CHAPTER - 5.

DETECTION OF ECG CHARACTERISTIC POINTS.
CHAPTER 5
DETECTION OF ECG CHARACTERISTIC POINTS.

The details regarding ECG are already discussed in Chapter 1. The various fiducial points like Q, R, S, P, T as well as the intervals are shown in Fig. 1.1(b). The accurate detection of these points by various methods is considered here.

5.1 R POINT AND QRS COMPLEX DETECTION.

The various methods for detection of R point are first considered. The simplest way would have been simple peak detector, but the reliability of such system is poor. Hence different methods are tried.

5.2 TOMPKIN'S METHOD FOR QRS DETECTION.

Pan and Tompkins develop a real time algorithm for QRS detection [24], which is further modified by Hamilton and Tompkins [15]. This is used for the R wave and QRS complex detection. As detailed in section 4.5, and as shown in Fig. 4.3, the various stages of such detector are i) A band pass filter prepared from a low-pass and a high-pass filter stage. ii) A derivative stage. iii) A squaring function stage, and iv) A moving window integrator.

The system was used to detect QRS complexes for different ECG waveforms available from MIT-BIH database as well as locally collected signals, by suitably changing the window size of the moving window integrator. For different sampling rates the window size is found out in number of samples, for a window of 150 ms. This corresponds to 30 samples for 200sps. The results are shown for a typical MIT-BIH file in Fig. 5.1 to Fig. 5.5. The ECG signal used (which is shown in Fig. 5.1), is a filtered signal as mentioned by Alste and Schilder [12], as detailed in section 4.5. The Fig. 5.2 is for result of low-pass and high pass filtering, which combines into a band-pass filter. Fig. 5.3 indicates output of derivative stage as well as squaring stage. The moving window integrator output is in Fig. 5.4, while Fig. 5.5 gives original signal and binarised window for R wave detection.

5.3 QRS DETECTION OR ECG BEAT DETECTION USING FILTER BANKS.

The method described by Afonso and Tompkins [25], for ECG beat detection and its results are already mentioned in section 4.8 and results are given in Table 4.2.
Figure 5.1. Filtered ECG signal.
Figure 5.2. Low-Pass and High-Pass Filter output.
Figure 5.3. Derivative stage and Squaring stage output.
Figure 5.4 Moving Window Integrator output.
Figure 5.5. Binarised window with original signal for R detection.
5.4 QRS DETECTION USING WAVELET TRANSFORM.

In this first some details of wavelet transform are considered, and then use of wavelet transform to detect QRS is discussed.

Wavelet Transform: -

Certain seismic signals can be modeled suitably by combining translations and dilations of a simple, oscillatory, function of finite duration called wavelet. It is also applied to in function representation, quantum mechanics, and signal processing.

A. Continuous-time Wavelet: -

Consider a real or complex value continuous time function $\psi(t)$ with the following properties.

1. The function integrates to zero

$$\int_{-\infty}^{\infty} \psi(t) \, dt = 0$$

2. It is square integrable or, equivalently has finite energy

$$\int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty$$

The function $\psi(t)$ is mother wavelet.

Property 1 is suggestive of a function that is oscillatory or that has wavy appearance and property 2 implies that most of the energy in $\psi(t)$ is confined to a finite duration.

Let $f(t)$ be any square integrable function. The continuous-time wavelet transform or CWT of $f(t)$ with respect to a wavelet $\psi(t)$ is defined as

$$\bar{W}(a, b) = \int_{-\infty}^{\infty} f(t) \, \frac{1}{\sqrt{a}} \psi^{*} \left( \frac{t - b}{a} \right) \, dt \quad \text{------------------------------------------(5.1)}$$

Where $a, b$ are real and $*$ denotes complex conjugation. Thus wavelet transform is a function of two variables. Both $f(t)$ and $\psi(t)$ belong to $L^2(R)$, the set of square
integrable functions, also called set of energy signals. Equation (5.1) can be written in more compact form by defining $\psi_{a,b}(t)$ as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t - b}{a}\right) \text{ dt}$$

(5.2)

Then equation 5.1 becomes

$$W(a,b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}^* (t) \text{ dt}$$

(5.3)

Notice that

$$\psi_{1,0}(t) = \psi(t)$$

(5.4)

The normalization factor of $1 / \sqrt{|a|}$ ensures that the energy stays the same for all $a$ and $b$; that is,

$$\int_{-\infty}^{\infty} \psi_{a,b}(t)^2 \text{ dt} = \int_{-\infty}^{\infty} |\psi(t)|^2 \text{ dt}$$

(5.5)

For all $a$ and $b$. For any given value of $a$ the function $\psi_{a,b}(t)$ is a shift of $\psi_{a,0}(t)$ by an amount $b$ along the time axis, thus the variable $b$ represents time shift or translation.

Form $\psi_{a,0}(t) = \frac{1}{\sqrt{|a|}} \psi \left(\frac{t}{a}\right)$

(5.6)

It follows that $\psi_{a,0}(t)$ is a time scaled and amplitude scaled version of $\psi(t)$. Since $a$, determines amount of time scaling or dilation, it is referred to as the scale or dilation variable. If $a>1$, there is stretching of $\psi(t)$ along the time axis, where as if $0<a<1$, there is a contraction of $\psi(t)$. Negative values of ‘$a$’ result in a time reversal in combination with dilation. Since the CWT is generated using dilates and translates of the single function $\psi(t)$, the wavelet for the transform is referred to as the mother wavelet.

