Linear and nonlinear analysis of normal and CAD-affected heart rate signals

ACHARYA, U Rajendra, FAUST, Oliver <http://orcid.org/0000-0002-0352-6716>, SREE, Vinitha, SWAPNA, G, MARTIS, Roshan Joy, KADRI, Nahrizul Adib and SURI, Jasjit S

Available from Sheffield Hallam University Research Archive (SHURA) at:
http://shura.shu.ac.uk/11430/

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version


Copyright and re-use policy

See http://shura.shu.ac.uk/information.html
Linear and Nonlinear Analysis of Normal and CAD-affected Heart Rate Signals

U Rajendra Acharya\textsuperscript{a,b}, Oliver Faust\textsuperscript{a,d}, Vinitha Sree\textsuperscript{c}, Swapna G\textsuperscript{f}, Roshan Joy Martis\textsuperscript{a}, Nahrizul Adib Kadri\textsuperscript{b}, Jasjit S Suri\textsuperscript{g}

\textsuperscript{a}Department of Electronics and Communication Engineering, Ngee Ann Polytechnic, Singapore 599489
\textsuperscript{b}Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
\textsuperscript{c}Global Biomedical Technologies Inc., CA, USA
\textsuperscript{d}School of Electronic Information Engineering, Tianjing University, China
\textsuperscript{e}School of Electrical Engineering, University of Aberdeen, Aberdeen
\textsuperscript{f}Department of Applied Electronics & Instrumentation, Government Engineering College, Kozhikode, Kerala 673005, India
\textsuperscript{g}Fellow AIMBE, CTO, Global Biomedical Technologies, CA, USA; Biomedical Engineering Department, Idaho State University (Aff.), ID, USA

*Corresponding author: Oliver Faust, email: faust_o@web.de

Abstract

Coronary Artery Disease (CAD) is one of the dangerous cardiac disease, often may lead to sudden cardiac death. It is difficult to diagnose CAD by manual inspection of electrocardiogram (ECG) signals. To automate this detection task, in this study, we extracted the Heart Rate (HR) from the ECG signals and used them as base signal for further analysis. We then analyzed the HR signals of both normal and CAD subjects using (i) time domain, (ii) frequency domain and (iii) nonlinear techniques. The following are the nonlinear methods that were used in this work: Poincare plots, Recurrence Quantification Analysis
(RQA) parameters, Shannon entropy, Approximate Entropy (ApEn), Sample Entropy (SampEn), Higher Order Spectra (HOS) methods, Detrended Fluctuation Analysis (DFA), Empirical Mode Decomposition (EMD), Cumulants, and Correlation Dimension. As a result of the analysis, we present unique recurrence, Poincare and HOS plots for normal and CAD subjects. We have also observed significant variations in the range of these features with respect to normal and CAD classes, and have presented the same in this paper. We found that the RQA parameters were higher for CAD subjects indicating more rhythm. Since the activity of CAD subjects is less, similar signal patterns repeat more frequently compared to the normal subjects. The entropy based parameters, ApEn and SampEn, are lower for CAD subjects indicating lower entropy (less activity due to impairment) for CAD. Almost all HOS parameters showed higher values for the CAD group, indicating the presence of higher frequency content in the CAD signals. Thus, our study provides a deep insight into how such nonlinear features could be exploited to effectively and reliably detect the presence of CAD.

**Keywords**: heart rate, CAD, ECG, HOS, Poincare plot, recurrence plot, EMD.

### 1. Introduction

Coronary arteries supply nutrients and oxygen to heart muscles. Coronary Artery Disease (CAD) is a pathological condition where the diameter of the arteries decreases either due to the formation of cholesterol plaque on its inner wall [Steinberget al., 1999] or due to the contraction of the whole wall for other reasons, such as tobacco smoking [Ockene et al., 1997] and environmental pollution [Brook et al., 2004]. The condition is often ominously silent, but progressive in nature. If it is not treated appropriately, it will eventually lead to ischemia (i.e., interruptions of blood supply) and then infarctions (i.e., the complete loss of blood supply). Usually one of the reasons for Sudden Cardiac Death (SCD) is CAD [Thompson et al., 2006]. Hence, early detection of CAD is essential to prevent SCD.

One of the most commonly used techniques for CAD detection is the Exercise Stress Test (EST). EST increases the workload of the heart and records exaggerated electrophysiological information. For this test to be accurate, a target Heart Rate (HR) has to be attained. Not all CAD patients can reach this rate. Furthermore there is considerable risk
for the patient, because such a stress test can trigger Ventricular Tachycardia (VT) or cardiac arrest [San Roman et al., 1998].

Electrocardiogram (ECG) could be a useful physiological measurement tool to detect the presence of CAD. However, visual interpretation of the ECG signals is not so effective as 50 - 70% of CAD patients do not show any notable difference in their ECGs [Silber et al., 1975]. However, the minute variations in the ECG signals have to be identified in order to diagnose specific type of heart disease. Due to the presence of noise and baseline wander, it is tedious to detect the minute variations by evaluating the morphological features of ECG signals. Hence, in this study, we extracted the HR from the ECG signals and used them for analysis. The study of Heart Rate Variability (HRV) is a better technique to diagnose CAD risk levels. HR is a nonlinear, non-stationary signal which indicates the subtle variations of the underlying ECG signal [Acharya et al., 2004a]. The HRV evaluates the changes in the consecutive heart rates and it assesses the health of the Autonomic Nervous System (ANS) non-invasively. The HRV analysis conveys information about homeostasis of the body [Lombardi 2000]. Standard methods to analyze the HRV were proposed in various domains [Task Force, 1996].

Various cardiac and non-cardiac diseases have been diagnosed using HR signals [Isler et al., 2007, Schumann et al., 2002, Acharya et al., 2004a, Gujjar et al., 2004, Carney et al., 2000]. They have analyzed the HR signals using various linear and non-linear techniques [Acharya et al., 2004a; 2007]. Huikuri et al. (1994) have analyzed the CAD subjects using HRV signals and showed that, the circadian rhythm decreases in CAD subjects. Hayano et al. (1990) have shown a correlation between CAD severity and a reduction in low-frequency power. Reduction decrease in high frequency power were shown in CAD subjects [Lavoie et al. (2004), Nikolopoulos et al. (2003)] and features of time and frequency domain were found to be lower for CAD subjects [Bigger et al. (1995)]. The statistical measures changes with time and hence time domain analysis is not effective and effectiveness of frequency domain analysis decreases with reduction in the signal to noise ratio [Acharya et al., 2006].

Nonlinear techniques are more in tune with the nature of physiological signals and systems, therefore, they outperform time and frequency domain methods. Hence, they are widely used in many biological and medical applications [Acharya et al., 2003; Fell et al.,
Owis et al. (2002) performed ECG-based arrhythmia detection and classification based on nonlinear modeling. Sun et al. (2000), Acharya et al. (2007) and Chua et al. (2008) used nonlinear techniques to analyze cardiac signals for the development of cardiac arrhythmia detection algorithms. Schumacher et al. (2004) elaborated the effectiveness of linear and nonlinear techniques in analyzing HR signals. The onset of various cardiovascular diseases like, Ventricular Tachycardia (VT) and Congestive Cardiac Failure (CCF) can be predicted using non-linear analysis of HR signals [Cohen et al. 1996]. Chua et al. (2006) introduced a method to extract features like bispectral entropy from HR signals by employing Higher Order Spectra (HOS) techniques. In their study, HOS features from HR signals were used to differentiate between a normal heart beat and seven arrhythmia classes. CAD results in reduced Baroreflex Sensitivity (BRS) and reduced vagal activity which can be understood by HRV analysis. BRS is an indicator of increased risk of SCD in myocardial infarction patients. Arica et al. (2010) used HR and systolic pressure signals to assess BRS.

The main aim of this paper is to present time, frequency and non-linear features for normal and CAD-affected HR signals. For this analysis, we extracted and analyzed features in the time domain, frequency domain, and also studied features derived using nonlinear methods. Furthermore, we have proposed various ranges for these features and presented unique nonlinear plots for the normal and CAD classes. Our results show that CAD subjects have less variability in their heart rate signal when compared to normal subjects. This reduced variability can be used as a single measure to diagnose CAD from ECG signals which were obtained under normal conditions. We predict that the consequent use of HRV measures will reduce the need to conduct stress ECG measurements, and therefore, expose patients to less risk.

2. Data Used

ECG signals from 10 CAD patients and an equal number of healthy volunteers were recorded using the BIOPAC™ equipment [http://www.biopac.com/]. The sampling frequency of ECG signal was 500 Hz. The average age of both normal and CAD subjects was 55 years (age varied from 40 to 70 years). The CAD patients used for this study, were taken from Iqraa Hospital, Calicut, Kerala, India. Subjects having normal blood pressure,
glucose level and ECG were considered in the normal category. For the CAD patients, coronary angiography (CAG) was performed. Patients with more than 50% narrowing in the left main artery were considered for this study. Patients suffering from bundle branch block (left or right bundle branch block), hypertrophy, atrial fibrillation, congestive heart failure, myopathy, and taking any cardiac medication are excluded in this study. The patients were selected by a cardiologist based on the similarity of their medications. It was assumed that the drug effects on the HR signal were similar. The data comprised a total of 61 normal and 82 ECG CAD datasets; each set had 1000 samples from 10 subjects. The variations of ECG signals in CAD and normal subjects are shown in Figure 1.

![Figure 1](image1.png)

(a) Typical RR signal: (a) normal (b) CAD.

The ECG beats were sent via a band pass filter with a lower cut off frequency of 0.3 Hz to eliminate baseline wander and higher cut off frequency of 50 Hz to eliminate the noise. A band-stop filter of cut-off frequency 50 Hz was used to eliminate power source influences. In the final step, the R peaks were located using Pan and Tompkins algorithm [Pan et al., 1985, Wariar et al., 1991]. The time duration between two consecutive R peaks is termed as RR interval \((t_{RR})\). Heart rate is defined as:

\[
HR_{bpm} = \frac{60}{t_{RR}} \quad \text{(Beats Per Minute)}
\]

(1)
3. Methods Used

This section discusses time, frequency and nonlinear domain techniques which were used for analyzing CAD and normal HR signals.

3.1. Time domain analysis

RR interval is one of the time domain parameter which reflects the combined influence of SNS and PNS. The mean of the RR time intervals is the mean RR parameter. Most of the time domain parameters (both short term and long-term variation indices) are derived from the RR intervals. Apart from the RR interval, we have calculated RMSSD (root mean square standard deviation), NN50 (number of pairs of consecutive NNs which vary greater than 50 ms) and pNN50 (ratio of NN50 divided by total number of NNs). RMSSD (in milliseconds) indicates the parasympathetic control of HR during the normal rhythm.

3.2. Frequency domain analysis

The above discussed time domain technique is easy to implement and use. But its ability to separate sympathetic and parasympathetic influences using heart rate signal is limited. The cardiac health of the subject can be evaluated using the power spectrum of the HR signal [Akselrod et al., 1981]. There are the three main frequency regions of the heart rate signal.

- The power in the frequency range from 0.15 Hz to 0.5 Hz is defined as high frequency (HF) power band.
- The power in the frequency range from 0.04 Hz to 0.15 Hz is defined as low frequency (LF) power band.
- The power in the frequency range from 0.0033 Hz to 0.04 Hz is defined as very-low-frequency (VLF) power band.

HF region is an indicator of the vagal activity and respiratory sinus arrhythmia (RSA), while LF refers to the baroreceptor control mechanisms and the combined effect of sympathetic and vagal systems. The VLF power spectrum indicates the vascular mechanisms and rennin-angiotension systems. In our work, we measured total power, HF, LF as well as the ratio of LF to HF power.
Frequency domain study is normally conducted using by using the using Fast Fourier Transform to estimate the Power Spectral Density (PSD). AR (Autoregressive) modeling is also a frequency domain analysis method [Faust et al., 2004]. The AR parameters are estimated by solving linear equations. A suitable filter order has to be selected. In this study, we have used order of AR model as 16 [Akaike, 1969; 1974; Anita et al., 2002]. Figure 2 shows the typical PSD of a normal HR signal (Figure 2(a)) and a CAD HR signal (Figure 2(b)).

![Figure 2 Typical PSD of heart rate signal: (a) normal subject (b) CAD subject.](image)

The frequency domain plots are divided into three regions. Each frequency band depicts a physiological processes and pathologies. The initial band is the VLF band followed by LF and then the unshaded portion is the HF range. The VLF variations are associated slow processes like thermal regulation, LF region relates to arterial blood pressure control (both sympathetic and parasympathetic effects) and HF band reflects respiration and parasympathetic activity. From the figures above, it can be observed that the PSD for the CAD subject is approximately half compared to the value of the normal subject for both VLF and LF regions. For the HF region, the PSD is almost the same for normal and CAD. This means that CAD reduces the activity of thermoregulatory and sympathetic systems, but the parasympathetic systems remain unaffected.

The Fourier transform indicates both amplitude and phase of different frequency components which are contained in a time domain signal. A drawback of this method is that
it does not specify the time when these frequencies occur in the signal. So, frequency domain
analysis is not preferable to study HR signals. Short Time Fourier Transform, based on a
finite length sliding window, solves this issue to a certain degree, but it doesn’t work well
for rapidly varying signals. HR signals are non-stationary. They are also nonlinear, and
hence, conventional time and frequency domain techniques cannot capture all the
information contained in the higher harmonics (order greater than 2) of the HR signal.
Nonlinear analysis methods and higher order spectrum can capture the higher harmonics
information contained in HR signals.

3.3. Nonlinear methods
The theory of nonlinear dynamics is widely to analyze the bio signals, which are nonlinear
in nature [Acharya et al., 2004b; 2006; 2007; Faust et al., 2012]. The following are the
nonlinear methods that were used in this work: Poincare plots, Recurrence Quantification
Analysis (RQA) parameters, Shannon entropy, Approximate Entropy (ApEn), Sample
Entropy (SampEn), Higher Order Spectra (HOS) methods, Detrended Fluctuation Analysis
(DFA), Empirical Mode Decomposition (EMD), Cumulants, and Correlation Dimension.
These methods are briefly explained in the following sub-sections.

3.3.1. Poincare geometry
It is a visual plot, which was adopted from nonlinear methods, to study the behaviours of RR
interval variability. It depicts the correlation between consecutive intervals in graphical
representation. This plot shows the comprehensive every beat to beat variation [Woo et al.
1992, Kamen et al. 1996]. These plots are studied mathematically by determining the
standard deviations of the lengths of RR intervals (RR(n)) [Tulppo et al., 1996]. The short
term variability (SD1) of the heart signal is measured by the points that are perpendicular to
the line-of-identity and long term variability by the points along the line-of-identity. By
visually examining the Poincare plot shapes, we can discriminate normal from CAD
subjects. In this paper, SD1 parameter was used to detect CAD. Figure 3 shows the
Poincare plots of normal (Figure 3(a)) and CAD (Figure 3(b)) subjects. The plots are ellipse
shaped and centre-aligned. SD2 describes the long term variability of RR(n) (instantaneous
RR), while SD1 indicates the shorter-term variability of RR(n). In the plot for CAD, SD2 and
SD1 are very low compared to the normal plot.
3.3.2 Recurrence quantification analysis (RQA)

Recurrence Plot (RP) indicates, for a given instant of time, the times at which path of the phase space meets the same location in the phase space. The duration and counts of recurrences of the dynamical systems are estimated by RQA. It measures the dynamicity and subtle rhythmicity in the HR signal. The RQA parameters evaluate the complexity and non-stationary nature of the time series [Webber et al., (1994), Zbilut et al., (1992) and Marwan et al., (2002)]. Zbilut at al. (2002) showed the usefulness of RQA in detecting randomness and complexity in non-stationary heart beats which cannot be analyzed easily by conventional techniques. In this study, following RQA features were used:

- **Mean diagonal line length** ($<L>$ or $L_{mean}$): depicts the average time of forecasting of the system. It can be written as:
  \[
  L_{mean} = \frac{\sum_{i=1}^{N_l} p_i}{\sum_{i=1}^{N_l} p_i} \tag{2}
  \]

- Max line length ($L_{max}$): It is the largest distance of the diagonal of the RP and is given by:
  \[
  L_{max} = \max\{|l_i|; i = 1,\ldots, N_l\} \tag{3}
  \]
  Here, $N_l$ indicates the number of diagonal lines in the RP.

- Recurrence Rate (REC): It describes the cloud of recurrence points existing in the plot.
REC is defined as:

\[
REC = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{ij}
\]  

Where \( R_{ij} \) represents the recurrence points, \( N \) is the amount of points on phase space path \( R_{ij} \). Determinism (DET): is the portion of \( R_{ij} \) that contribute to the diagonal lines in the plot. It explains the predictability of the dynamical system.

\[
DET = \frac{\sum_{l=l_{\text{min}}}^{l_{\text{max}}} P(l)}{\sum_{l=1}^{l_{\text{max}}} R_{ij}}
\]

Here, \( P(l) \) is the distribution in the frequency domain of the diagonal line with lengths \( l \) and the minimum diagonal line length is given by \( l_{\text{min}} \).
Figure 4 presents typical RP for normal (Figure 4(a)) and CAD subjects (Figure 4(b)). More variation (dots) is predominant in the normal signal compared to the CAD class. Moreover, there is a more regular pattern in the recurrence plot of CAD. This indicates that there is more rhythmicity with respect to normal subjects.

3.3.3. Approximate entropy (ApEn)

It indicates the fluctuation in the time domain signal [Pincus, 1991]. The value of ApEn is higher for more varying data. Hence, more varying time domain signals will have higher ApEn values, while regular and predictable time series signals will have lower ApEn values. ApEn is given by:

$$ApEn(m, r, N) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log C_i^m(r) - \frac{1}{N-m} \sum_{i=1}^{N-m} \log C_i^{m+1}(r)$$

where

$$C_i^m(r) = \frac{1}{N-m+1} \sum_{j=1}^{N-m+1} \theta(r - \|x_i - x_j\|)$$

is the correlation integral.

Furthermore, $x_i, x_j, x_k, x_l \rightarrow$ Phase space trajectory points,
In this work, we have used, ‘$r$’ is the embedding dimension($m$) to 2 and the radial distance $r$ was set to $0.2 \times$ the time series standard deviation [Thakor et al., 2004].

3.3.4. Sample Entropy (SampEn)

It quantifies the complexity in signal [Richman et al., 2000]. