# HRV AND ECG SIGNAL ANALYSIS OF SMOKERS AND NON-SMOKERS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

**Bachelor of Technology**

**in**

**Biomedical Engineering**

Submitted

By

**RUCHIKA GOEL**

Under the Guidance of

**Dr. Kunal Pal**

**Department of Biotechnology & Medical Engineering**

**National Institute of Technology**

**Rourkela 769008**

**Department of Biotechnology & Medical Engineering**

**National Institute of Technology, Rourkela- 769008**

**Dated: May 14, 2011**

## CERTIFICATE

This is to certify that the thesis entitled "**HRV AND ECG SIGNAL ANALYSIS OF SMOKERS AND NON-SMOKERS**" submitted by **MS. RUCHIKA GOEL** in partial fulfilment of the requirements for the degree of **Bachelor of Technology in Biomedical Engineering** embodies the bonafide work done by her in the final semester of her degree under the supervision of the undersigned. The thesis or any part of it has not been submitted earlier to any other University / Institute for the award of any Degree or Diploma.

**(Dr. KUNAL PAL)**

**Assistant Professor**

**Indian Institute of Technology Bhubaneswar**
**Bhubaneswar - 751013. Orissa**

Date: 13th May 2011

## Certificate

This is to certify that the thesis entitled "HRV AND ECG SIGNAL ANALYSIS OF SMOKERS AND NON-SMOKERS" submitted by Ms. Ruchika Goel for the partial fulfilment of the requirements for the degree of B.Tech embodies the bonafide work done by her in the final semester of her degree under the co-supervision of the undersigned. The thesis or any part of it has not been submitted earlier anywhere for any degree or diploma or any other qualification.

Subhr

(Dr. S.R.Samantaray)
School of Electrical Sciences
Indian Institute of technology, Bhubaneswar

# ACKNOWLEDGEMENT

I would like to take this opportunity to extend my hearty gratitude to my thesis supervisor, Dr. Kunal Pal, Assistant Professor, Department of Biotechnology and Medical Engineering; National Institute of Technology, Rourkela, for whose constant guidance and encouragement, the successful completion of my B.Tech thesis was possible.

I would also like to thank my co-supervisor, Dr. S.R.Samantaray, Associate Professor, School of Electrical Sciences, Indian Institute of Technology, Bhubaneswar, for his extended support and direction to the course of the project.

Gratitude would be incomplete without the mention of Prof. D. N. Tibarewala, School of Bioscience & Engineering, Jadavpur University, for his generous contribution of an in-house developed ECG data acquisition system for collaborative research.

I am really thankful to National Institute of Technology, Rourkela, for permitting me to utilise the facilities in its laboratories for the smooth execution of my experiment.

I extend my warm gratitude to my friend, Mr. Anup Puri, for his constant motivation and support throughout the course of my B.Tech research.

Finally, I would like to thank my seniors, juniors and my fellow students who enthusiastically supported me by giving their ECG readings as well as their personal information without hesitation.

**RUCHIKA GOEL**

**Department of Biotechnology and Medical Engineering**

## TABLE OF CONTENTS

# ABSTRACT

The current study deals with the study of Heart Rate Variability and wavelet-based ECG signal analysis of thirty-two volunteers, who were divided into two groups of smokers and non-smokers. Although the preliminary results of frequency domain analysis of HRV showed some dominance towards the sympathetic nervous system activity in smokers, they were not found to be statistically significant. Hence the bias of results towards the increase of sympathetic activity might be attributed to the masking affect of some other factors, apart from smoking, which were not included in our experiment. The wavelet decomposition of the ECG signal was done using the Daubechies (Db 6) wavelet family. No significant difference was observed between the smokers and non-smokers which apparently suggested that HRV does not affect the conduction pathway of heart.

# 1. INTRODUCTION

In recent days, the smoking of tobacco-based products has increased in the college and university going students. In general, it has been found that students take up smoking during the age of 11-17 yrs, which have been attributed to their misinterpretation of smoking habits as a sign of higher social status. It has become one of the major causes of preventable death in the world population. The scenario is worst in the developing countries, where smoking has been inculcated into the social lifestyle of the people. The common side effects of smoking include manifestations of cancer, cardio-pulmonary diseases (of varying intensities) and melanocytic pigmentation in the smokers. Of late, various cutaneous manifestations have also been reported by various researchers. Most of the side effects of smoking have been attributed to the inhalation of the soot and nicotine present in the cigarette smoke.

Nicotine, an alkaloid, is one of the potent neurophysiologic modulator. The inhalation of cigarette smoke results in the rapid absorption of the same which in turn reaches the brain within 10-16 sec. It promotes the release of dopamine acetylcholine receptors. The release of the dopamine induces a short-term behavioral change in people leading to the feeling of pleasure and excitement apart from the sense of alertness and arousal. The nicotine-dependence of the smokers makes it very difficult for them to quit smoking. This may be attributed to the reduction of nicotine concentration within monoamine oxidase inhibitor enzymes, responsible for the metabolism of catecholamine. Apart from the dopamine levels, nicotine has also been found to modulate the physiological concentrations of other neurotransmitters, e.g. glutamate, GABA, acetylcholine, dopamine, norepinephrine, and serotonin. The effect of nicotine on the neurophysiological system results in the increased activity of the sympathetic system.

Taking note from the above, it seems quite feasible to study the activity of the autonomic nervous system from the frequency domain features of the ECG signal. This is due to the fact that the autonomic nervous system plays a direct role in the stimulation of the SA node, which is responsible for beginning the electrical activity of the heart. In a normal and healthy person, there is a balance between the sympathetic and parasympathetic system known as the sympathovagal balance. The interval between the R-R peaks plays an important role in understanding the activity of the autonomic nervous system. In general, a parasympathetic system increases the R-R intervals while the sympathetic system reduces the same. The measure of the change in the R-R intervals is expressed as Heart Rate Variability (HRV).

The evolution of the joint time-frequency analysis (JTFA) has opened up a completely new arena of bio-signal processing. Since the analysis of ECG signal in JTFA domain has not seen much progress, it becomes quite justifiable to analyze the signal in this domain.

The current study deals with the ECG analysis of thirty-two volunteers who were categorized under two groups viz. smokers and non-smokers. The frequency domain HRV analysis was carried out for both the groups. Wavelet analysis, a JTFA tool, of the ECG signals was also carried out to understand if there was any change in the electrophysiology of the heart.

## 2. LITERATURE REVIEW

Normally clinicians compute the overall heart rate by counting the number of QRS complexes over an interval of one minute. However, during HRV analysis the time gap between identical events in consecutive cardiac cycles are considered – and how its value changes with the progression of time is noted. At the same time how the instantaneous heart rate changes with time is also taken into account. 'Heart Rate Variability' has become the conventionally accepted term to describe variations of both instantaneous HR and R-R intervals.[1] In order to describe oscillation in consecutive cardiac cycles, other terms are cycle length variability, heart period variability, R-R variability and R-R interval tachogram, and they more appropriately emphasize the fact that it is the interval between consecutive beats that is being analyzed rather than the heart rate per second. Even in physiological state the heart rate of an individual does not remain absolutely constant – it varies even on the beat-to-beat basis owing to the dynamic control of ANS on the SA node.[2]

3

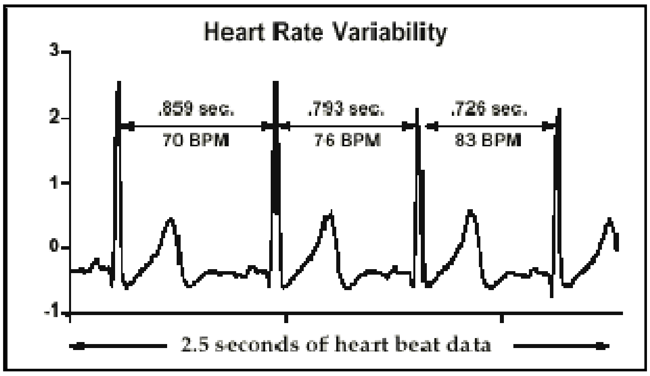



**Figure 1: Heart Rate Variability is a measure of the changes in RR interval of the ECG signal**

Although other methods used to detect beats including EMG, blood pressure etc. has been developed, ECG is still considered to be a better option due to its clear waveform which enables easy elimination of heartbeats not originating in the SA node. Normal sinus rhythm is continuously generated by the SA node which acts like a pacemaker tissue and is located in the right atrium of the heart. Cells in the SA node naturally discharge beats at about 60-100 beats/min, i.e. in the absence of the influence of extrinsic neural and hormonal control.

[4]

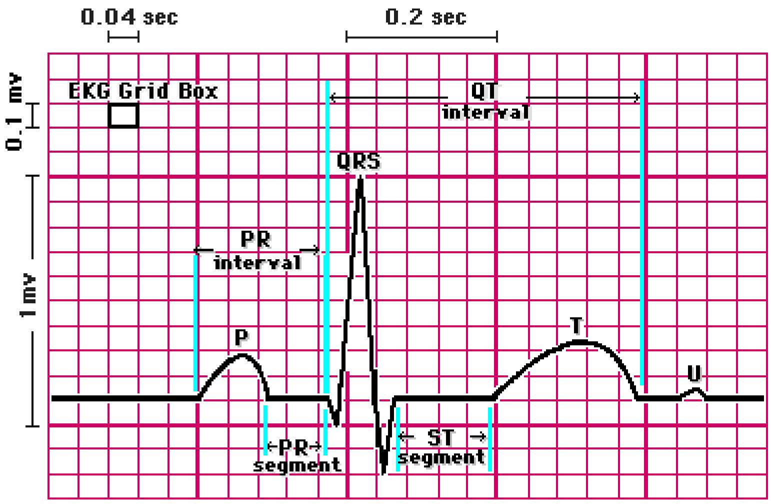



Figure 2: ECG Wave

## 2.1 Interdependence of ANS and HRV:

The nervous system is divided into two parts:

- Somatic nervous system which is in charge of voluntary control of organs which include mainly muscles.
- Autonomic Nervous System (ANS) also known as the visceral/automatic system regulates individual organ function and homeostasis, and for the most part is subject to involuntary control.

