional scale if necessary. It was observed that the Dy WT-based QRS detector is robust to noise.

The algorithm of the Dy WT-based QRS detector is as follows:

Algorithm computes the Dy WT of a windowed portion of length L_w seconds of an ECG signal at the dyadic scales $a=2^i$, $i=i_m, i_{m+1}, \ldots i_u$. L_w was set to 2.05 s (512 samples). The starting index i_m and the ending index were chosen based upon known

physical constraints. QRS complexes were detected by making use of the property that the absolute value of D_y WT has localized maxima across several consecutive scales at the instant of the occurrence of transients. For each scale there is located the local maxima of the absolute value of D_y WT (b, 2^i) which exceed a given local threshold. A choice of 60% as the local threshold value (i.e., 60% of the maximum value of the $|D_y$ WT| in each windowed portion of data) gave the best results. Hence, 0.6 was used as the local threshold value. Both the number and the locations of thresholded local maxima (peaks) of the $|D_y$ WT| were considered at the scale, $a=2^i$ and at the scale $a=2^{i+1}$. If the number of peaks is the same, and the location of the peaks align within ±25 samples (±0.1s) neighborhood of time across two consecutive scales, it was assumed that the locations of these maxima correspond to the location of possible QRS complexes. However, if the number of peaks agrees but the locations exceed ±25 samples time neighborhood, the thresholded data is pruned of the offending misaligned peaks. The locations of peaks that align across the two consecutive scales are stored in a vector of possible QRS complexes. If the number of peaks does not agree across two consecutive scales, the D_yWT is computed at the next dyadic scale, and the procedure described above is repeated while $I <= i_a$.

Next, the vector of possible QRS complexes was searched for refractory peaks. Any peak in the vector of possible QRS complexes occurring within a refractory period (0.2 s or 50 samples) of a previously thresholded peak was disregarded since the refractory period represents the interval immediately following a QRS complex during which no further excitation of the cardiac tissue can initiate another QRS complex. Thus, detection of false peaks could be minimized, and the locations represent the time of occurrence of the transient waves. The heart rate was then estimated by computing the inverse of the time interval between two consecutive waves.

The D_y WT is generally computed at scales 2^i for, theoretically, all i. However, one can restrict the range of scale parameters needed to compute the Dy WT based on the nature of the signal under study. Using the procedure outlined in [36], they designed a cubic spline mother wavelet with center frequency equal to 120 Hz and a bandwidth of 240 Hz. Since the average spectral support of the QRS complexes is 6–30 Hz and most of the spectral power of motion artifact and muscle noise is within 0–5 Hz as shown in [39], they chose the scales such that they cover the spectral support of QRS complexes. Such a selection of scales helps in filtering out noise.

Consequently, the algorithm starts with the scale, 2^{im} (index $i_m = 1$), corresponding to center frequency of 60 Hz and bandwidth 120 Hz, and continues up to a maximum scale, 2^{iu} (index $i_u = 3$), corresponding to center frequency of 15 Hz and bandwidth of 30 Hz. This is the upper limit. The algorithm does not compute the D_y WT beyond this scale. If the algorithm fails to find any matching peaks at this scale, then it declares that there are no QRS complexes in the data segment and proceeds to the next data segment. After choosing the lowest and highest scales, they compute the D_yWT at the lowest scale and double the scale parameter (if the number of peaks and their locations do not match at two successive scales) until the highest scale is reached. The flow chart of the algorithm is shown in **Figure 7.5.**

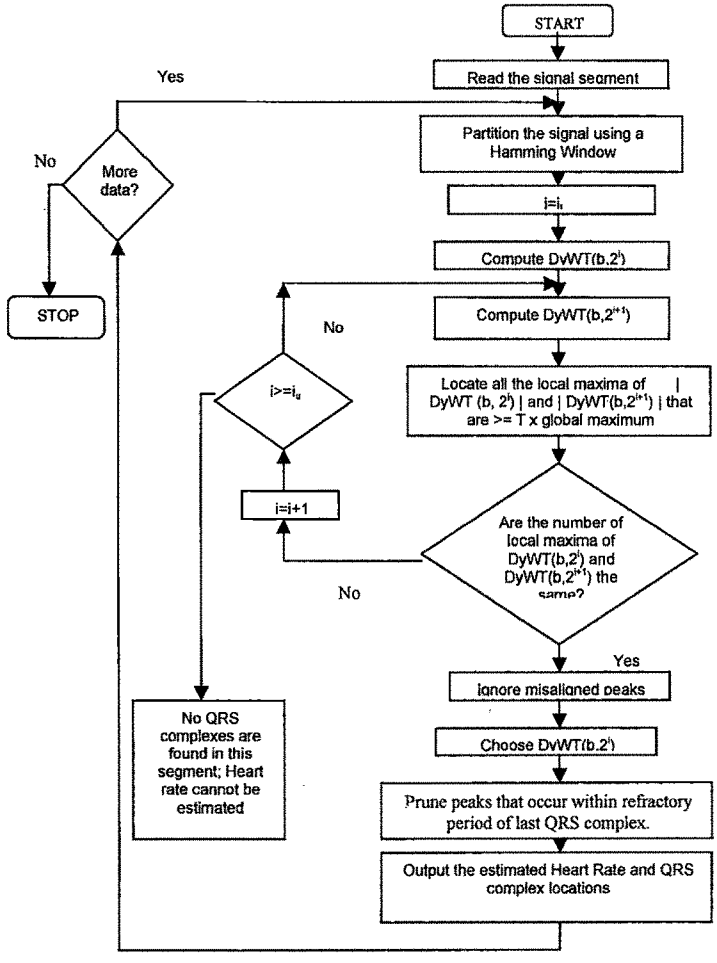


Figure 7.5 Flow chart of the D_yWT-based QRS detector

7.4.1 Detection of ECG characteristic points

A good performance of an automatic ECG analyzing system depends heavily upon the accurate and reliable detection of the QRS complex, as well as the T and P waves. The detection of the QRS complex is the most important task in automatic ECG signal analysis. Once the QRS complex has been identified, a more detailed examination of ECG signal, including the heart rate, the ST segment, etc., can be performed.