The set of square integrable functions forms a linear vector space under addition and scalar multiplication. Furthermore, this vector space comes with a well-defined inner product. Given two finite energy signals $x(t)$ and $y(t)$, their product, denoted by $<x(t), y(t)>$, is given by
The total energy in \( x(t) \) is given by

\[
\int_{-\infty}^{\infty} |x(t)|^2 \, dt \quad \text{--- (5.8)}
\]

Which is \( <x(t), x^*(t)> \). The square root of energy is called norm of \( x(t) \) and is denoted by \( ||x(t)|| \). From equation (5.7) one can see that the CWT, is essentially a collection of inner products of a signal \( f(t) \) and the translated and dilated wavelet \( \psi_{a,b}(t) \) for all \( a \) and \( b \):

\[
W(a,b) = <f(t), \psi_{a,b}(t)> \quad \text{--- (5.9)}
\]

Although the time and frequency resolution problems are results of a physical phenomenon (the Heisenberg uncertainty principle) and exist regardless of the transform used, it is possible to analyze any signal by using an alternative approach called the multiresolution analysis (MRA). MRA, as implied by its name, analyzes the signal at different frequencies with different resolutions.

MRA is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. This approach makes sense especially when the signal at hand has high frequency components for short duration and low frequency components for long duration. Fortunately, the signals that are encountered in practical applications are often of this type.

As seen in the above equations, the transformed signal is a function of two variables, \( b \) and \( a \), the translation and scale parameters, respectively. \( \psi(t) \) is the transforming function, and it is called the mother wavelet. The term mother wavelet gets its name due to two important properties of the wavelet analysis as explained below:
The term wavelet means a small wave. The smallness refers to the condition that this (window) function is of finite length (compactly supported). The wave refers to the condition that this function is oscillatory. The term mother implies that the functions with different region of support that are used in the transformation process are derived from one main function, or the mother wavelet. In other words, the mother wavelet is a prototype for generating the other window functions.

The parameter scale in the wavelet analysis is similar to the scale used in maps. As in the case of maps, high scales correspond to a non-detailed global view (of the signal), and low scales correspond to a detailed view. Similarly, in terms of frequency, low frequencies (high scales) correspond to a global information of a signal (that usually spans the entire signal), whereas high frequencies (low scales) correspond to a detailed information of a hidden pattern in the signal (that usually lasts a relatively short time).

**B. Discrete wavelet transform:**

Continuous wavelet transform (CWT) maps a one-dimensional function \( f(t) \) to a function \( w(a,b) \) of two continuous real variables \( a \) and \( b \), which are the wavelet dilation and translation respectively. The region of support of \( W(a,b) \) is defined as the set of ordered pairs \( (a,b) \) for which \( W(a,b) \neq 0 \). In principle the region of support of a CWT is unbounded; that is it can be the entire plane defined by \( \mathbb{R}^2 \), the set of all ordered real pairs. The CWT provides a redundant representation of the signal in the sense that entire support of \( W(a,b) \) need not be used to recover \( f(t) \).

A type of representation, which is nonredundant wavelet representation, can be given in the form

\[
\begin{align*}
\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} d(k,l) 2^{-k/2} \psi(2^{-k}t - l) &= \cdots \cdots \cdots \cdots (5.10)
\end{align*}
\]

Observe that in contrast to CWT, which involves a continuum of dilation and translations, equation (5.10) uses discrete values for these parameters. The dilation take values of the form \( a = 2^k \) where \( k \) is an integer. At any dilation \( 2^k \), the translation parameter takes values of the form \( (2^k l) \) where \( l \) is again an integer. The values \( d(k,l) \) are related to values of the wavelet transform \( W(a,b) = W(f(t)) \) at \( a = 2^k \) and \( b = 2^k l \).
corresponds to sampling of coordinates \((a,b)\) or a grid in the way that consecutive values of the discrete scales as well as the corresponding sampling intervals differ by a factor of two. This process is called dyadic sampling. The two dimensional sequence \(d(k, 1)\) is commonly referred to as the discrete wavelet transform (DWT) of \(f(t)\) observe that DWT is still the transform of a continuous time signal. The discretization is only in the \(a\) and \(b\) variables.

The main idea is the same as it is in the CWT. A time-scale representation of a digital signal is obtained using digital filtering techniques. Recall that the CWT is a correlation between a wavelet at different scales and the signal with the scale (or the frequency) being used as a measure of similarity. The continuous wavelet transform was computed by changing the scale of the analysis window, shifting the window in time, multiplying by the signal, and integrating over all times. In the discrete case, filters of different cutoff frequencies are used to analyze the signal at different scales. The signal is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies.

The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations, and the scale is changed by upsampling and downsampling (subsampling) operations. Subsampling a signal corresponds to reducing the sampling rate, or removing some of the samples of the signal. For example, subsampling by two refers to dropping every other sample of the signal. Subsampling by a factor \(n\) reduces the number of samples in the signal \(n\) times.