Researchers have examined the effect of emotions on the autonomic nervous system by the analysis of heart rate variability, which serves as a dynamic comparison between autonomic function and balance.[5] We know that in a healthy volunteer, the parasympathetic nervous system fibers and the sympathetic nervous system fibers richly innervate the sinoatrial node. Parasympathetic innervations of the heart are mediated by the vagus nerve which causes a decrease in the SA node thereby decreasing the heat rate whereas stimulation by the sympathetic fibers causes an increase in the heart rate.

Thus the variability in the heart rate is due to the action of balance and synergy between the two branches of the Autonomic system, which is mainly enforced through neural, mechanical, humoral and other physiological mechanisms. It maintains cardiovascular parameters in their most favorable ranges and permits suitable reactions to change in external or internal stimuli.[6]

This balance between the effect of the sympathetic nervous system and the parasympathetic nervous systems is known as the sympathovagal balance and is believed to be echoed in the beat-to-beat changes of the cardiac cycle.[7] The heart rate is defined by the reciprocal of the RR interval with units of beats/min.

## 2.2 Relation between Cardiac health and ANS:

HRV is reflective of the general state of well-being of the organism. It is predominantly dependent on the extrinsic regulation of the Heart rate. HRV is thought to reflect the heart's ability to adapt to changing circumstances by detecting and quickly responding to unpredictable stimuli.[8] Recent experimental confirmation for a connection between a tendency for fatal arrhythmias and indications of either increased sympathetic or reduced parasympathetic activity has motivated to the development of quantitative markers to adjudge the autonomic activity. HRV represents one of these most potent markers. It is a strong and independent forecaster of death following an acute myocardial infarction.[9]

It has been proved in research studies that during the period of a mental or an emotional anxiety, an increase in the sympathetic activity and a simultaneous decrease in the parasympathetic activity were observed. This results in amplified strain on the heart, immune as well as other important hormonal systems. The increase of sympathetic activity is related to a reduced ventricular fibrillation threshold and thus an augmented threat of fibrillation, in contrast to an increase in parasympathetic activity, which protects the heart.[10]

## 2.3 Ways of measurement of HRV:

### 2.3.1 Time domain methods:

- *Statistical measures:*
  1. SDNN- It is the standard deviation of NN intervals which are often calculated over a period of 24 hrs.
  2. SDANN- It is standard deviation of the average NN intervals which are measured over short periods of 5 minutes. Thus SDANN can be seen as a determination of changes in the heart rate for cycles longer than 5 min.
  3. RMSSD- It is the square root of the mean squared difference of successive NN intervals.
  4. SDSD- It is the standard deviation of the differences between consecutive NN intervals.
  5. NN50- It is the number of pairs of consecutive NNs which have a difference of more than 50 ms in the entire span of recording.
  6. pNN50 – It is the count of NN50 which is divided by the total number of all NNs.

- *Geometric measures:*
  1. HRV triangular index- It is the total number of all NN intervals which is divided by the height of the histogram of all NN intervals and measured on a discrete scale.
  2. Differential index- It is the variation between the widths of the histogram of differences between neighboring NN intervals which is measured only at selected heights.[11]
  3. Logarithmic index- Coefficient 'phy' of the negative exponential curve $k * e^{-phy\, t}$ which is the best estimate of the histogram of differences between the adjacent NN intervals.

### 2.3.2 Frequency Domain Methods:

The transformation of HRV data mathematically is used to distinguish and measure the sympathetic and parasympathetic activity along with the activity of the autonomic nervous system, thus decomposing the HRV signal into its constituent frequency components and thus computing the relative power of the components.[12]

The three main frequency bands of interest are referred to as:

- Very Low frequency (VLF)- 0.003-0.04 Hz.
- Low frequency band (LF) - 0.04 to 0.15 Hz.
- High frequency band (HF) -0.15 to 0.4 Hz.

The magnitude of the HF component provided an index of vagal activity and the magnitude of the LF component provided an index of sympathetic activity with vagal modulation. The LF/HF ratio was used as a marker of instant sympathovagal balance.[13]

The distribution of the low frequency and high frequency power as well as their central frequencies may not be fixed but may vary with changes in autonomic modulations of the heart. Measurements of VLF, LF and HF power components is usually made in absolute values of power(ms^2), but LF and HF may also be measured in normalized units (n.u.), which represents the relative value of each power component in fraction to the total power excluding the VLF components.

The representation of LF and HF in n.u. emphasizes the controlled and balanced behavior of the two branches of the ANS. The advantage of the n.u. units is that normalization tends to minimize the effect on the values of LF and HF components of the changes in total power.[14]

### 2.4 Why short-term recordings are preferred?

Problem of being stationary is frequently discussed with long-term recordings. If mechanisms responsible for heart period modulations of a certain frequency remain unchanged during the whole period of recording, the corresponding frequency components of HRV may be used as a measure of these modulations. If the modulations are not stable, interpretation of the results of frequency analysis is less well defined.

In particular, physiological mechanisms of heart period modulations responsible for LF and HF power components cannot be considered stationary during the 24 hr period. Thus, spectral analysis performed in the entire 24 hr period as well as spectral results obtained from shorter segments (e.g. 5 min) averaged over the entire 24 hour period provide averages of the modulations attributable to the LF and HF components.

Such averages obscure detailed information about autonomic modulation of RR intervals available in shorter recordings.

## 2.5 Requirements to obtain a reliable spectral estimation:

1. Sampling Rate: Optimal range should be in the range of 250-500 Mz or higher.
2. Baseline and trend removal (if used) may affect the lower components in the spectrum, hence should be taken care of.
3. A stable and noise-independent fiducial point should be chosen.
4. Ectopic beats, arrhythmic events, missing data and noise effects may later the estimation of the PSD of HRV.[15]

For certain experimental setups, the respective environmental variables have to be controlled besides the fact that the nature of the environment should also always be expressed. Besides, individual volunteers should face similar recording environment.

## 2.6 Applications of HRV in the detection of:

- ✓ Myocardial infarction
- ✓ Diabetic neuropathy
- ✓ Cardiac transplantation
- ✓ Myocardial dysfunction
- ✓ Tetraplegia

## 2.7 Influence of cigarette smoking on HRV:

It has been observed that cigarette smoking increased sympathetic activity or decreased vagal cardiac activity and so has been recognized as a major mechanism for increased risk of coronary artery disease in smokers. Many recent studies have reported increased susceptibility to sudden coronary death and increased subsequent mortality after myocardial infarction in patients with decreased vagal cardiac control assessed by heart rate variability.

The decrease in vagal cardiac activities may be present in smokers and may be a reason for the association between smoking and cardiac death.[16]

Generally is seen that smoking causes a significant:

1. increase of the LF/HF ratio
2. Increase of the normalized values of the LF component
3. Decrease in absolute values of total power, LF and HF.
4. Significant decrease in the time-domain indices.

## 2.8 What Are Wavelets?

A wavelet is a waveform of limited duration that has an average value of zero. Unlike sinusoids that theoretically extend from minus to plus infinity, wavelets have a beginning and an end. Sinusoids are smooth and predictable and are good at describing constant-frequency (stationary) signals. Wavelets are irregular, of limited duration, and often non-symmetrical. They are better at describing anomalies, pulses, and other events that start and stop within the signal.

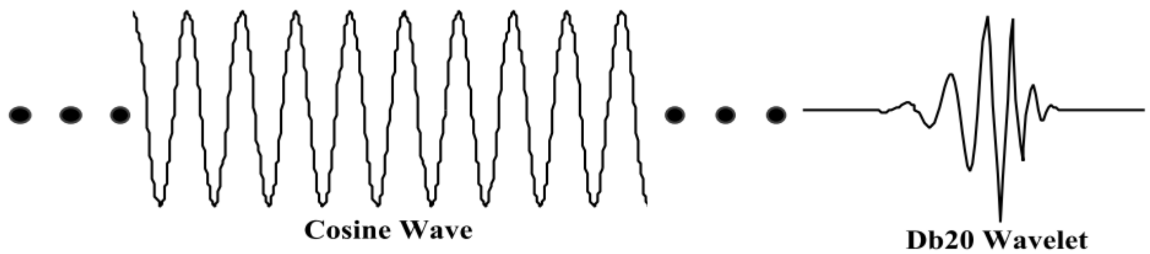



**Figure 3: A portion of an infinitely long sinusoid (a cosine wave is shown here) and a finite length wavelet. Notice the sinusoid has an easily discernible frequency while the wavelet has a pseudo frequency in that the frequency varies slightly over the length of the wavelet.[17]**

There exists different types of wavelets which are matched against the shape of the desired signal whose wavelet transform is to be done. If the wavelet family is close to the desired physical characteristics of the signal, the particular wavelet family is selected for use.

17

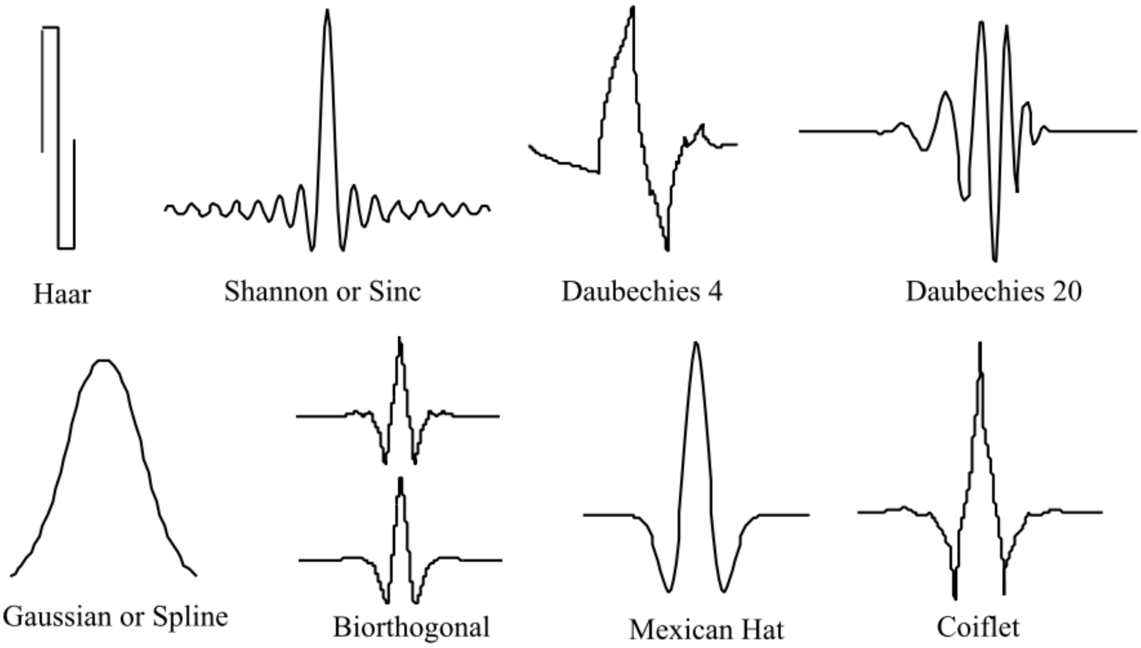



**Figure 4: Types of wavelet families**

### 2.8.1 Comparison between Fast Fourier Transform and Continuous Wavelet Transform

Time domain analysis was used in earlier times to study the features of ECG signals, but due to its insufficiency to describe all the details of ECG features, attention was diverted to the frequency representation of a signal. To accomplish this, FFT (Fast Fourier Transform) technique was developed which allowed viewing signals in the frequency domain.

FFT decomposes the original signal into the constituent sinusoids of different frequencies called as spectrum thus allowing manipulation of the transformed data from time to frequency domain, and then by the method of inverse Fourier Transform helps to undergo custom filtering such as elimination of constant frequency noise. In spite of this advantage, FFT doesn't give information about the time that a particular frequency occurred.

**18**

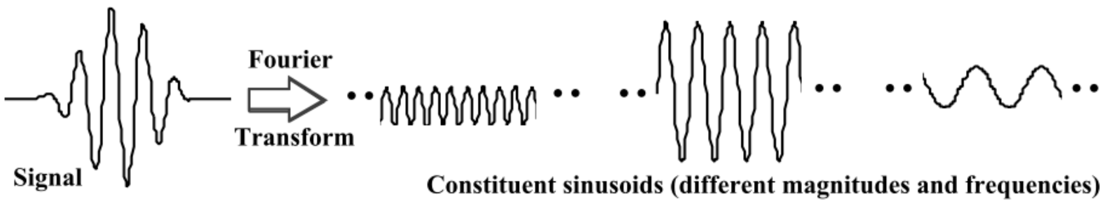



**Figure 5: Signal transformed into a number of sinusoids of various sizes and frequencies.**

Figure 16 demonstrates the stretching and shifting process for the continuous wavelet transform. Waveform (B) shows a Db 20 wavelet family of the length of 1/8 second starting at the beginning i.e. t = 0 and ending before 1/4 second. The zero values are extended to the whole 1 second. The point-by-point comparison with the pulse signal (A) would be very poor obtaining a very small correlation value.

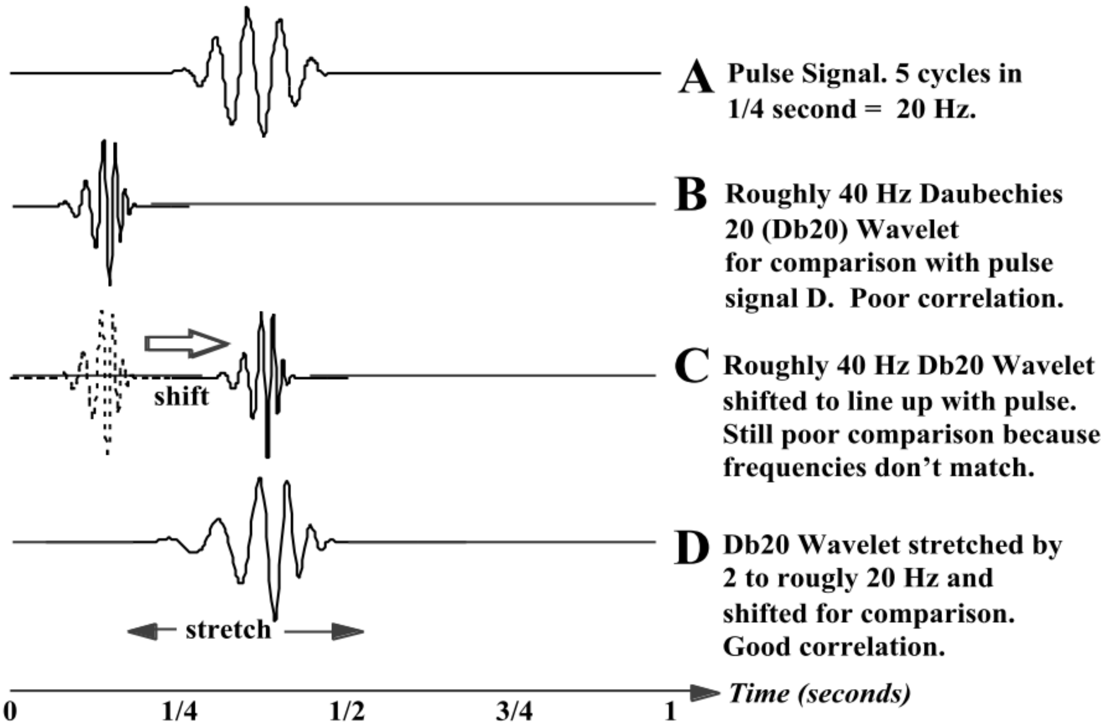



Figure 6

An actual wavelet transform compares many stretched and shifted wavelets (analysis wavelets) to the original pulse. The unstretched wavelet is often referred to as the mother wavelet. The Db20 wavelet filter we are using here starts out as 20 points long (hence the name) but can be stretched to many more points. A counterpart low pass filter used in the upcoming discrete wavelet transform is often called a father wavelet.

Waveform (D) demonstrates the Db20 wavelet stretched to the position where the frequency is almost the same as the pulse (A) and moved to the right until the peaks and valleys match up almost perfectly. At these amounts of translating and stretching, we should ideally get a good comparison and a great correlation value. However, if we further shift to the right even at this same stretching, it will progressively yield more poor correlations. Further stretching doesn't help at all because even when lined up, the pulse and the over-stretched wavelet won't be the same frequency.

In the CWT we have a respective correlation value for every single shift of the stretched wavelet.

The generalized equation for CWT is a shortcut that shows that the correlation coefficients depend on both the stretching and the shifting of the wavelet, $\psi$, to match the signal ($x_n$ here) as we have just seen. The equation shows that when the dilated and translated wavelet matches the signal the summation will produce a large correlation value.

$$\boxed{C\,(\text{stretching, shifting}) = \sum X_n\, \psi\,(\text{stretching, shifting})}$$

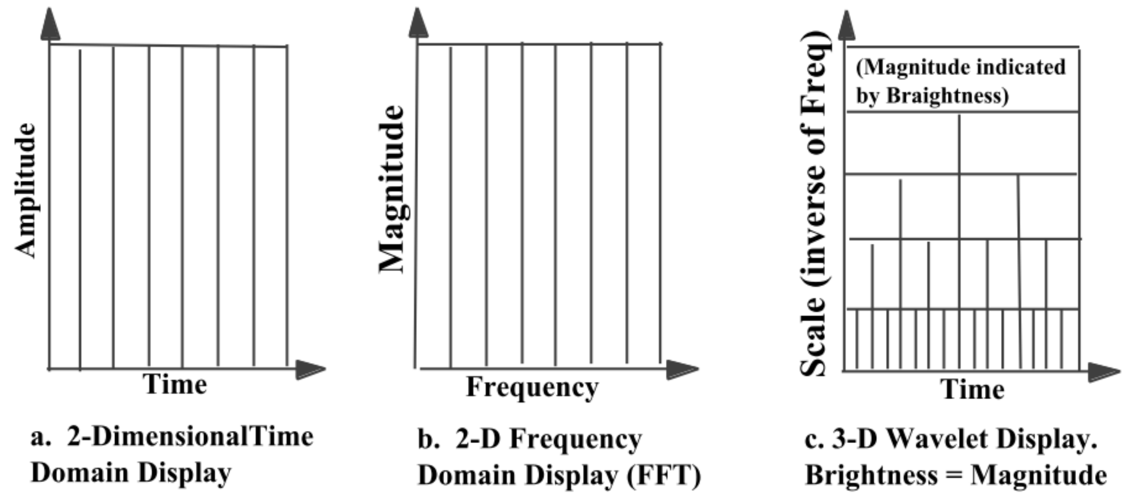



Figure 7

Figure 7 shows a wavelet transform display (c) and how it compares to an ordinary time-domain (a) and frequency-domain (b) display. Note that the wavelet display (c) is drawn inverted ("flipped" vertically). In other words the high frequencies are on the bottom and the lower frequencies are on top. The y axis on most wavelet displays shows increasing scale (stretching of the wavelet) rather than increasing frequency.

### 2.8.2 Advantages of Wavelet Analysis over Traditional Frequency Domain Analysis:

1. By stretching and shifting a wavelet, we can match it to the hidden event and thus discover its frequency and location in time. In addition, a

particular wavelet shape may match the event unusually well, also telling us about the shape of the event. For example, the Haar wavelet would match an abrupt discontinuity while the Db20 would match a chirp signal.

2. An important advantage of a wavelet transform is that, unlike an FFT, we can threshold the wavelet coefficients for only part of the time. Suppose we had a binary signal that had a great deal of noise added which changed frequency as time progressed (e. g. chirp noise). Using a 7-level DWT the noise would appear at different times in the different frequency sub-bands which could be threshold at the appropriate times.[19]

### 2.8.3 <u>Wavelet Decomposition</u>

The left half of the Discrete Wavelet Transform is called the decomposition or analysis portion and comprises the forward transform. The right half is called the reconstruction or synthesis portion and comprises the inverse transform.

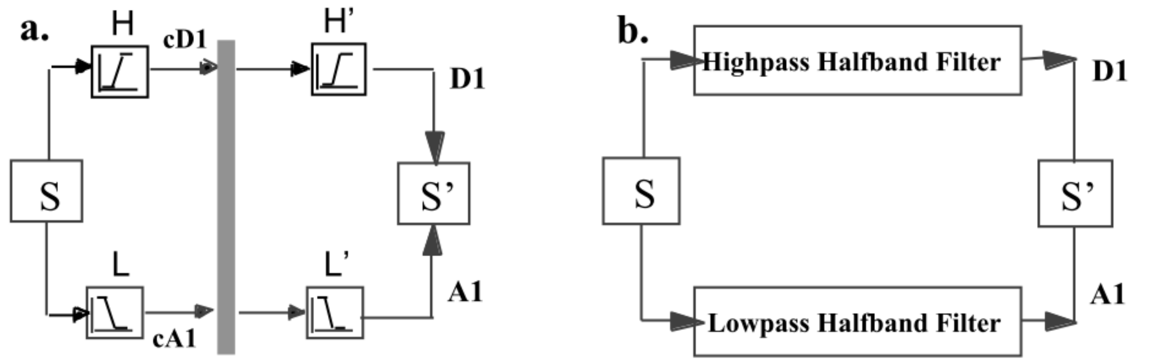



**Figure 8: Single level decomposition wavelet**

"Decomposition" in wavelet terminology means splitting the signal into 2 parts using a highpass and a lowpass filter. Whereas "Reconstruction" means using filters to combine the parts.

On the upper path of Fig. 18 (a), the signal, S, is first filtered by H (highpass decomposition filter) to produce the coefficients cD1. At this point we can do further decomposition (analysis) for compression or denoising, but for now we proceed directly to reconstruct (synthesize) the signal. cD1 is next filtered by H' (highpass reconstruction filter) to produce the Details (D1). The same signal is also filtered on the lower path by the lowpass decomposition filter L to produce the coefficients cA1 and then by the lowpass reconstruction filter L' to produce the Approximation (A1). "Approximation" in Wavelets is the smoothed signal after all the lowpass filtering. "Details" are the residual noise after all the highpass filtering. cA1 designates the approximation coefficients and cD1 designates the details coefficients. These coefficients can be broken down (decomposed) into further coefficients in higher level systems.

The details have a special property called orthogonality, which means that they are completely independent of each other and therefore can be added together in any sequence.

When the Details and Approximations are added together they reconstruct S (cD1 + cA1 = S) which is identical to the original signal. For a simple denoising, we could simply reject these high frequencies in D1 and A1 by itself would be a rudimentary "denoised" signal.[20]

### 2.8.4 Choice of Prototype Wavelet

The huge number of known wavelet families provides a good scope to search for a wavelet which would efficiently represent a signal of interest. The choice of the wavelet depends on its application.[21] The Haar wavelet has the advantage that it can be computed easily and and is also easy to understand. The Daubechies algorithm involves complex concepts along with being a little difficult to compute. But, the Daubechies wavelet selects detail that the Haar wavelet may perhaps ignore. It is of primary significance that the selection of the wavelet family should closely match the signal of interest.[22]

Daubechies wavelets have structural similarity with QRS complex and their energy spectrums are concentrated around low frequencies. Thus it is expected that some detail coefficients from multiresolution decomposition will show better resemblance with QRS complex of the ECG wave in time scale domain.[23][24]

# 3. MATERIALS AND METHODS

## 3.1 VOLUNTEERS

In this study, ECG of 32 volunteers within the age group of 18-26 yrs was recorded. The volunteers were divided into two categories of smokers and non-smokers comprising sixteen volunteers in each. Before recording the ECG signal, all volunteers were asked to read and fill up a questionnaire divulging their background information, medical history, smoking habits etc. If the volunteers agreed to participate in the study, they were asked to sign the consent form at the end of the questionnaire (Annexure-I).

## 3.2 BASIC INFORMATION

The sex and mean age± standard deviation of the volunteers have been shown in Table 1. Individuals with smoking history of few months to 8 years, who smoked 1-11 cigarettes per day, comprised the smoking group. No participating volunteers had any history of diabetes mellitus, cardiac disease and/or hypertension.

**Table 1: Classification of volunteers**

| CATEGORY | MALE | | FEMALE | | TOTAL |
|---|---|---|---|---|---|
| | Number | Mean age±SD | Number | Mean age±SD | Number |
| **Smokers** (21.4 ±1.9) | 16 | 21.4 ±1.9 | 0 | 0 | 16 |
| **Non-smokers**(22.8±1.6) | 11 | 22.5±1.7 | 5 | 23.4±1.5 | 16 |
| **Total** (22.1±1.9) | 27 | 21.9±1.9 | 5 | 23.4±1.5 | 32 |

## 3.3 DESIGN OF EXPERIMENTAL STUDY

The volunteers were checked for their height and weight. They were seated comfortably in chairs to minimize motion artifacts. The volunteers put on shoes throughout the proceedings of the experiment in order to avoid any contact with ground. All metallic objects in contact with their body including watch, bracelet, cell phone etc. were removed and kept separately.

### 3.4 CUSTOM MADE ECG DEVICE

An in-house developed ECG data acquisition system was provided by Prof. D. N. Tibarewala, School of Bioscience & Engineering, Jadavpur University for collaborative research. ECG was recorded continuously for six minutes. The ECG system mainly composed of two parts:

3.4.1 ECG Electrodes: A bio-potential two electrode system is a transducer which provides an electrical contact with human body and converts bio-potential signals (having an ionic origin) into electronic signals to be processed by subsequent circuitry.

3.4.2 HRV-DAQ: HRV-DAQ is the interface between ECG signal and data acquisition system. This is a custom made device consisting of pre-amplifier, opto-coupler, post amplifier, USB based ADC, DC to DC converter and USB of computer.

### 3.5 HRV Analysis

HRV is due to the combined activity of the parasympathetic and sympathetic nervous systems, which in turn gives an indication of the autonomic nervous system. Frequency domain HRV analysis of the ECG signals was carried out using the NI Biomedical Startup Kit 3.0. The software package allows detrending the signal before analysis could be made. Various frequency parameters like LF power, HF power and LF/HF ratio was determined to estimate the balanced functioning of the autonomic nervous system.

### 3.6 Wavelet-based ECG signal processing

Wavelet, a JTFA technique, allows simultaneous analysis of a signal both in time and frequency domain. Of late, this technique has been extended for the analysis of the biosignals by various researchers. The wavelet, Daubechies 6 (db 6) was used for the multiresolution analysis of the ECG signals of smokers and non-smokers. From the preliminary studies, it was found that the signal decomposed at D7 and D8 levels provided some distinct variations amongst the groups. Hence various features like mean, mode, covariance, variance, standard deviation, log energy entropy, Shannon entropy and energy density were extracted for both the D7 and D8 levels. The distinct features were represented in feature space to analyze whether the same could be used for classifying the signals.

# 4. RESULTS AND DISCUSSIONS

## 4.1 FFT analysis of HRV

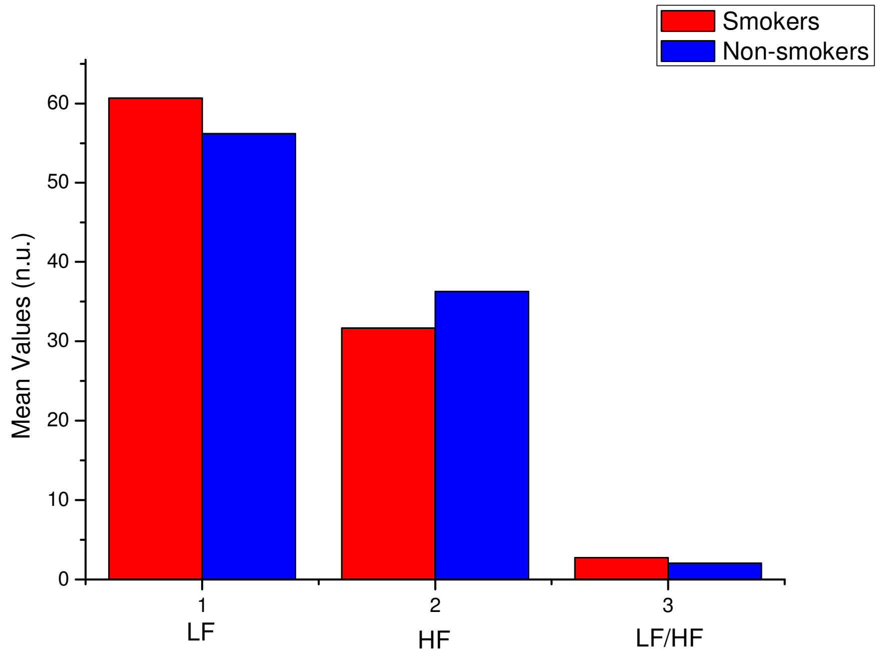



**Figure 9: Graph plotted to depict the normalized values of low frequency and high frequency components.**

| Category | Low Frequency power (n.u.) ± SD | High Frequency power (n.u.) ± SD | LF/HF ratio ± SD |
|---|---|---|---|
| **Smokers** | 60.649±16.853 | 31.671±14.753 | 2.7675±2.263 |
| **Non-smokers** | 56.153±16.926 | 36.264±15.584 | 2.0675±1.439 |

**Table 2: The above table shows the Heart rate variability spectral components measured in smokers vs. non-smokers.**

HRV frequency domain results have been tabulated in table 2. Two sample t test was done at p=0.05 for checking the statistical significance between the differences of the population means with the test difference. But no statistical significance was observed at 0.05 levels. The results suggest that there was an increase in the LF components with a subsequent reduction in the HF components in smokers compared to non-smokers. Also, a rise in the LF/HF ratio was observed in smokers compared to non-smokers. Though the preliminary results suggested shift of the frequency components towards sympathetic dominance, there were other factors (e.g. fooding habits, stress, physical activities etc.), which were not taken into account during HRV analysis, which might have a masking effect on the analysis. The magnitude of the HF component gives an indication about the vagal/ parasympathetic activity while the magnitude of the LF component gives an indication about the sympathetic activity with vagal modulation. LF/HF ratio is generally used as a marker of instant sympathovagal balance, which in turn provides an indication about the cardiac health.[25] A balanced sympathovagal activity indicates a healthy cardiac function. An increase in this ratio in smokers is indicative of an increased risk of coronary artery disease in smokers.