They [43] first filter the ECG signal with a bandpass filter (or a matched filter) to suppress the *P* and *T* waves and noise. Then the signal is passed through a nonlinear transformation to enhance the QRS complexes. The decision rules are used to determine whether QRS complexes are present in the signal. A matched filter [40] can improve the signal-to-noise ratio; its effect is limited by the variability of QRS waveforms for different beats of the same subject. The detection of the *P* wave is difficult because this wave is small and sometimes is embedded in noise. Gritzali proposed a simple method to detect *P* and *T* waves by "length transformation" [41], but it is not robust to noise.

Wavelet transform is a very promising technique for time frequency analysis [41]. By decomposing signals into elementary building blocks that are well localized both in time and frequency, the WT can characterize the local regularity of signals [42]. This feature can be used to distinguish ECG waves from serious noise, artifacts and baseline drift. In this paper, an algorithm based on the WT for detecting QRS complex, .P and T waves is proposed. A dyadic wavelet transform is used for extracting ECG characteristic points. The local maxima of the WT modulus at different scales can be used to locate the sharp variation points of ECG signals. The algorithm [43] first detects the QRS complex, then the *T* wave, and finally the *P* wave.

7.5 ANN based detection

The section describes methods used for detection based on Artificial Neural Network. ART2 is used to store pattern which are to be detected. It requires some preprocessing of ECG including cycle separation. They can serve two different functions when detecting QRS waveforms: adaptive filtering (Preprocessing) and waveform prediction.

7.5.1. Self-Organizing Network for QRS recognition

7.5.1.1 Introduction

[44] describes a self-organizing QRS wave recognition system for electrocardiograms

(ECG's) using neural networks. An ART2 (adaptive resonance theory) networks are employed in this self organizing neural networks system. The system consists of preprocessor detects R points in the ECG and divides the ECG into cardiac cycles. A QRS-wave is the part of the ECG that is between a Q points and an S point. The input to the ART2 network is one cardiac cycle from which the ART2 network indicates the approximate locations of both the Q and S points. The recognizer establishes search regions for the Q and S points. Then, it locates the Q and S points in each search region. The system uses this method to recognize a QRS wave. Then, the ART2 network learns the new QRS wave pattern from the incoming ECG. The ART2 network self organizes in response to the input ECG. The average recognition error of the present system is less than 1 ms in the recognition of the Q and S points.

7.5.1.2 The Principle of QRS-Wave recognition with ART2 network

It was assumed in [44] that the portion of the ECG from the R point to the Q point and the portion from the R point to the S point can be approximated by a straight line. This is only a preliminary approximation. After learning takes place with the ART2 network, this assumption is no longer necessary. One can find a well-fitted right-angled triangle pattern for the portion of the ECG.

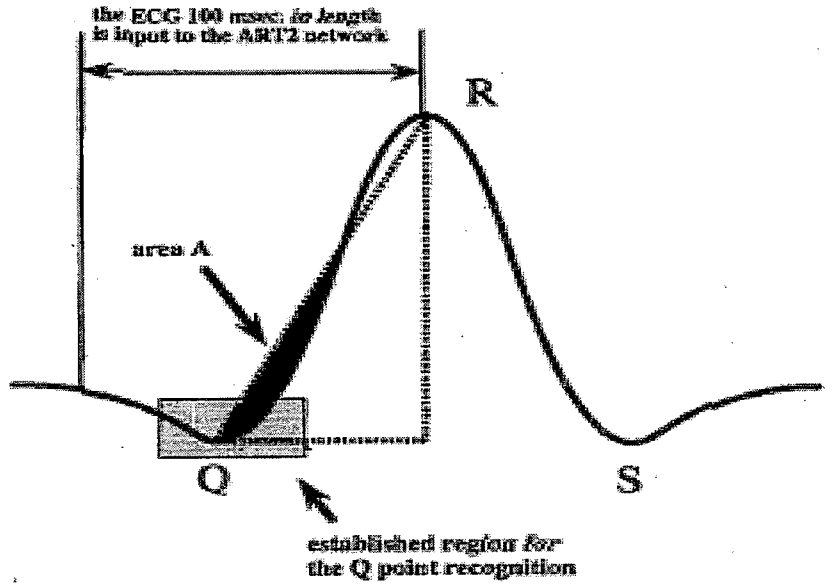


Figure 7.6 Q point recognition

A straight line approximating the portion of the ECG is the hypotenuse of the right-angled triangle pattern as shown in **Figure 7.6.** One hundred right-angled triangle

patterns are memorized by long- term memories **(LTM's)** of the ART2 network. The length of the base of each triangle patterns is different. In the input layer of the ART2, the input patterns are normalized within a predetermined level by the competitive interaction among neurons. An automatic gain control ensures that all the right-angled triangle patterns memorized in the LTM have the same height. These right-angled triangle patterns thus memorized are the initial template patterns for the ART2 network. As the portion of the ECG **100** ms in length from the R point in the direction towards the P point, as shown in **Figure 7.6,** is input to the ART2 network, the ART2 network associates the right-angled triangle pattern that is closest to the input ECG from those stored in the LTM's. The ART2 network receives a raw ECG signal, and each **F1** neuron represents a 1 ms sample of the ECG. At that time, the left end of the associated right-angled triangle pattern indicates the approximate location of the Q point in the ECG. The approximate location of the S point is determined in a similar manner. The recognizer of the system establishes search regions for the Q and S points, i.e., 4 ms before and after the points indicated by the ART2 network. The recognizer locates the Q and S points in each established region. The recognizer considers a location of the ECG as the Q point when the slope

$$\text{Slope} = (x(nT) - x(nT + T)) / T \qquad \ldots\ldots (7.1)$$

is less than 16.0 in the Q point search region, where T is the sampling rate of the ECG. The S point is the location where the following condition is satisfied

$$(x(sT + nT) - x(sT + nT + T)) / T =$$
$$<0 \quad n = -2, -1, 0$$
$$>0 \quad n = 1, 2 \qquad \ldots\ldots (7.2)$$