Upsampling a signal corresponds to increasing the sampling rate of a signal by adding new samples to the signal. For example, upsampling by two refers to adding a new sample, usually a zero or an interpolated value, between every two samples of the signal. Upsampling a signal by a factor of \(n\) increases the number of samples in the signal by a factor of \(n\).

Although it is not the only possible choice, DWT coefficients are usually sampled from the CWT on a dyadic grid, i.e., \(a_0 = 2\) and \(b_0 = 1\), yielding \(a = 2^k\) and \(b = l \times 2^k\), as described before. Since the signal is a discrete time function, the terms function and
sequence will be used interchangeably in the following discussion. This sequence will be denoted by $x[n]$, where $n$ is an integer.

The procedure starts with passing this signal (sequence) through a half band digital low pass filter with impulse response $h[n]$. Filtering a signal corresponds to the mathematical operation of convolution of the signal with the impulse response of the filter. The convolution operation in discrete time is defined as follows:

$$
x[n] \ast h[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[n]
$$

A half band low pass filter removes all frequencies that are above half of the highest frequency in the signal. For example, if a signal has a maximum of 1000 Hz component, then half band low pass filtering removes all the frequencies above 500 Hz.

The unit of frequency is of particular importance at this time. In discrete signals, frequency is expressed in terms of radians. Accordingly, the sampling frequency of the signal is equal to $2\pi$ radians in terms of radial frequency. Therefore, the highest frequency component that exists in a signal will be $\pi$ radians, if the signal is sampled at Nyquist's rate (which is twice the maximum frequency that exists in the signal); that is, the Nyquist's rate corresponds to $\pi$ rad/s in the discrete frequency domain. Therefore using Hz is not appropriate for discrete signals. However, Hz is used whenever it is needed to clarify a discussion, since it is very common to think of frequency in terms of Hz. It should always be remembered that the unit of frequency for discrete time signals is radians.

After passing the signal through a half band low pass filter, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of $\pi/2$ radians instead of $\pi$ radians. Simply discarding every other sample will sub sample the signal by two, and the signal will then have half the number of points. The scale of the signal is now doubled. Note that the low pass filtering removes the high
frequency information, but leaves the scale unchanged. Only the subsampling process changes the scale. Resolution, on the other hand, is related to the amount of information in the signal, and therefore, it is affected by the filtering operations. Half band low pass filtering removes half of the frequencies, which can be interpreted as losing half of the information. Therefore, the resolution is halved after the filtering operation. Note, however, the subsampling operation after filtering does not affect the resolution, since removing half of the spectral components from the signal makes half the number of samples redundant anyway. Half the samples can be discarded without any loss of information. In summary, the low pass filtering halves the resolution, but leaves the scale unchanged. The signal is then subsampled by 2, since half of the number of samples are redundant. This doubles the scale.

This procedure can mathematically be expressed as

\[ y[n] = \sum_{k=\infty}^{\infty} b[k] \cdot x[2n - k] \] (5.12)

Having said that, how the DWT is actually computed is considered: The DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information. DWT employs two sets of functions, called scaling functions and wavelet functions, which are associated with low pass and high pass filters, respectively. The decomposition of the signal into different frequency bands is simply obtained by successive high pass and low pass filtering of the time domain signal. The original signal \( x[n] \) is first passed through a half band high pass filter \( g[n] \) and a low pass filter \( h[n] \). After the filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of \( \pi/2 \) radians instead of \( \pi \). The signal can therefore be subsampled by 2, simply by discarding every other sample. This constitutes one level of decomposition and can mathematically be expressed as follows:
\[ y_{\text{high}}[k] = \sum x[n] \cdot g[2k-n] \]
\[ y_{\text{low}}[k] = \sum x[n] \cdot h[2k-n] \]

Where \( y_{\text{high}}[k] \) and \( y_{\text{low}}[k] \) are the outputs of the high pass and low pass filters, respectively, after subsampling by 2.

This decomposition halves the time resolution since only half the number of samples now characterizes the entire signal. However, this operation doubles the frequency resolution, since the frequency band of the signal now spans only half the previous frequency band, effectively reducing the uncertainty in the frequency by half.

The above procedure, which is also known as the subband coding, can be repeated for further decomposition. At every level, the filtering and subsampling will result in half the number of samples (and hence half the time resolution) and half the frequency band spanned (and hence double the frequency resolution). Figure 5.6 illustrates this procedure, where \( x[n] \) is the original signal to be decomposed, and \( h[n] \) and \( g[n] \) are low pass and high pass filters, respectively. The bandwidth of the signal at every level is marked on the figure as "f".
Figure 5.6. The Subband Coding Algorithm.
As an example, suppose that the original signal $x[n]$ has 512 sample points, spanning a frequency band of zero to $\pi$ rad/s. At the first decomposition level, the signal is passed through the high pass and low pass filters, followed by subsampling by 2. The output of the high pass filter has 256 points (hence half the time resolution), but it only spans the frequencies $\pi/2$ to $\pi$ rad/s (hence double the frequency resolution). These 256 samples constitute the first level of DWT coefficients. The output of the low pass filter also has 256 samples, but it spans the other half of the frequency band, frequencies from 0 to $\pi/2$ rad/s. This signal is then passed through the same low-pass and high pass filters for further decomposition. The output of the second low pass filter followed by subsampling, has 128 samples spanning a frequency band of 0 to $\pi/4$ rad/s, and the output of the second high pass filter followed by subsampling has 128 samples spanning a frequency band of $\pi/4$ to $\pi/2$ rad/s. The second high pass filtered signal constitutes the second level of DWT coefficients. This signal has half the time resolution, but twice the frequency resolution of the first level signal. In other words, time resolution has decreased by a factor of 4, and frequency resolution has increased by a factor of 4 compared to the original signal. The low pass filter output is then filtered once again for further decomposition. This process continues until two samples are left. For this specific example there would be 8 levels of decomposition, each having half the number of samples of the previous level. The DWT of the original signal is then obtained by concatenating all coefficients starting from the last level of decomposition (remaining two samples, in this case). The DWT will then have the same number of coefficients as the original signal.