## 4.2 Wavelet Decomposition of ECG Wave

Figure 10 and 11 shows a typical multiresolution analysis of the 3 sec ECG signal obtained from non-smokers and smokers. A conscious effort was made to analyze the signals at higher levels of decomposition so as to ensure the presence of low frequency components of original signal. The figures indicate that at lower levels, high-frequency components are pronounced whereas at higher levels, low frequency components become more pronounced.[26] The results suggest that the coefficients at levels d7 and d8 contain distinct details of the signal. The wavelet coefficients from the levels d7 and d8 indicate their close resemblance with db6 wavelet scaling function. Hence, d7 and d8 coefficients were identified for the detection and extraction of features (e.g. mean, mode, covariance, variance, standard deviation, log energy entropy, Shannon entropy and energy density) for ECG signal classification amongst the smokers and non-smokers groups. The reconstructed signals from the wavelet coefficients of other levels showed noisy signal components and hence were not considered for further analysis. Graphs generated by a MATLAB program (Annexure-II) developed in our laboratory for the ECG features are shown below for the d7 coefficient (figure 12). Figure 12 shows the features extracted from the ECG signal. The results showed that there might be significant differences in the **mean** and **log energy entropy** features, for both d7 and d8 levels, and were used further for ECG signal classification. The other features graphs showed almost completely overlapping features with no further scope for classification.

The mean vs. log energy entropy of the ECG signals of the volunteers were represented in feature space (figure X), which basically represents both the ECG features at a time in a graph. It is used preliminarily to predict the possibilities of further classification using linear or non-linear classifiers. The results indicate that there was no significant difference amongst the ECG signals of the smokers and non-smokers.

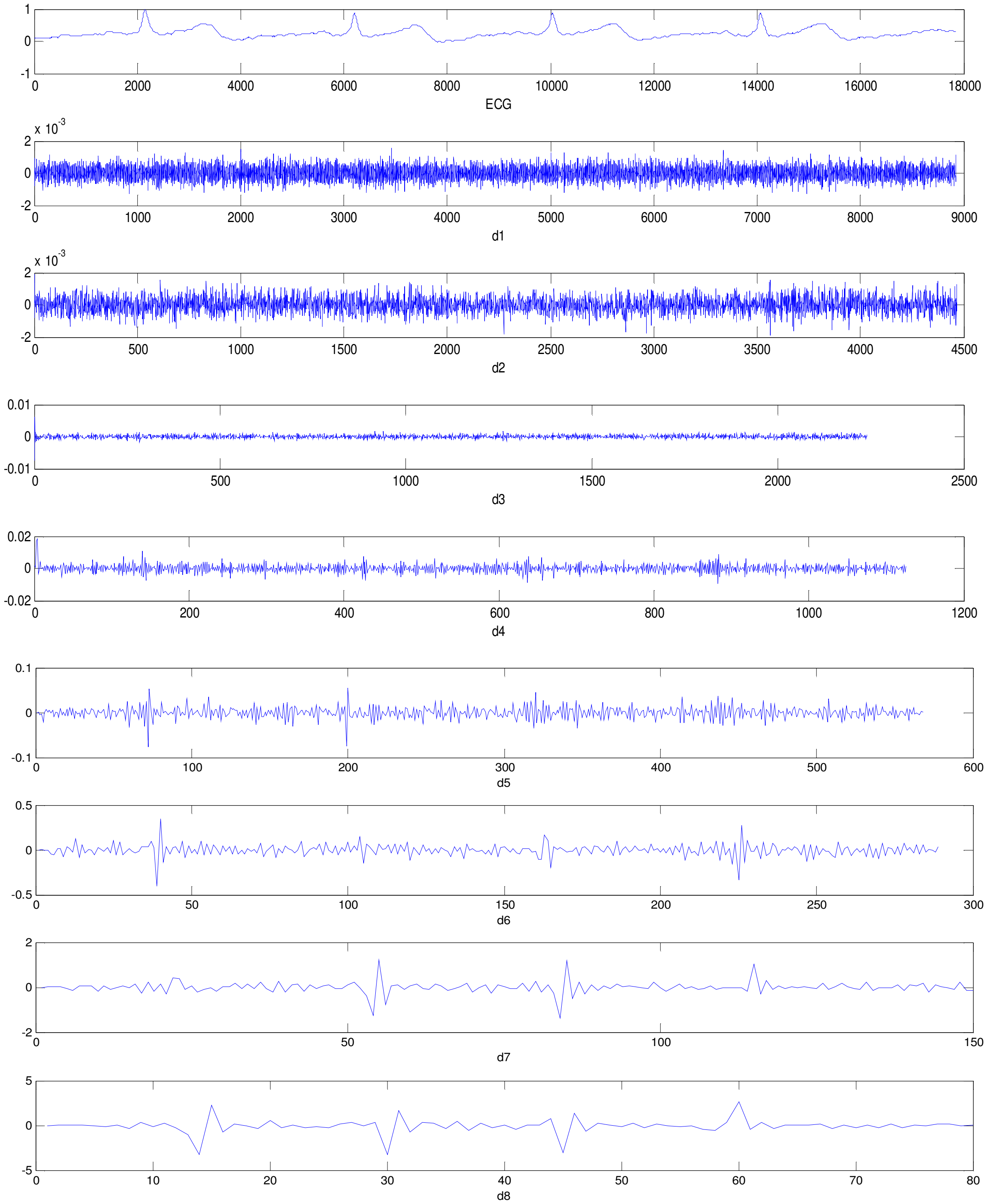



**Figure 10: Wavelet decomposition upto 8 detail coefficients in non-smokers.**

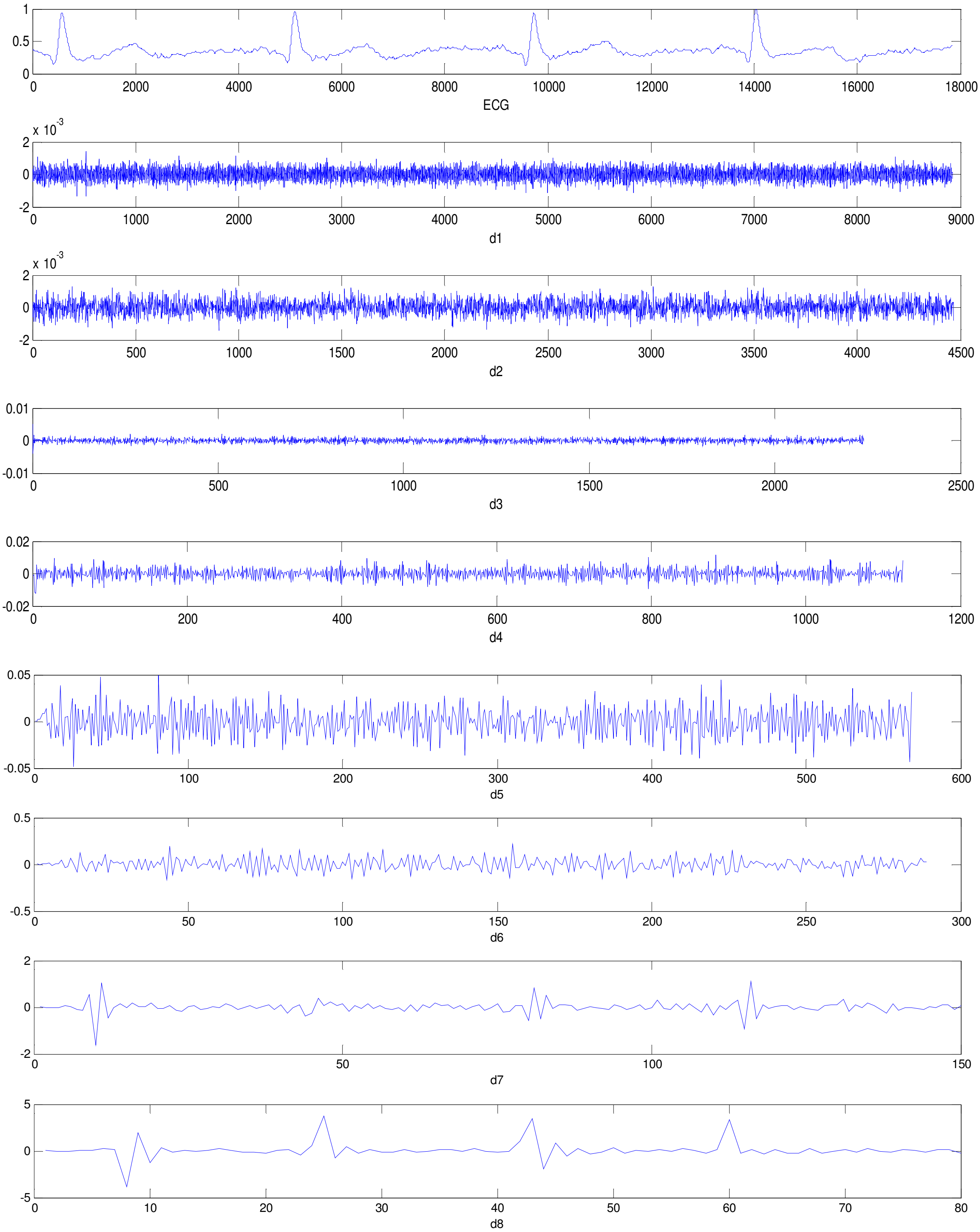
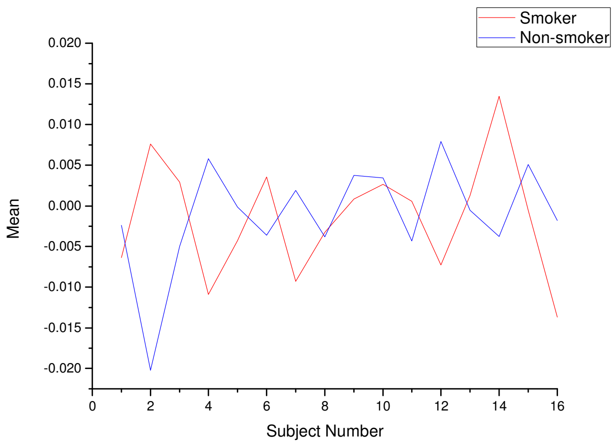
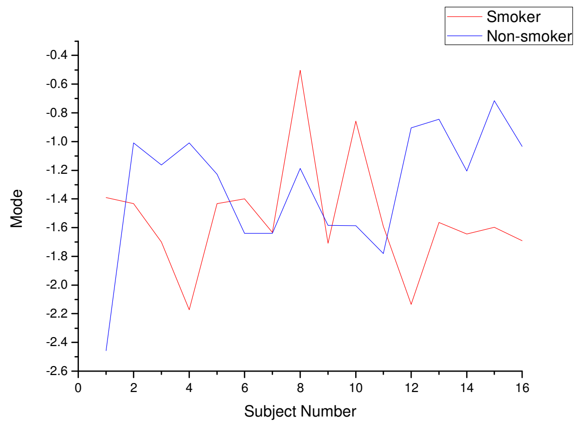



**Figure 11: Wavelet decomposition upto 8 detail coefficients in smokers.**

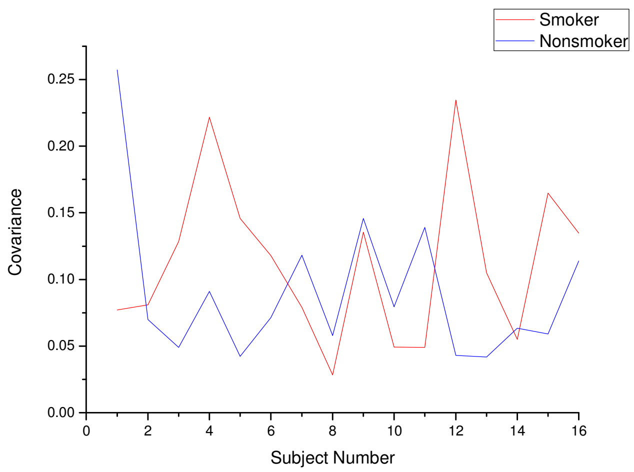
## 4.3 ECG Features Extracted:



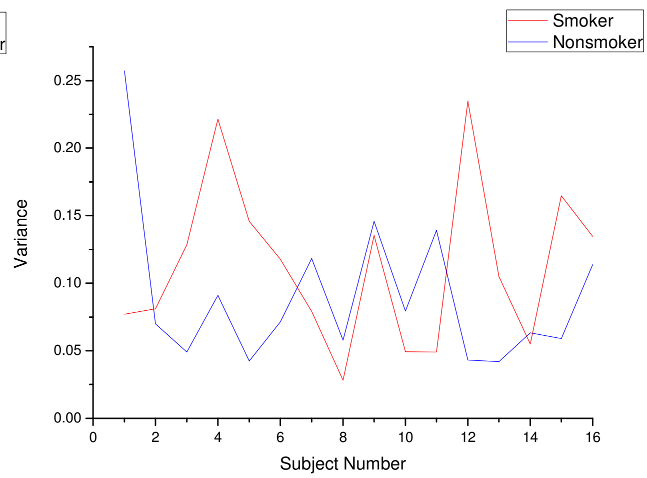

**(a)**



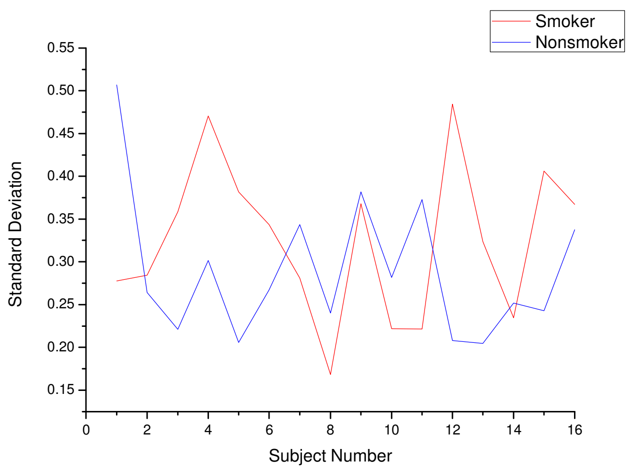

**(b)**



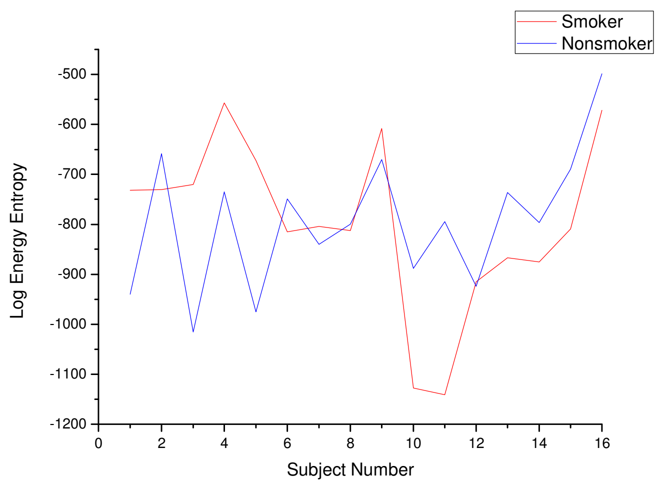

**(c)**



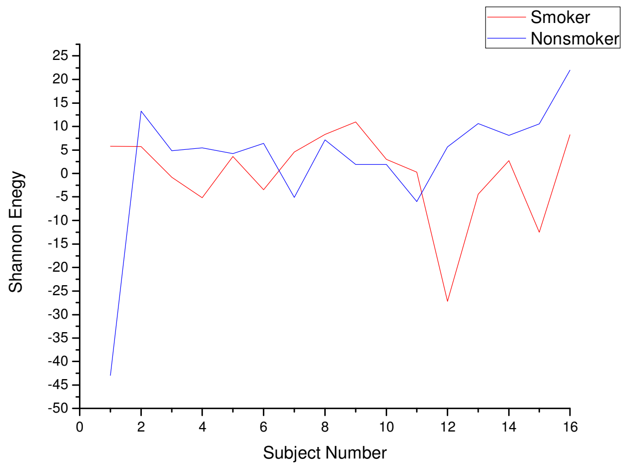

**(d)**



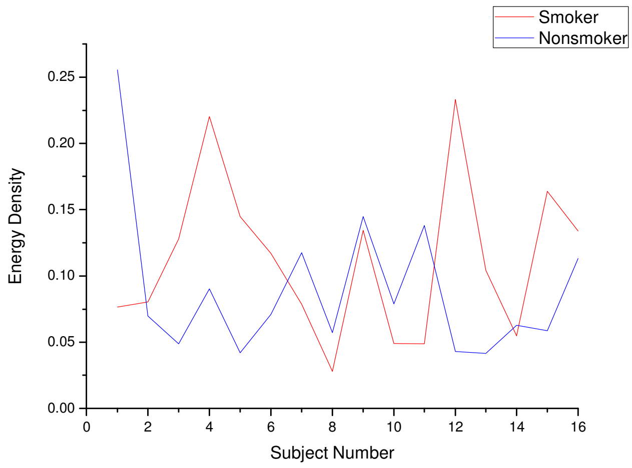

**(e)**




**(f)**




(g)




(h)

**Figure 12: Features extracted from the ECG signals. (a) Mean, (b) Mode, (c) Covariance, (d) Variance, (e) Standard Deviation, (f) Log energy entropy, (g) Shannon Entropy, (h) Energy density.**

## Comparison of Mean between d7 and d8 level:

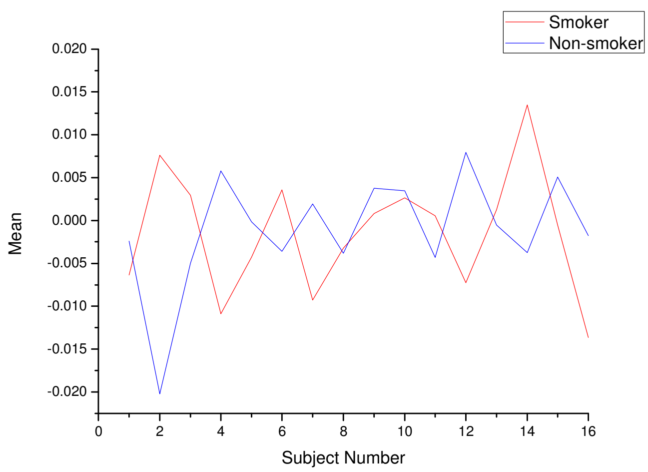



Figure 13(a): d7 level

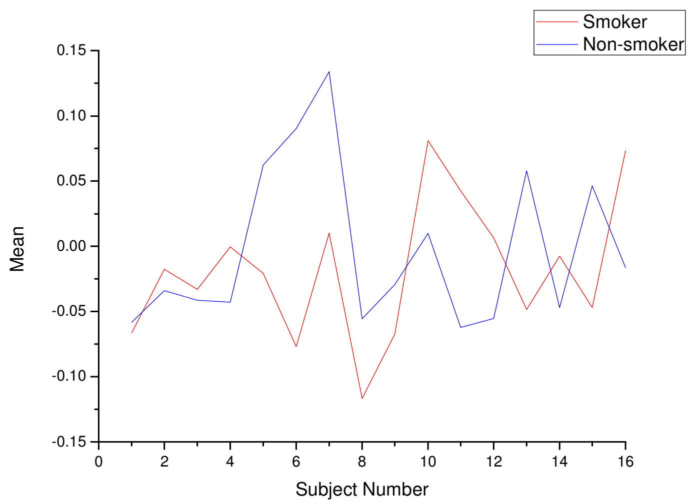



Figure 13(b): d8 level

## Comparison of Log Energy Entropy between d7 and d8 level

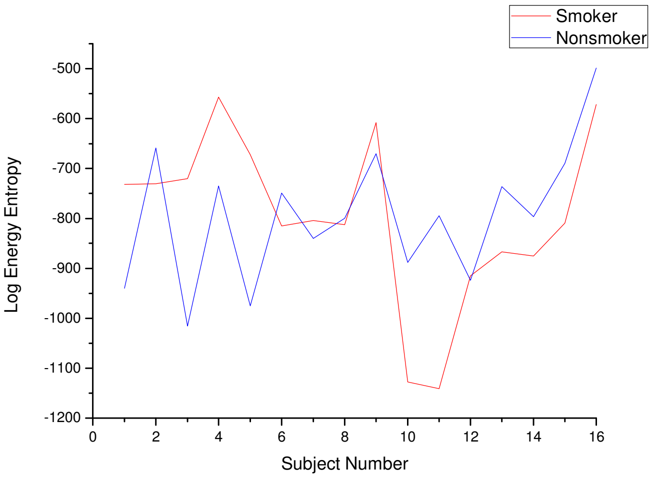



Figure 14(a): d7 level

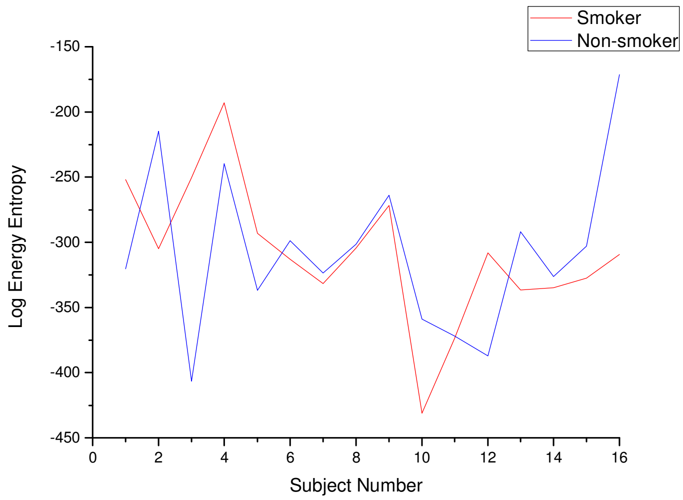



Figure 14(b): d8 level

## Feature Space Plot

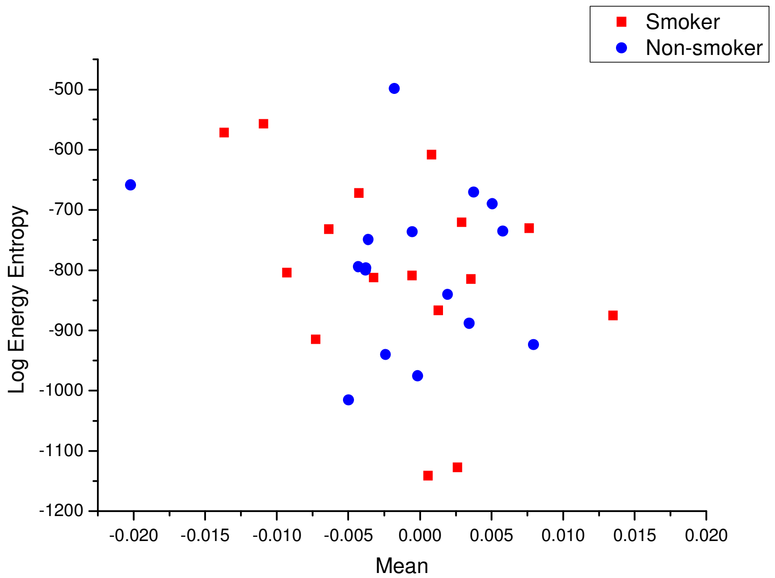



Figure 15(a): Log energy vs. mean plot for d7 level

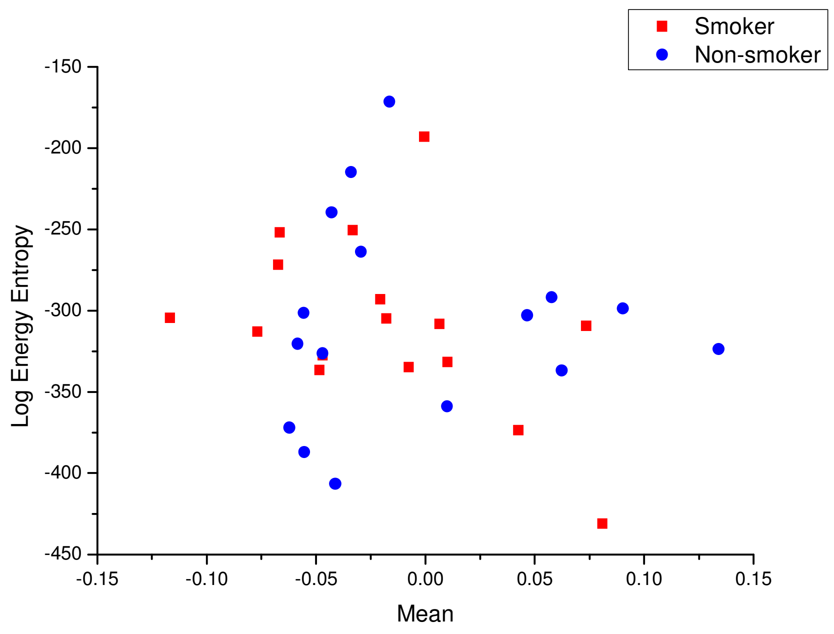



Figure 15(b): Log energy vs. mean plot for d8 level

# 5. CONCLUSIONS

The current study deals with the HRV and wavelet based ECG signal analysis of the volunteers within the age group of 18 and 26 yrs. The HRV analysis showed that the autonomic nervous system functioning of the smokers were slightly biased towards the sympathetic activity as compared to the non-smokers. The factors which were not taken into account for the HRV analysis might have played a masking effect on the analysis due to which no clear trend was obtained. The masking effects, if any, may be overcome by increasing the sample size of the control and test groups. The ECG signal analysis was carried out using wavelet transformation (Db6 wavelet) which was used for extracting the features of the ECG. The best features (mean and log energy entropy) were plotted in the feature space to figure out whether any classifier may be used for classification. The results indicate that the features were overlapping with each other and hence no classifier may be used for signal classification. It can be concluded from wavelet-based signal analysis that smoking does not apparently affect the conduction path of the ECG signals. The effect of smoking on the cardiovascular system may be due to its action on other cardiopulmonary tissues like cardiac muscles, lungs activity and blood.

## REFERENCES


[1] Sumana Chatterjee, Subhrangsu Aditya, D. N. Tibarewala. A comparative study between females of pre-pubertal and reproductive age groups to explore how hpg-axis affects the autonomic control over cardiac activity, National Conference on Biomechanics, March 7-8, 2009, IIT Roorkee, Page- 47-48.

[2] Sumana Chatterjee, Subhrangsu Aditya, D. N. Tibarewala. An Initiative to Study the Inter-Relation between the Activity of Heart and Hormonal System in Female, Socio-Economic Development, Challenges before Women Scientists, Technologists & Engineers, February 13-15, 2009, NIT Rourkela, Page- 64-67, ISBN No.- 93-80043-03-1.

[3] http://www.heartmath.org/research/science-of-the-heart/heart-rate-variability

[4] http://noodle.med.yale.edu

[5] http://www.faqs.org/patents

[6] http://www.faqs.org/patents

[7] http://www.heartmath.org/research/science-of-the-heart/heart-rate-variability

[8] Sumana Chatterjee, Study of HRV in women of different age groups, School of Bio-Science and Engineering, Jadhavpur University, submitted on May 2009.

[9] Heart Rate Variability**,** Standards of measurement, physiological interpretation, and clinical use, Task Force of The European Society of Cardiology and The North American society of Pacing and Electrophysiology.

[10] http://www.heartmath.org/research/science-of-the-heart/heart-rate-variability

[11] Heart Rate Variability**,** Standards of measurement, physiological interpretation, and clinical use, Task Force of The European Society of Cardiology and The North American society of Pacing and Electrophysiology.

[12] http://www.physionet.org/

[13] Influence of Cigarette Smoking on Heart Rate Variability in Young Healthy Volunteers, A.N.E. April 1999 Vol. 4, No. 2 Andrikopoulos, et al. Influence of Cigarette Smoking on HRV 21 1

[14] Heart Rate Variability**,** Standards of measurement, physiological interpretation, and clinical use, Task Force of The European Society of Cardiology and The North American society of Pacing and Electrophysiology.

[15] Heart Rate Variability**,** Standards of measurement, physiological interpretation, and clinical use, Task Force of The European Society of Cardiology and The North American society of Pacing and Electrophysiology.

[16] Short- and Long-Term Effects of Cigarette Smoking on Heart Rate Variability, The american journal of cardiology volume 65.

[17] D.Lee Fugal, Conceptual Wavelets in Digital Signal Processing, An in-depth, Practical Approach for the Non-Mathematician.

[17] D.Lee Fugal, Conceptual Wavelets in Digital Signal Processing, An in-depth, Practical Approach for the Non-Mathematician.

[18, 19, 20] D.Lee Fugal, Conceptual Wavelets in Digital Signal Processing, An in-depth, Practical Approach for the Non-Mathematician.

[21] Daubechies, I., Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.

[22] Amara Grap, An Introduction to Wavelets, IEEE Comp.Sc. And Eng., Vol. 2, No. 2, 1995.

[23] S.Z. Mahamoodibad, A. Ahmadian, M.D. Abolhasani, ECG feature extraction using Daubechies wavelets, in: Proceedings of Fifth IASTED International Conference, 2005, pp. 343–348.

[24] S. Majumder, S. Dhar, A. Sinha, A. Roy, A hybrid wavelet and time plane based method for QT interval measurement for ECG signals, in: Proceedings of Ninth International Conference on Signal Processing, ICSP 2008, pp. 2120–2123.

[25] Influence of Cigarette Smoking on Heart Rate Variability in Young Healthy Volunteers, Andrikopoulos, et al.,A.N.E. April 1999 Vol. 4, No. 2.

[26] .Saurabh Pal, Madhuchhanda Mitra, Detection of ECG characteristic points using Multiresolution Wavelet Analysis based Selective Coefficient Method, Measurement 43 (2010) 255–261

## APPENDIX-I

### Smoking Questionnaire

Study to observe the effect of smoking on Heart Rate Variability

Research initiative by National Institute of Technology-Rourkela

**Instructions**: This questionnaire attempts to discern patterns of heart rate variability as a result of smoking by individuals. The research program is being conducted for a research thesis. Please take your time to fill in the information accurately and to the best of your knowledge. The information you have provided will be used only for the purposes of this project, and will be kept strictly confidential. Thank you.

**Section I. General Information**

Participant Code: …………………….   Date: ……………..

**1.1** Name: ……………………………………   **1.2**. D.O.B: …………………

**1.3** Age: ……………   **1.4** Sex: M / F: ……………..

**1.5** Height: …………   **1.6** Weight: ………………….

**1. 7** Contact address (including Tel. Ph # and E-mail id.): ….……………………………………………………………………………………..

**1.8** Do you have any objection or difficulty in giving this information to us? Yes/ No

**1.9** When did you first start smoking?(with age) …………………….

**1.10** Places of residence and education from birth till date:

| **Education** | **Place** | **From Year** | **To year** |
|---|---|---|---|
| Kindergarten | | | |
| Primary school(upto stnd 8) | | | |
| High school(stnd 9-12) | | | |
| College | | | |

**1.11** Have you been on medication for the past few days, months? YES/ NO.

If yes, please specify the drug.

**1.12** Please state medical history of self and family regarding:

- Blood sugar levels
- Cardiac problems/ Hypertension
- Any other disorders

**<u>Section II. Domain of smoking</u>**

**2.1** What is the frequency of your smoking: (Please mention the number of cigarettes)

**1**. on a normal day
**2**. Tired / stressed out
**3**. Angry
**4**. Confused
**5**. Nervous / worried
**6**. Happy
**7**. Very excited

**2.2** Consumption of alcohol or related products :

If yes, state the frequency.

**2.3** Recreational drugs taken, if any :

Frequency of intake:

**2.4** Food habits :

<table>
<tr><th>Eating Habits</th><th>Response</th></tr>
<tr><td></td><td></td></tr>
<tr><td>a. Veg./ Non-veg.</td><td></td></tr>
<tr><td>b. Skip meals</td><td><table><tr><td>Breakfast</td><td>Lunch</td><td>Evening snacks</td><td>Dinner</td></tr></table></td></tr>
<tr><td>c. Junk food (days in a week)</td><td></td></tr>
<tr><td>d. Chocolates (do)</td><td></td></tr>
<tr><td>e. Fruits (do)</td><td></td></tr>
</table>

**2.5** Do you exercise ( morning/evening walks, gym):

**2.6** Do you involve yourself in any physical activity like outdoor sports, cycling etc :

If yes, please specify the frequency and activity

**2.7** Do you practice meditation?

**2.8** Did you smoke in the day today before coming here?

If yes, to what extent?

**2.9** Any other comments you may want to make:

**2.10** Declaration:

I, ………………………………………….., hereby give my consent to abide by the terms and conditions of the research experiment and understand that it is being conducted under my sole discretion.

…………………………………

Signature of the participant

**2.11** People in charge of administering the questionnaire:

Dr. Kunal Pal

Phone # 9178812505

Email id: pal.kunal@yahoo.com

Ruchika Goel

Phone # 9937649765

Email id: ruchika.goel27@gmail.com

## APPENDIX-II

## MATLAB CODE FOR GENERATION OF ECG FEATURES

```
clc;
clear all;
close all;
%%
file_path={
        'C:\Users\Rajendra Prasad\Documents\Desktop\Final year B.tech
project\For analysis(3 sec data)\Non-smokers\'
        'C:\Users\Rajendra Prasad\Documents\Desktop\Final year B.tech
project\For analysis(3 sec data)\Smokers\'
    };

colo=['r','b'];
no_case={'smokers' 'nonsmokers'};
 ecg_feature={'mean' 'mode'  'cov' 'var' 'std' 'log energy' 'wentropy'
'nenergy'  };

%no_feature=1;
for no_feature=1:1:8

 for j=1:1:length(file_path)

file_filter='*.txt';
file_a=strcat(char(file_path(j)),file_filter)
files=dir(file_a);
[m,n] = size(files)
i=1;
display('    ');

display(sprintf('%%%%%%%%%OPENING SUBJECT %s %%%%%%%%%',
char(file_path(j))));
z=1;

while i<=m
fname=strcat(char(file_path(j)),char(files(i).name));
display(sprintf('%%%%%%%%%PROCESSING FILE %s %%%%%%%%%', fname));
fid = fopen(fname,'r');
MyData = fscanf(fid,'%g');
status = fclose(fid);

c=MyData(10:17000);
c=c/max(c);
[c1,l1]=wavedec(c,1,'db6');
a1=appcoef(c1,l1,'db6',1);
d1=detcoef(c1,l1);

[c2,l2]=wavedec(c,2,'db6');
a2=appcoef(c2,l2,'db6',2);
d2=detcoef(c2,l2);

[c3,l3]=wavedec(c,3,'db6');
a3=appcoef(c3,l3,'db6',3);
d3=detcoef(c3,l3);
```

```
[c4,l4]=wavedec(c,4,'db6');
a4=appcoef(c4,l4,'db6',4);
d4=detcoef(c4,l4);

[c5,l5]=wavedec(c,5,'db6');
a5=appcoef(c5,l5,'db6',5);
d5=detcoef(c5,l5);

[c6,l6]=wavedec(c,6,'db6');
a6=appcoef(c6,l6,'db6',6);
d6=detcoef(c6,l6);

[c7,l7]=wavedec(c,7,'db6');
a7=appcoef(c7,l7,'db6',7);
d7=detcoef(c7,l7);

[c8,l8]=wavedec(c,8,'db6');
a8=appcoef(c8,l8,'db6',8);
d8=detcoef(c8,l8);

f=d7;
 %f=d8;
switch no_feature
        case 1,
            s(z)=( mean(f));%%%%%%%%
        case 2,
            s(z)=( mode(f));%%%%%%%%%%%%%%%%%%%%
        case 3,
            s(z)=( cov(f)); %%%%%%%
         case 4,
            s(z)=( var(f));%%%%%%
        case 5,
            s(z)=( std(f));
        case 6,
            s(z)= wentropy(f,'log energy');%%%%%%%%%
        case 7,
            s(z)= wentropy(f,'shannon');%%%%%%%%%%%
        case 8,
            s(z)=sum((f.^2))/length(f); %energy density

end

z=z+1;

figure(1)
subplot(511)
plot(c); xlabel('ECG')
subplot(512)
plot(d1); xlabel('d1')
subplot(513)
plot(d2); xlabel('d2')
subplot(514)
plot(d3); xlabel('d3')
subplot(515)
plot(d4); xlabel('d4')
```

```
figure(2)
subplot(411)
plot(d5); xlabel('d5')
subplot(412)
plot(d6); xlabel('d6')
subplot(413)
plot(d7); xlabel('d7')
subplot(414)
plot(d8); xlabel('d8')
%input('');
i=i+1;
end

figure(10+no_feature),plot(s,colo(j));
title(ecg_feature(no_feature));
hold on;
Str = 65+no_feature-1; %A=65 Z=90
  colm= sprintf('%s ', Str)


  xlscol=fn_xlcolumngen(ecg_feature(no_feature),colm,2+(m*(j-1)),m,2)
%fn_xlcolumngen(cell_title,cell_name,start_col,datalen,no_class)%eg
a,2,50,5


if j==1
xlswrite('r.xls',xlscol(1),'db',char(xlscol(2)));
 %input('');
end

xlswrite('r.xls',s','db',char(xlscol(3)));


%input('');

%xlscol=fn_xlcolumngen(no_case(j),colm,1,1,2)
%fn_xlcolumngen(cell_title,cell_name,start_col,datalen,no_class)%eg
a,2,50,5
%input('');

%input('');
%%
 end

hold off;
end
```

# APPENDIX-III

## MATLAB CODE FOR FEATURE SPACE PLOT

```
clc;
clear all;
close all;
%%

file_path={

    'C:\Users\Rajendra Prasad\Documents\Desktop\Final year B.tech
project\For analysis(3 sec data)\ECG feature extraction\'

        };

    for j=1:1:length(file_path)

file_filter='*.xls';
file_a=strcat(char(file_path(j)),file_filter);
files=dir(file_a);
[m,n] = size(files);
i=1;
display('    ');
display(sprintf('%%%%%%%%%OPENING FOLDER %s %%%%%%%%%',
char(file_path(j))));

while i<=m
%sheetname=regexprep(char(files(i).name), '.xls', '')
sheetname='db'
[feature1,feature1_name]=xlsread(strcat(char(file_path(j)),char(files(i).na
me)), sheetname,'A1:A33');
[feature2,feature2_name]=xlsread(strcat(char(file_path(j)),char(files(i).na
me)), sheetname,'F1:F33');

fcmdata(:,1)=feature1;
fcmdata(:,2)=feature2;
length(fcmdata(:,1))
length(fcmdata(:,2))

handel=figure(1),plot(fcmdata(1:16,1),fcmdata(1:16,2),'o','color','k');
hold on
figure(1),plot(fcmdata(17:32,1),fcmdata(17:32,2),'o','color','r');
hold off;

title('feature space');
xlabel('mean');
ylabel('log energy');

s=strcat('cluster',char(files(i).name))

s = regexprep(s, '.xls', '')
s=strcat(char(file_path(j)),s);
```

```
s=strcat(s,'.jpg')
saveas(handel,s,'jpg');
qw=input('');
%close all;
i=i+1;
end
    end
```