where s is the location of the S point in the search region. If there is no location in the ECG satisfying (7.1) or (7.2), the system considers that the approximate locations indicated by the ART2 network are the Q and S points, respectively. These two conditions are determined experimentally. In the search region, when there are a few locations that satisfy (7.1) or (7.2), the system may recognize a point as the Q and S point that is not the true Q or S point. After completing a QRS-wave recognition of one cardiac cycle, the ART2 network changes the pattern of the right- angled triangle associated by obtaining new information from the input ECG. The ART2 network extracts the features of the input ECG pattern and changes stored patterns in the LTM so that the ART2 network will be able to create new template patterns for the recognition of the QRS-wave. The patterns changed in the LTM replace the template

patterns of the initial values of the LTM. The right-angled triangle patterns are used as the initial template patterns of the QRS-wave. As the system self-organizes with respect to the incoming ECG patterns, the template patterns no longer retain the right-angled triangular shape. With the present algorithm, however, the initial length of the base of the right-angled triangle patterns is held even if the initial template patterns change according to the incoming ECG. As the process goes on, many characteristic QRS-wave patterns, which are different depending on each patient, will be stored in the LTM. Successive Q and S points recognition is carried out using these stored patterns. This is the self- organizing process of the present system in response to newly input ECG's. The vigilance parameter determines whether the input pattern is the same as a pattern stored in the LTM. When the matching rate between two patterns (one is the input to the system and the other one is the stored pattern in the LTM) is higher than the vigilance level, the system considers the two patterns as the same. Therefore, when the matching rate between QRS-waves from a different patient is higher than the vigilance level, learning is carried out by lumping together these QRS-waves. It is considered that the two patterns are the same when the matching rate between the two Patterns is higher than the vigilance level. Learning by lumping similar patterns will have no effect on correct pattern recognition. When a physician wants to know the characteristic QRS-wave patterns of a patient for diagnosis, clusters of the LTM are added for each patient to the system. The QRS-wave classification of each patient is performed using a corresponding cluster. At that time, one cluster of LTM stores QRS-wave patterns for one patient.

7.5.1.3 ART2 network for QRS-Wave recognition

Carpenter and Grossberg designed the neural-network model of the ART2 network. The ART2 network self-organizes in response to newly input patterns [45]. The ART2 network consists of an attentional subsystem and an orienting subsystem.

The attentional subsystem (**Figure 7.7**) recognizes the input patterns through a hypothesis testing cycle. It consists of a feature detector (F1), a category representation (F2), and a set of bottom-up and top-down LTM's. The connections between the neurons in the F1 layer and the corresponding neurons in the F2 layer form the LTM. A pathway from the F1 layer to the F2 layer is the bottom-up LTM. The pathway from the F2 layer to the F1 layer forms the top-down LTM. In the hypothesis testing cycle, the F1 layer receives an input pattern and extracts the features of the input pattern using a Non-linear signal function.

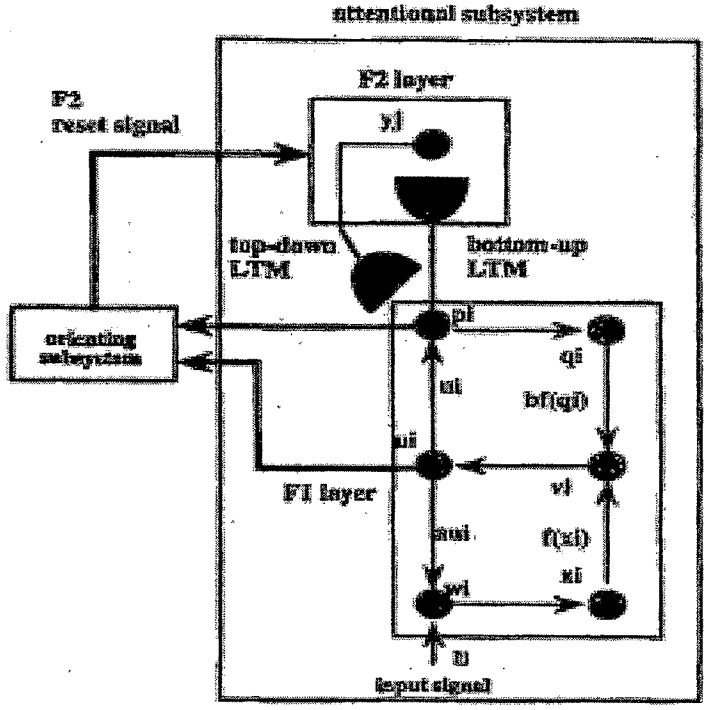


Figure 7.7 Basic architecture of ART2 network

7.5.1.4 Results

Some fiducial point recognition systems have been developed in the past. In these systems, an ECG 10 s in length is taken and divided into cardiac cycles. The cardiac cycles are stacked one over another to reduce noise. The fiducial points are recognized in one cardiac cycle of the ECG that emerged as a result of the stacking. They are the average of all the fiducial points in the cardiac cycles before stacking. The present system, however, recognizes the Q and S points in each cardiac cycle before stacking. The past systems recognize the fiducial points by slope detection. They have compared the recognition rate of the present system with the recognition rate of the slope detection method. Since the ART2 network is a self-organizing neural network to the input patterns, no training is needed for pattern recognition. We gave right-angled triangle patterns to the LTM's of the ART2 network as the initial values. The right- angled triangle patterns are used to locate the approximate location of the Q or S point. The recognition error rate of the QRS-wave by the present system was compared with the recognition error rate of the slope-detection method. In the slope-detection method, high- frequency disturbance is removed with a low-pass filter. Then the Q and S points are recognized by examining the slope of the ECG from the R point to the Q or S point. The slope from the R point towards the P wave is

calculated for the Q point recognition. The S point is recognized in a similar manner. The low-pass filter was employed to remove high-frequency disturbance. The Q and S points are recognized when the conditions of (7.1) for the Q point and (7.2) for the S point are satisfied. The recognition error of the QRS- wave of the present system was evaluated by the following method. A part of the ECG including the QRS-wave, 100 ms in length from the R point in the direction towards the P wave and 100 ms in length from the R point in the direction towards the T wave was printed on paper. Both the Q and S points were recognized by the present system and by the slope-detection method. Both the Q and S points were visually recognized by a well-trained person. Visual recognition is carried out by examining the slope change of the ECG wave form and the wave form of the ECG before and after the points were visually considered to be Q and S points. In evaluating recognition errors, if we visually recognized the Q and S points clearly, we calculated the recognition error. They considered that the visually recognized Q and S points were the true Q and S points. Then, the recognition errors by the present system and slope detection were evaluated using the visually recognized Q and S points. A total of 1500 cardiac cycles taken from the ECG's of five subjects were processed to check recognition errors by both the present system and the slope-detection method.