The frequencies that are most prominent in the original signal will appear as high amplitudes in that region of the DWT signal that includes those particular frequencies. The difference of this transform from the Fourier transform is that the time localization of these frequencies will not be lost. However, the time localization will have a resolution that depends on which level they appear. If the main information of the signal lies in the high frequencies, as happens most often, the time localization of these frequencies will be more precise, since they are characterized by more number of samples. If the main information lies only at very low frequencies, the time localization will not be very
Figure 5.7. ECG signal and Haar Detailed coefficients at Level 1.
Figure 5.8. Haar Detailed coefficients at Level 2 and Level 3.
precise, since few samples are used to express signal at these frequencies. This procedure in effect offers a good time resolution at high frequencies, and good frequency resolution at low frequencies. Most practical signals encountered are of this type.

The frequency bands that are not very prominent in the original signal will have very low amplitudes, and that part of the DWT signal can be discarded without any major loss of information, allowing data reduction.

Li et al. detected ECG characteristic points using wavelet transform [46]. With the multi-scale feature of the wavelet transform (WT), the QRS complex was distinguished from high P or T waves, noise, baseline drift or artifacts. The accuracy reported is 99.8% for MIT/BIH database. The P and T waves were also detected. Corresponding to 250 Hz as sampling rate, the power spectra for ECG signal, noise and artifact is plotted. The 3 dB bandwidth of equivalent filter Q1 at different scales is found out. For scale s1 = 21, it is 62.5 to 125 Hz, similarly for scales as 22, 23, 24, 25, the bandwidths are (18-58.5), (8-27), (4-3.5), (2-6.5), Hz respectively. The WT used was quadratic spline wavelet with compact support and one vanishing moment. It was found that energy of QRS was largest in (8-27) Hz bandwidth. It was shown that the corresponding to R wave two modulus maximum lines with opposite signs.

Practical Results and Discussions:— The QRS peak corresponds to central zero crossing of the modulus maximum pair, and the Q and S points are detected at the beginning and end of the negative and positive modulus (which are also the zero crossing points). This is seen in the Fig. 5.7, and 5.8. Here the Haar wavelet is used, and it is decomposed at scales 21, 22, 23, respectively and the relationship between modulus maxima and the QRS can be easily seen. At scales 22, the width is found out as 6 samples, which when reflected to 256 samples scale, becomes 24 samples, which is the width of QRS complex in samples. This is converted to the QRS width by multiplying the sample period which is 2.7778ms (360 Hz sampling frequency) in this case. The signal used for testing is one segment of x_300 signal available from MIT-BIH database. The sampling frequency is 360 Hz. Thus the sampling period is (1000/360) ms. The exact calculations of QRS width in this case will be 66.66 ms by Haar wavelet method. If the width of QRS is found out from time domain representation then it is found to be 63.8889
ms for x_300 same segment which is used for testing with Haar wavelet as well as by ART2 theory discussed in next section.

5.5 QRS DETECTION BY ADAPTIVE RESONANCE THEORY 2 (ART2) NEURAL NETWORK.

ART1 is the first model for adaptive resonance theory for neural networks developed by Gail Carpenter and Stephen Grossberg [70]. This theory was developed to address the stability-plasticity dilemma. The network is supposed to be plastic enough to learn an important pattern. But at the same time it should remain stable when, in short-term memory (STM), it encounters some distorted versions of the same pattern.

In ART1, classification of an input pattern in relation to stored patterns is attempted, and if unsuccessful, a new stored classification is generated. Training is unsupervised. There are two versions of training: slow and fast. They differ in the extent to which the weights are given the time to reach their eventual values. Slow training is governed by differential equations, and fast training by algebraic equations.

ART2 is the analog counterpart of ART1, which is for discrete cases. These are self-organizing neural networks, this is surmised from the fact that training is present but unsupervised. The ART3 model is for recognizing a coded pattern through a parallel search, and is developed by Carpenter and Grossberg. It tries to emulate the activities of chemical transmitters in the brain.

From the above mentioned neural network models, in the present system i.e. for the QRS-wave detection, ART2 neural network is used due to the following two main reasons:

1) ART2 network possesses the self-organizing property.
2) ART2 network addresses the stability-plasticity dilemma.

Thus ART2 is a class of adaptive resonance architectures which rapidly self-organizes pattern recognition categories in response to arbitrary sequences of either analog or binary input patterns. In order to cope with arbitrary sequences of analog input patterns, ART2 architectures embody solutions to a number of design principles, such as the stability-plasticity tradeoff, the search-direct access tradeoff, and the match-reset tradeoff. In these architectures, top-down learned expectation and matching mechanisms are critical in self-stabilizing the code learning process. A parallel search scheme updates...
itself adaptively as the learning process unfolds, and realizes a form of real-time hypothesis discovery, testing, learning, and recognition. After learning self-stabilizes, the search process is automatically disengaged. Thereafter input patterns directly access their recognition codes without any search. Thus recognition time for familiar inputs does not increase with the complexity of the learned code. A novel input pattern can directly access a category if it shares invariant properties with the set of familiar exemplars of that category. A parameter called the attentional vigilance parameter determines how fine the categories will be. If vigilance increases (decreases) due to environmental feedback, then the system automatically searches for and learns finer (coarser) recognition categories. Gain control parameters enable the architecture to suppress noise up to a prescribed level. The architecture's global design enables it to learn effectively despite the high degree of nonlinearity of such mechanisms.

5.5.1 ART2 system for analog input patterns.

For ART2 to match and learn sequences of analog input patterns in a stable fashion, its feature representation field F1 includes several processing levels and gain control systems. Bottom-up input patterns and top-down signals are received at different locations in F1. Positive feedback loops within F1 enhance salient features and suppress noise. Although F1 is more complex in ART2 than in ART1, the LTM equations of ART2 are simpler.

How the signal functions and parameters of the various ART2 architectures can best be chosen to categorize particular classes of analog input patterns for specialized applications is the subject of ongoing research. In particular, since ART2 architectures are desired to categorize arbitrary sequences of analog or digital input patterns, an arbitrary preprocessor can be attached to the front end of an ART2 architecture.
Figure 5.9. Typical ART2 architecture.
Specialized applications is the subject of ongoing research. In particular, since ART2 architectures are desired to categorize arbitrary sequences of analog or digital input patterns, an arbitrary preprocessor can be attached to the front end of an ART2 architecture.

In each ART2 architecture, combinations of normalization, gain control, matching and learning mechanism are interwoven in generally similar ways. Although how this is done may be modified to some extent, in all the ART2 variations, F1 needs to include different levels to receive and transform bottom-up input patterns and top-down expectation patterns, as well as an interfacing level of inter-neurons that matches the transformed bottom-up and top-down information and feeds the results back to the bottom and top F1 levels.

5.5.2 Properties of ART 2 Network

There are many properties of this type of network. A few of them are described below.

Stability-Plasticity Trade-Off:

Assume an arbitrary sequence of analog input patterns. Since the plasticity of an ART2 system is maintained for all time, and since input presentation times can be of arbitrary duration, STM processing must be defined in such a way that a sustained new input pattern does not wash away previously learned information. Removal or ablation of one part of the F1 internal feedback loop in Fig. 5.9 can lead to a type of instability in which a single input, embedded in a particular input sequence, can jump between categories indefinitely.

A. Search-Direct Access Trade-Off:

An ART2 system carries out a parallel search in order to regulate the selection of appropriate recognition codes during the learning process, yet automatically disengages the search process as an input pattern becomes familiar. Thereafter the familiar input pattern directly accesses its recognition code no matter how complex the total learned recognition structure may have become.

B. Match-Reset Trade-Off:

An ART2 system needs to be able to resolve several potentially conflicting properties which can be formulated as variants of a design trade-off between the requirements of sensitive matching and formation of new codes.
The system should, on the one hand, be able to recognize and react to arbitrarily small differences between an active F1 STM pattern established by a bottom-up input exemplar should be nearly identical to the learned top-down F2 → F1 expectation pattern in order for the exemplar to be accepted as a member of an established category. On the other hand, when an uncommitted F2 node becomes active for the first time, it should be able to remain active, without being reset, so that it can encode its first input exemplar, even though in this case there is no top-down/bottom-up pattern match whatsoever. A combination of an appropriately chosen ART2 reset rule and long term memory (LTM) initial values work together to satisfy both of these processing requirements. In fact, ART2 parameters can be chosen to satisfy the more general property that learning increases the system’s sensitivity to mismatches between bottom-up and top-down patterns.

5.5.3 ART2 Equations.

There are equations for STM, and LTM. Reset equations for orienting subsystem as well as equations for choice of initial top-down LTM values are also written. Implementation of property that learning increases mismatch sensitivity can be translated into suitable parameter constraints. For choice of a new category, bottom up LTM initial values, can be selected by suitable relations. The top down initial LTM values need to be chosen small otherwise top-down LTM read out by an uncommitted node can lead immediate F2 reset rather than learning of new category. Here only ART2 STM equations are given the remaining can be found out in [70].