7.5.2 Adaptive back propagation

7.5.2.1 Introduction

A Supervised neural network (NN) based algorithm was used for automated detection of ischemic episodes resulting from ST segment elevation or depression [46]. The performance of the method was measured in terms of beat-by beat ischemia detection and in terms of the detection of ischemic episodes. The algorithm used to train the NN was an adaptive back propagation (BP) algorithm. This algorithm drastically reduces training time when compared to the classical BP algorithm. The recall phase of the NN is then extremely fast, a fact that makes it appropriate for real time detection of ischemia episodes. The resulting NN is capable of detecting ischemia independent of the lead used. It was found that the average ischemia episode detection sensitivity is 88.62% while the ischemia duration sensitivity is 72.22%

7.4.2.2 Selection of the Training Set from the ST-T European Database-Ischemia Detection Rules

The training set was constructed using patterns from the ST-T European database. The ECG's included in this database are 2-h-long recordings digitized at 250 Hz. This database is intended to be used for the evaluation of algorithms for ischemia analysis based on ST and T wave changes. The database consists of 90 continuous two-channel records. The leads used included modified leads V1, V2, V3, V4, and V5 and

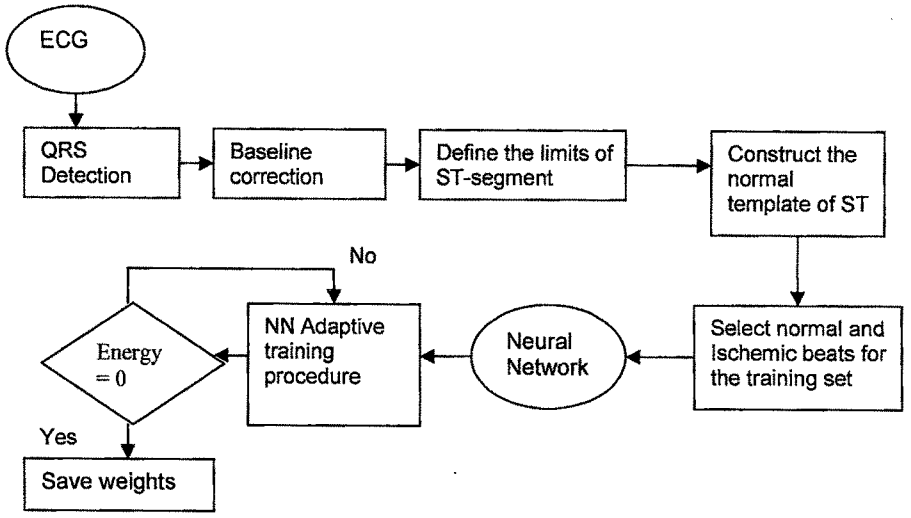
modified limb leads MLI and MLIII. Each record contains at least one ST or T ischemic episode. Specifically, for the ST episode annotations, the expert cardiologists used the following criteria:

1) ST segment deviations are measured in relation to a reference normal waveform selected from the first 30 s of each record.
2) Measurements of ST deviation are taken 80 ms after the J-point or in the case of tachycardia 60 ms after the J-point.
3) ST episodes must contain an interval of at least 30 s during which the absolute value of the ST deviation is no less than 0.1 mV.
4) The beginning and the end of each episode are annotated searching backward and forward, respectively, from the ST-episode, until a beat is found with absolute ST segment deviation less than 0.05 mV.

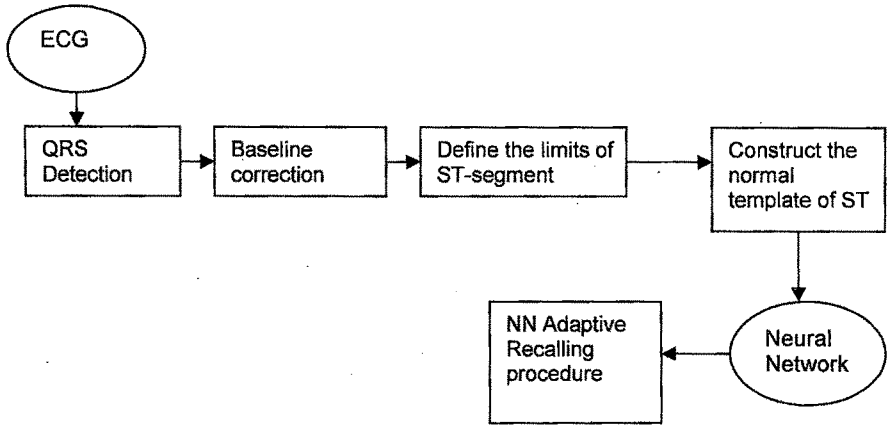
All conventions and definitions of the European ST-T database for ischemic episode classification were followed in this work. Specifically, the ECG recordings are broken into ten-beat groups. The number of normal and abnormal beats in each ten-beat group is then counted. If the number of normal beats is greater than or equal to the number of abnormal beats, then all beats of the group are assigned to the normal state. If the number of ischemic beats is greater than the number of normal beats, then all beats of the group are assigned to abnormal. Algorithm is designed to detect an ischemic segment if consecutive ten-beat groups are ischemic for at least 30 s. If this is not the case, a normal ECG segment is diagnosed. This way, the false detection of a single normal beat in an ischemic segment will not result in misclassification. One-hundred-twenty patterns, resembling normal and ischemic cases encountered in clinical practice, were selected from the European ST-T Database for the construction of the training set. Such patterns included flat ST depression, sloped ST depression, and noisy signals. Specifically, 50% of the patterns used were normal, 25% had ST depression, and 25% ST elevation. Ischemic patterns were chosen from the patterns belonging to the most depressed area of the ischemic episodes.

7.5.2.3 Description and Training of the NN

Figure 7.8 shows the general configuration of the system used. The NN used to identify the ST segment, consisted of three layers. The number of neurons in the input layer (N_i) was chosen to be 20. This value was determined experimentally, so as to accelerate and stabilize the performance of the NN. In particular, the 40 points normally present in a typical ST segment sampled at 250 Hz were reduced to 20 by dividing the input points into successive pairs, and then taking the average value of each pair as the input to the NN.



7.8 (a) Training Phase



7.8 (b) Recalling Phase

Figure 7.8 Flow for Adaptive back propagation

This accelerated both training and recall of the NN with no adverse effect on its classification accuracy. The number of neurons (Nh) in the hidden layer was chosen to be ten. The value of N_h was determined heuristically, so as to decrease the training time of the NN.

For the third layer (output layer), two neurons were used. The output of each of these neurons assumes a value between zero and one, which is rounded to one if higher than 0.5 or to zero otherwise. The resulting four output combinations serve to identify four beat classes: normal beats, ST depressed, ST elevated, or unclassifiable.

An adaptive BP algorithm was used for the training procedure [47]. This BP algorithm changes the weights of NN so as to minimize the error or energy function, defined by the equation

$$E = \sum_{k=1}^{M} \| o_k - t_k \|^2 \quad \text{..... (7.3)}$$

where M is the size of the training set, o_k the output vector of NN, and t_k the target vector for each training pair k . Each unit of the network uses the sigmoid activation function

$$f(x) = 1/(1+e^{-\alpha x+\beta}) \quad \text{..... (7.4)}$$

where α and β and are constants that determine the transition of the neural unit.

The procedure for the training of the NN is based on an adaptive algorithm with the parameter α changing, so as to help avoid entrapment into local minima. Specifically, as with classical BP methodology, purpose is the minimization of the energy function

$$\Delta w = \alpha * \frac{\partial E}{\partial w} \quad \text{..... (7.5)}$$

where w is the weight vector of the weights between the input and the hidden layers. If in the above

equation $\frac{\partial E}{\partial w} = 0$, a minimum has been reached.

7.5.2.4 Recall Phase of the NN

The main goal of the ECG preprocessing is to formalize the ST-segment in order to prepare an input suitable for the NN without loss of information. This was accomplished by computing the difference between an ischemic ST segment template and the normal (reference) template.

7.5.2.5 Results

There are defined two sets of performance indexes: 1) sensitivity and positive predictivity for ischemic ST episode detection and 2) sensitivity and positive predictivity for ischemia duration.

Correctly detected episodes are termed true positive (TP) episodes. Missed episodes are termed false negatives (FN). Erroneously detected nonischemic episodes are

termed false positives (FP). Finally, correctly identified normal beats are termed true negative (TN). Four indexes calculated in [46] are defined as follows.

1) Ischemic ST Episode Sensitivity (ST Se): Is defined as the ratio of the number of detected episodes matching the database annotations to the number of annotated ischemia episodes. This index expresses the sensitivity of the algorithm to the identification of ST episodes
2) Ischemic ST Episode Predictivity (ST P): Defined as the ratio of the number of correctly detected (matching) episodes to the number of episodes detected
This index is a measure of the inclination to incur false detection. Here, the denominator is the number of ischemic ST episodes detected by the NN algorithm.
3) Ischemia Duration Sensitivity (IS Se): Defined as the ratio of the duration of true matched ischemia to the total duration of annotated ischemia in the database
4) Ischemia Duration Predictivity (IS P): Defined as the ratio of the duration of true matched ischemia to the total duration of the ischemia detected by the NN

Since the total number of ischemia episodes in the database is small, aggregate statistics was used. The gross statistics are obtained by evaluating the numerators and denominators of the four indexes indicated above over the whole database. The average statistics are obtained by evaluating the above indexes for each file of the database separately, and averaging over the number of files. The indexes for the episode detection are better than the ones for duration of ischemia. In particular, the gross statistics give an episode detection sensitivity of 85% versus 88.62% of the average statistics method. The ischemic episode detection positive predictivity is 68.69% and 78.38%, respectively. On the other hand, the overall ischemia duration sensitivity stands at 73% for gross and at 72.22% for average statistics. The indexes for ischemia duration predictivity are 69.45% and 67.49% for gross and average statistics, respectively. Finally, the overall sensitivity and positive predictivity for the detection of the normal beats are 91% and 83%, respectively. From the above mentioned figures, it was observed that except in the case of episode predictivity, the performance of the algorithm is described consistently by all types of aggregate statistics.

Also, the average statistics produce better performance indexes as regards episode detection than gross statistics, while the opposite is true for the ischemia duration.

The performance indexes differ considerably from lead to lead, with the most problems encountered in lead MLIII (modified lead III). The lead with the best overall

performance as compared with the overall figures of merit, is modified lead I (MLI). Also NN-based algorithm can reliably detect ischemia episodes even in leads not included in the training set such as MLI.

7.5.3 Competitive Learning based detection

Multiplication of Backwards Difference (MOBD) is defined. MOBD keeps a value of the maximum of the signal. Whenever the signal drops by half of the maximum, a peak is detected and the running maximum value is halved. In a QRS complex, the QR portion would set the running maximum to the highest value and the RS portion is what would be detected. The problem with derivatives and MOBD is that they give too many false positives and true negatives as compared to more modern methods.

In competitive learning the output neurons of a neural network compete among themselves to become active (fired). Whereas in a neural network based on Hebbian learning several output neurons may be active simultaneously. In competitive learning only a single output neuron is active at any one time. It is this feature that makes competitive learning highly suited to discover statistically salient features that may be used to classify a set of input patterns.

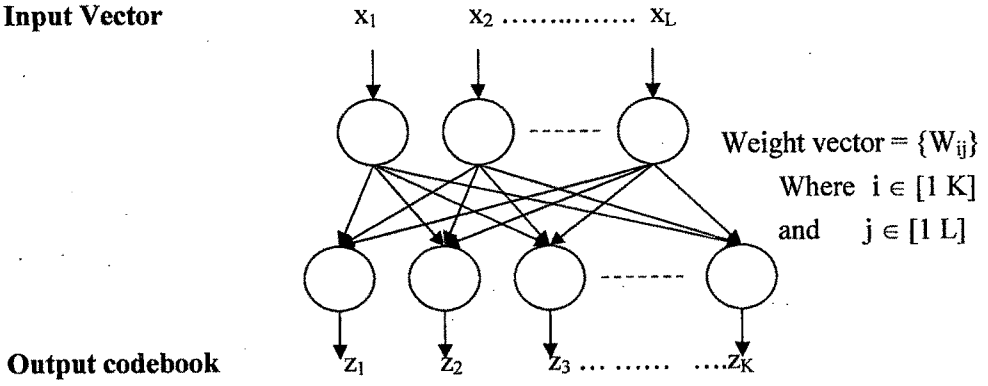


Figure 7.9 Competitive Learning Neural Network for QRS detection

In the simplest form of competitive learning, the neural network has a single layer of output neurons, each of which is fully connected to the input nodes. The network may include feedback connections among the neurons, as indicated in **Figure7.9.** In the network architecture the feedback connections perform lateral inhibition, with each neuron tending to inhibit the neuron to which it is laterally connected. In contrast, the feed forward synaptic connections in the network of Figure **7.9** are all excitatory.

When the synaptic weights are properly scaled they form a set of vectors that fall on the same N-dimensional unit sphere. However, for this function to be performed in a

"stable" fashion the input patterns must fall into sufficiently distinct groupings to begin with. Otherwise the network may be unstable because it will no longer respond to a given input pattern with the same output neuron.

During testing, under utilization problem occurs. Some neurons are left and never win the competition. Hence, frequency sensitive competitive learning algorithm is used. A record of how frequent each neuron is the winner is kept to maintain that all neurons in the network are updated an approximately equal number of times. To implement this scheme, the distance is modified to include the total number of times that the neuron i is the winner.

7.5.3.1 Algorithm for QRS Detection

Figure 7.10 depicts the template to be matched as QRS complex. During training weights are updated to come closer to input patterns. Each pattern appearing in input vectors will be assigned a unique index. Hence, the pattern matching the specified QRS complex will also be given unique index which will be stored as special value. During testing, for each pattern, again index is generated and matched with index of QRX complex. If matches count for number of QRS complex will be incremented.