**ART2 STM EQUATIONS : F1**

The following dimensionless equations (5.14) – (5.20) characterize the STM activities, pi, qi, ui, vi, wi, xi, computed at F1:

\[
pi = ui + \sum g(y_j) Z_{ji} \tag{5.14}
\]

\[
qi = pi / (e + \|p\|) \tag{5.15}
\]

\[
iu = vi / (e + \|vi\|) \tag{5.16}
\]
\[
vi = f( xi ) + b f( qi )
\]

\[
wi = li + aui
\]

\[
xi = wi / ( e + ||wi|| )
\]

Where \( ||V|| \) denotes the L_2 norm of a vector V [ i.e. \( ||V|| = ( \sum_{i} v_i^2 )^{1/2} \) ] and where \( y_j \) is the STM activity of the jth F2 node. The nonlinear signal function \( f \) in equation (5.17) is typically of the form

\[
f(x) = \begin{cases} 
2\theta x^2 / ( x^2 + \theta^2 ) & \text{if } 0 \leq x \leq \theta \\
x & \text{if } x \geq \theta 
\end{cases}
\]

which is continuously differentiable, or

\[
f(x) = \begin{cases} 
0 & \text{if } 0 \leq x \leq \theta \\
x & \text{if } x \geq \theta 
\end{cases}
\]

which is piecewise linear. The graph of function \( f(x) \) in equation (5.20) may also be shifted to the right, making \( f(x) = 0 \) for small \( x \), as in equation (5.21). Since the variables \( xi \) and \( qi \) are always between 0 and 1 [Eqs.(5.15) and (5.19)], the function values \( f(x) \) and \( f( qi ) \) also stay between 0 and 1. Alternatively, the signal function \( f(x) \) could also be chosen to saturate at high \( x \) values. This would have the effect of flattening pattern.
Similarly ART2 STM equations for F2, LTM equations, Reset equations, Match-Reset trade off equations, Mismatch sensitivity equations [70] can be solved in realisation of the ART2 system.

5.6 PRINCIPLE OF QRS WAVE RECOGNITION BY ART2.

In this section the principle of recognizing a QRS-wave using ART2 network is explained. It is assumed 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 approximation is no longer necessary. A well-fitted right-angled triangle pattern can be found for the portion of the ECG. A straight line approximating the portion of the ECG is the hypotenuse of the right-angled triangle pattern as shown in fig. (5.9). One hundred right-angled triangle patterns are memorized by long-term memories (LTM) 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.

The R point is already recognized by peak detection method, hence timings are stated with respect to R point only. As shown in Fig.5.9, input to ART2 network is the portion of the ECG 100 ms in length from the R point in the direction towards the P point. Then ART2 network associates the right-angled triangle pattern that is closest to the input ECG from those stored in the LTM. The ART2 network receives a raw ECG signal, and each FI neuron represents a 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 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,

\[
\frac{(x(nT) - x(nT + T))}{T}
\]

(5.22)
Figure 5.10. The principle of Q point recognition.

ESTABLISHED REGION FOR THE Q POINT RECOGNITION.
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,

$$x(sT + nT) - x(sT - nT + T)/T = \begin{cases} < 0 & \text{for } n = -2, -1, 0 \\ > 0 & \text{for } n = 1, 2 \end{cases}$$

where $s$ is the location of the S point in the search region. If there is no location in the ECG satisfying equations (5.22) or (5.23), the system considers that the approximate locations indicated by the ART2 network are the Q and S points respectively [71]. These two conditions are determined experimentally. In the search region, when there are a few locations that satisfy equation (5.22) or (5.23), 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 for 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.
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 different patients 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, the clusters created in lumping process 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.

The ART2 network consists of an attentional subsystem and an orienting subsystem. The attentional subsystem 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. 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 nonlinear signal function.

\[
f(x) = \begin{cases} 
2\theta x^2 / (x^2 + x^2 \theta^2) & \text{if } 0 \leq x \leq \theta \\
x & \text{if } x \geq \theta 
\end{cases} \quad (5.24)
\]

where \( \theta = 0.2 \) is determined experimentally [71]. The portion of an ECG 100 ms in length from the R point is truncated and input to the ART2 network as shown in fig.(5.8) to locate the approximate location of the Q or S point. At that time, the portion of the ECG from the Q or S point to the terminal of the truncated ECG may curve upward. This
upwardly curved section of the ECG cannot be approximated by a right-angled triangle pattern. The nonlinear signal function resolves this problem by inhibiting the curved section. The gain of the F1 layer is automatically controlled in response to the input pattern through competitive interactions among neurons. Therefore, the size of the input pattern is normalized into a predetermined dynamic range. When the activity of a neuron in the F1 layer is large enough, it generates an excitatory signal along the bottom-up LTM to the neurons in the F2 layer. At that time, the signal is multiplied by the LTM. The F2 layer is the competitive network that chooses the neuron receiving the largest input. Only the neuron thus chosen is activated and the other neurons are inhibited. The activated neuron associates a pattern through the top-down LTM at the F1 layer. When the associated pattern does not match the input pattern, the orienting subsystem sends a reset signal to the F2 layer so that the activated neuron may be strongly inhibited and may not be activated again until the hypothesis testing cycle ends. After that, the input pattern is reinstated and generates an output to the F2 layer again. When the associated pattern approximately matches the input pattern, the hypothesis testing cycle ends. After that, the bottom-up and top-down LTMs which are linked to the activated neuron in F2 learn new information about the input pattern. The F1 layer consists of 101 neurons and the F2 layers consist of 100 neurons. The parameters used in the ART2 network are shown in Table (5.1).