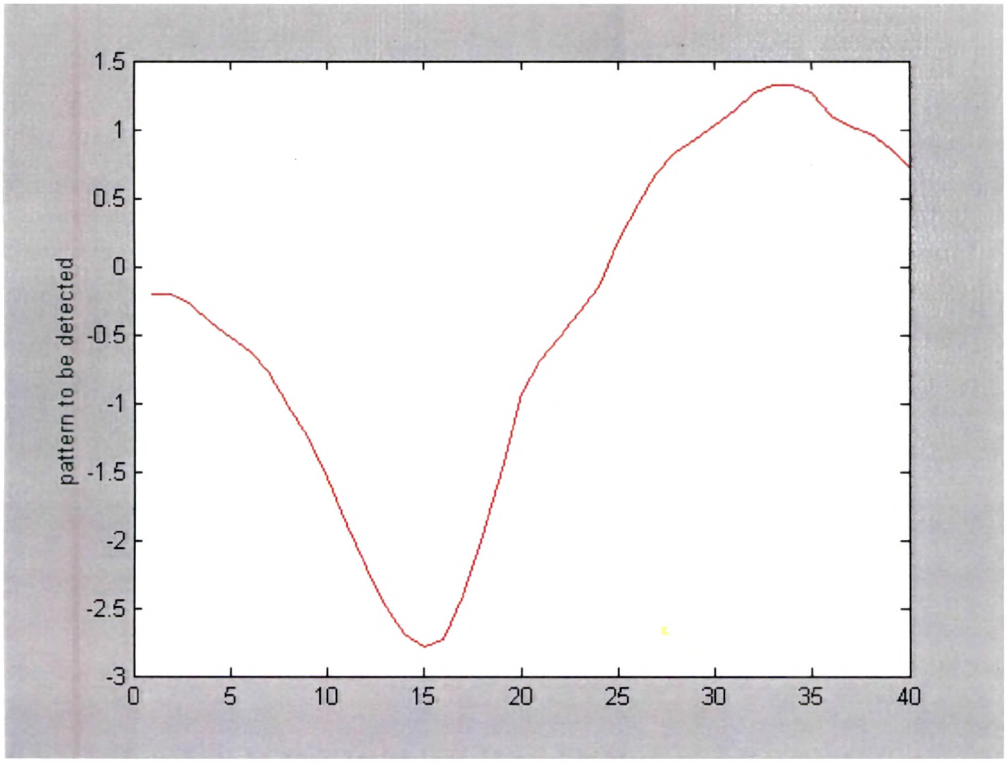


Figure 7.10 Template of QRS wave to be detected

7.5.3.2 MATLAB Implementation and waveforms

The competitive learning algorithm described in the section is implemented using MATLAB. The script is given below:

```
close all;
clear all;

%-------------------------------------------------------------------------------------------
I=40;
%H=200;%Hidden nodes
eta=0.02;
time=5;%time epoches
tests=200;
H=tests;
N=I+tests;
fid = fopen('c1ecg.dat','r');
[x,count]=fscanf(fid,'%f',N);
fclose(fid);
%from 120 to 159, pattern is to be detected
pattern=x(120:159);
figure(1);
plot(pattern);ylabel('pattern to be detected');
% Initialize weights:
w=ones(H,I)/I;
for j =1 :tests                    % tests:  no of training sets
for i = j:I+j-1
n=i-(j-1);
Up(n,j)= x(i,1);                   %overlapping
end
end
decode=zeros(I,H);
for t = 1:time
      t
f = ones(H,1);%frequency of winning
z = zeros(H,1);
location=0;
for l = 1:tests
min = sum(abs(Up(:,l)-w(1,:)'))*f(1);
```

```
loc=1;%Initialization for loc

%-----------------------------------------------------------------------------------------
%find loc:
%-----------------------------------------------------------------------------------------
for k = 1:H
if  min> sum(abs(Up(:,l)-w(k,:)'))*f(k)
min= sum(abs(Up(:,l)-w(k,:)'))*f(k);
loc = k;
else
loc=loc;                                %no operation required
end     %of if
 end   %of k

%-----------------------------------------------------------------------------------------
if Up(:,l)==pattern
location=loc;           % the exact location for pattern match
else
location=location;
end                                % of if
%-----------------------------------------------------------------------------------------
decode(:,loc)=Up(:,l);                  %store the vector
%-----------------------------------------------------------------------------------------
z(loc) = 1;
f(loc)=f(loc)+30;
for tin =1:4
for j = 1:I
w(loc,j)=w(loc,j)+eta*(Up(j,l)-w(loc,j)');
end                     %of j
end                             %of tests
end                             %tin
end                             %of t

%-----------------------------------------------------------------------------------------
                        %testing for the same input patterns:Up:
%-----------------------------------------------------------------------------------------
tests=2000;
N=I+tests;
```

```
fid = fopen('c1ecg.dat','r');
[x1,count]=fscanf(fid,'%f',N);
fclose(fid);
for j =1 :tests  % tests is no of training sets
for i = j:I+j-1
n=i-(j-1);
Up(n,j)= x1(i,1);%overlapping
end
end          %Up is training input to the MLP
code=zeros(tests,1);
ztest = zeros(H,tests);
count=0;
for i = 1: tests
min = sum(abs(Up(:,i)-w(1,:)'));
rank= 1;
for j = 1:H
op(j,i) = sum(abs(Up(:,i)-w(j,:)'));
if min>op(j,i)
min=op(j,i);
rank =j;
else
rank=rank;
end %of if
end     %OF j

%-------------------------------------------------------------------------------------------
if rank == location
count = count+1;
else
count = count;
end

%-------------------------------------------------------------------------------------------
end %tests
figure(2);
plot(x1);ylabel('signal for detection');
```

7.5.3.3 Parameters used

The parameters used for testing g and training are as follows:

- Number of input nodes I = No. of samples used to form input vector = 40
- Number of hidden layer nodes H= No.of tests performed = 200
- Learning Rate, $\eta = 0.02$
- Time epochs = 5
- Inner loop time epochs tin = 4
- Increment of frequency= 30

One important point about training and testing phases is the type of input fed to it as testing pattern. Patterns as input are at 1 sample advancement as compared to previous pattern. Testing is also done with same kind of shifted sample vectors only. Testing is carried out by taking signals of the same type but with different length. Shape of signal is shown in **Figure 7.11(a), (b)** and **(c).**

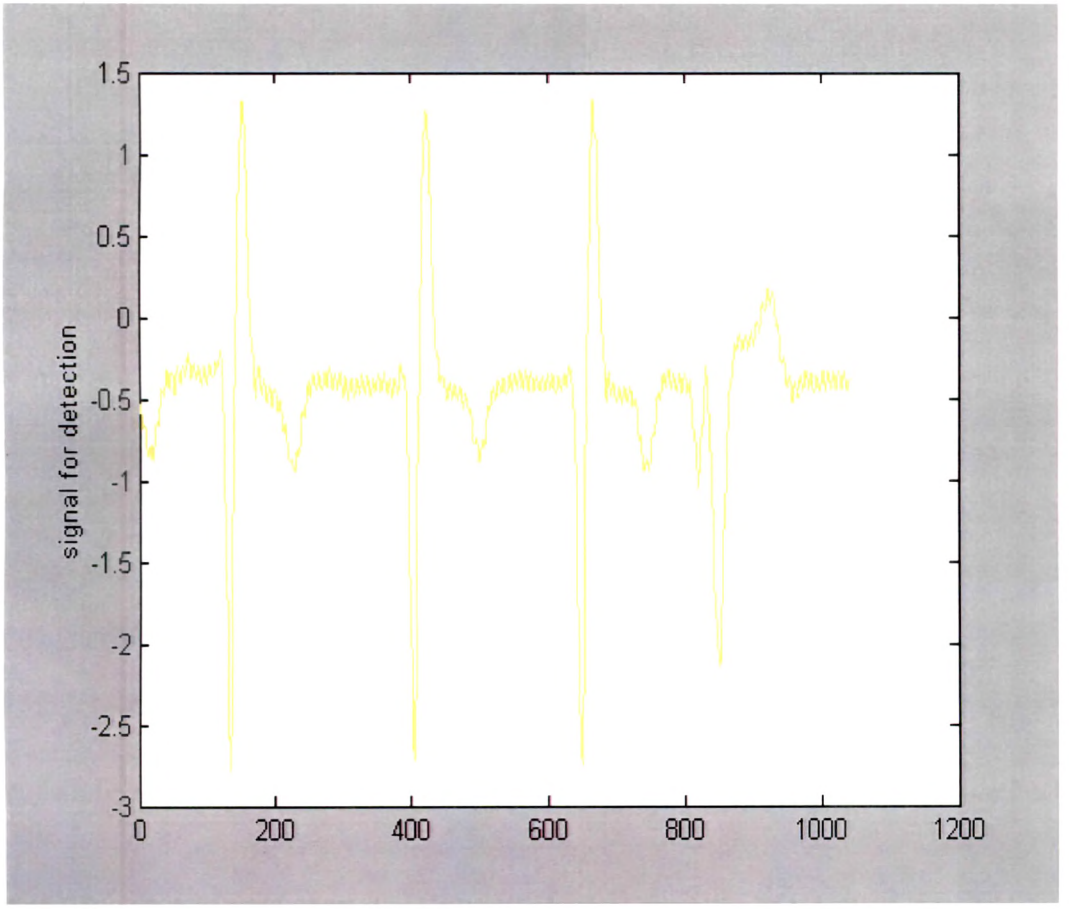


Figure 7.11(a) part of test waveform for which count=4

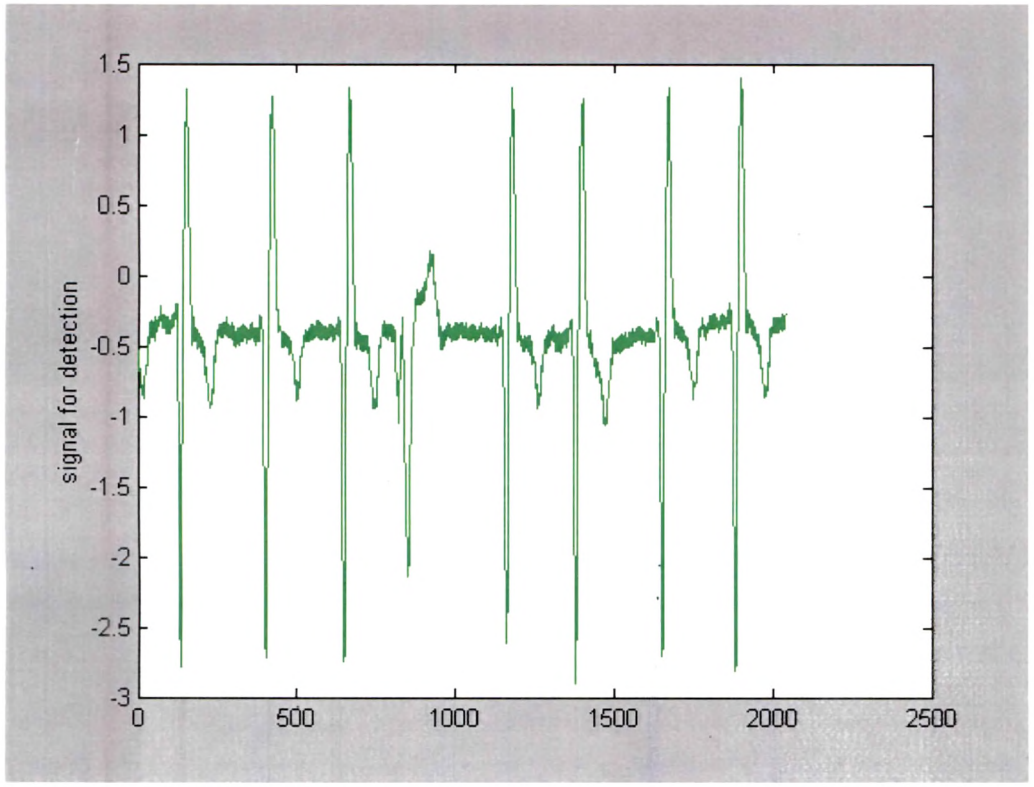


Figure 7.11(b) part of test waveform for which count=8

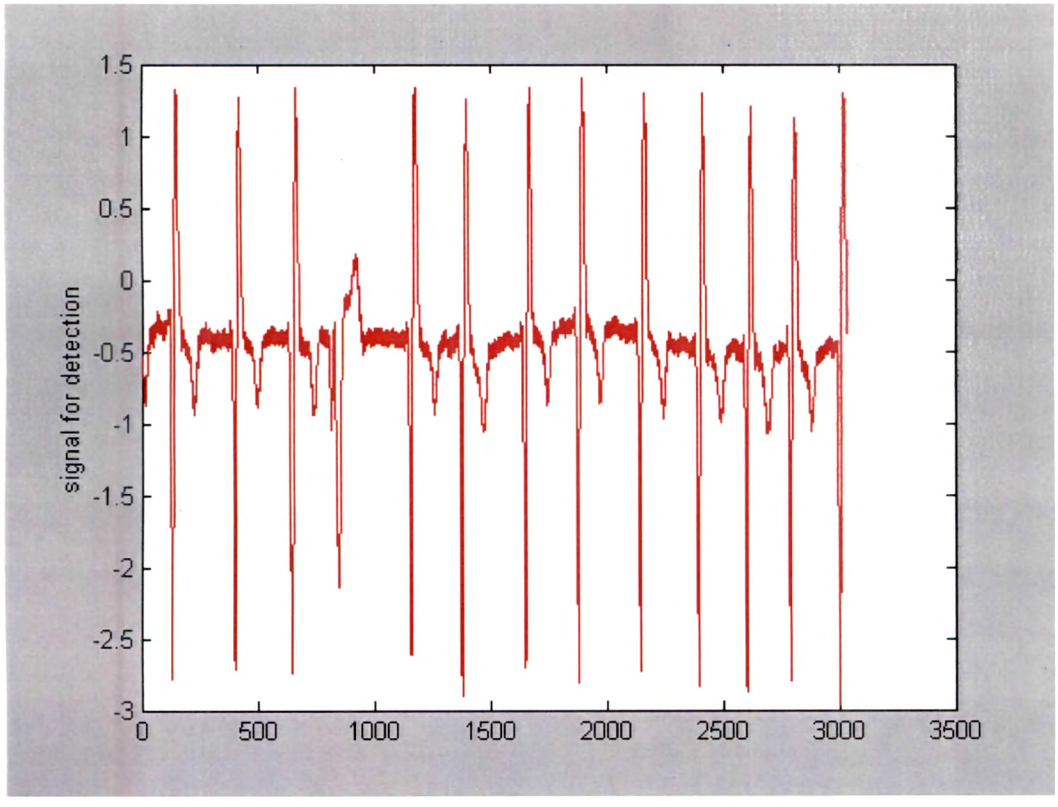


Figure 7.11(c) part of test waveform for which count = 13

Figure 7.11 Waveforms showing signal for detection by VQANN