A vigilance parameter of the ART2 network regulates the coarseness of the categories of the input pattern. As the vigilance parameter becomes closer to 1.0, the system becomes more sensitive to the difference between the input pattern and the associated pattern. If the vigilance parameter is too high, the system regards the newly input pattern as a slightly different pattern. If this occurs, unnecessary patterns may be stored in the LTM. On the other hand, if the vigilance parameter is too low, the ART2 network will create too few categories. Therefore, an appropriate vigilance level should be determined.

There are 100 right-angled triangle patterns in the LTMs as the initial values. The ART2 network stores 100 right-angled triangle patterns as 100 different patterns when the vigilance parameter is set to 1.0, and a vigilance value of 0.0 during network performance ensures that the network resonates by choosing the category whose template
is closest to the incoming ECG without creating new categories. The vigilance value, however, is not set 1.0 during initialization, and it is not 0.0 during the network performance in the present system. When the system fails to detect the R point correctly, the pattern that does not include the Q or S point is input to the ART2 network. Then, the pattern input to the ART2 network is much different from the template patterns. In this case, the ART2 network does not resonate and put that irrelevant pattern into a new category. The vigilance value of 0.98 categorizes initial 100 right-angled triangle patterns as 100 different patterns and makes the ART2 network not to resonate to irrelevant patterns.

**AREA CONSIDERATIONS:**

Another condition is added in the present system for the QRS-wave recognition, which is expressed by equations (5.25) and (5.26).

\[
\sum_{i=0}^{100} \text{sum}(i) \leq \text{lim.}
\]

Equations (5.25) and (5.26) are not modification to the ART2 network. They are used for the QRS-wave recognition in addition to the ART2 network. When the ECG is input to the ART2 network, the network resonates by choosing the category whose template is closest to the incoming ECG. The vigilance parameter of the ART2 network judges the degree of match between two patterns. Even when the degree of the match between two patterns is higher than the vigilance level, another condition has to be satisfied to recognize the QRS-wave correctly. When the amplitudes of the right-angled triangle patterns are larger than those of the ECG, a section is created by the difference in amplitude between the right-angled triangle pattern and the ECG. The area of that section...
is temporarily called area A [see Fig. 5.10]. The larger the area A, the further the left end of the right-angled triangle pattern from the Q or S point. Area A is smaller than the predetermined parameter ‘lim’ which is actually limit for the area A. For the QRS-wave recognition, the ART2 network associates the template when the degree of match between patterns is higher than the vigilance level and area A is smaller than ‘lim’. In present system ‘lim’ is determined experimentally and has a value of 1.2 for Q point recognition and it is 1.75 for S point recognition.

5.7 VALUES OF THE PARAMETERS USED IN ART2 NETWORK

The parameters used in the ART2 network system for QRS-wave recognition are shown in Table 5.1 below:-

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Description of Definition</th>
<th>Parameter Notation</th>
<th>Parameter Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Feedback Parameter in F1 Layer</td>
<td>a</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Feedback Parameter in F2 Layer</td>
<td>b</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Parameter in Orienting Subsystem</td>
<td>c</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Activation Level in F2 Neuron</td>
<td>d</td>
<td>0.9</td>
</tr>
<tr>
<td>5</td>
<td>Threshold Parameter</td>
<td>( \theta )</td>
<td>0.2</td>
</tr>
<tr>
<td>6</td>
<td>Learning Rate</td>
<td>( \varepsilon )</td>
<td>0.1</td>
</tr>
<tr>
<td>7</td>
<td>Vigilance Parameter</td>
<td>( \rho )</td>
<td>0.98</td>
</tr>
<tr>
<td>8</td>
<td>Constant</td>
<td>e</td>
<td>0</td>
</tr>
</tbody>
</table>
Practical Results with ART2:

The results of ART2 for x_300 segment are as below.

degree of match is 0.997621 & req is 0.98

width of R from Q is 41 ms

width of R from S is 21 ms

total QRS width 62 ms.

5.8 ST SEGMENT DETECTION.

A supervised ANN based algorithm was developed by Maglaveras et al. [66], and is used for automated detection of ST segment elevation or depression. In order to minimize the probability of false detection of ST depression and to eliminate low frequency noise, the isoelectric level must be correctly identified. The method used is based on the assumption that the isoelectric level of the signal lies in the area approximately 80 ms left of R peak, where the first derivative becomes equal to zero for at least 10 ms or in the flattest 20 ms segment. The ST segment is assumed to begin at 60 ms after R peak in normal sinus rhythm case. The R peak is already detected by Pan and Tompkins algorithm [24]. In the case of tachycardia (RR interval < 600 ms), the beginning of ST segment is taken at 40 ms after the R peak. The ST segment for each heartbeat has a predefined length of 160 ms which means that the end point is 220 ms after R peak in normal case and 200 ms otherwise. After the isoelectric level is fixed the ST-segment measurement is done by three basic approaches [26]. These are,

1) \( J + x \) ms
2) \( R + x \) ms
3) Windowed search.

The \( J + x \) method (Fig. 5.11) is the most prevalent among researchers. The ST point is defined as the part of the ECG located at \( J + x \) ms. The J point is the first inflection point after the S point or may be the S point itself in certain ECG wave forms. Typical measurement is made at \( J + 80 \) ms or in case of tachycardia 60 ms after the J point.
Figure 5.11: (J+X)ms method

Figure 5.12: (R+X)ms Method

Figure 5.13: Windowed Search Method
In the windowed search method (Fig. 5.13), two end points are defined to the search window. Normally J + 20 ms and the onset of T wave are taken as the window boundaries.

The R + x method (Fig. 5.12) of locating the ST point uses the peak of the R wave as the base point. Measurement is made x ms after the R peak. This method has the advantage that the R wave is easier to identify than the J point and hence R + x method has been used in the project work. The ST segment is assumed to begin at 60 ms after the R peak in normal sinus rhythm. In the case of tachycardia (RR - interval < 600 ms) the beginning of the ST segment is taken at 40 ms after R peak. The ST segment for each heart beat has a predefined length of 160 ms. This means that the end point is 220 ms after R-peak in normal case and 200 ms otherwise.

ST - segment levels can be expressed as absolute changes relative to the isoelectric line [26] or as changes normalized to the R-wave peak. The work uses absolute voltage measured from the isoelectric voltage measured of the ST segment with respect to the isoelectric voltage levels. Due to the R wave amplitude changes ST levels may change. Hence in order to avoid false indications of ST changes the measurement is made with respect to the isoelectric voltage level. For R + x method the number of samples used for different sampling frequency are,

i) 32 points for sampling frequency of 200 Hz.
ii) 40 points for sampling frequency of 250 Hz.
iii) 58 points for sampling frequency of 360 Hz.

As per above the ST segment is decided and then ANN can be used to detect elevation or depression. The details of the ANN approach are discussed in Chapter 6. Fig. 6.3 and 6.1q indicate how isoelectric line is decided and plotted on one segment of s20011-2 file and its corresponding ST segment decided by R + x method.

5.9 T WAVE DETECTION.

In the ECG abnormality detection the width of T wave alone is not important. When calculation of corrected QT interval (QTc) factor is done, then width of T wave is included in measurement. (QTc = QT interval/ (preceding RR interval )^{1/2}). When efforts
were made to measure the width of T wave alone, it was found out that the alignment of ECG segment as well as the normalization of ECG segment both are equally important. The ECG segments are thus aligned to have the R point at sample point equal to 100, and the maximum amplitude is aligned to a value as 1. Abnormality in T waves is Tall T waves or peaked T waves or small, flattened T waves or inverted T waves. Since all these were related to magnitude of the T waves, measurement was easily done in time domain only.

T wave width detection by ART2-theory was also tried. It was found out that the factor ‘lim’ in Equation 5.26 was decided experimentally and the factor for T waves was found as 0.431 for the T wave section from QRS side to peak of T waves, and it was found as 0.914 for the remaining half of T wave. Here the measurement was done by right angled triangle method and correction factor ‘lim’ was decided. However it is mentioned in most of the references that magnitude and nature of T wave indicates the abnormality. Hence the rigorous testing for T wave using ART2 method was not done and the obtained values of ‘lim’ also were not checked practically for the available complete database. As is noted earlier the factor ‘lim’ has a value of 1.2 for Q point detection and 1.75 for S point recognition. Suzuki [71] found out the ‘lim’ values by computer simulation and he checked results for 1500 cardiac cycles of 5 different subjects. From this point of view the values of ‘lim’ for T waves are not checked rigorously in this work.

5.10 PR INTERVAL MEASUREMENT.

Normal value of PR interval is 0.12 to 0.20 sec [56]. For short PR interval than 0.12 sec, consider Wolff-Parkinson-White syndrome or Lown-Ganong-Levine syndrome, other causes can be Duchenne muscular dystrophy, type II glycogen storage disease, etc. For long PR interval the diagnosis is first degree heart block. The heart block indicates problem in conduction system of the heart. The topic of this thesis is coronary artery disease so the conduction system problems are not covered in this. Thus short or long PR interval is not relevant from the point of view of the scope of this thesis work. However the PR interval in time domain is found out by suitable program. It was not thought necessary to obtain PR interval measurement for 300 ECG segments. Thus in few cases it
was measured to check accuracy of software prepared. 25 PR interval values are checked. The accuracy was found satisfactory.

5.11 CONCLUDING REMARKS.

In this Chapter the ECG characteristic points and the various intervals of ECG are found out by different methods. The QRS detection by Haar wavelet transform and by ART2 method provided very good results. The width of 300 QRS complexes were found out from various 16 ECG files (refer Chapter 1), and were stored for the testing of neural network based system. The ST point detection was done by R+x method and the 300 ST segments, each originally of 40 points (for 250 Hz sampling frequency) converted to 20 points and were stored. Suitable changes in the number of sample points collected for ST segment were done for changed sampling frequency. Thus 300 ST segment templates were stored. The QTc factors similarly were calculated for all 300 ECG segments and stored. The method for PR interval measurement was decided from software point of view. Only few PR intervals were generated for each file of ECG. Thus the work discussed in this Chapter generates the data needed for various neural network systems whose design and performance is discussed in Chapter 6.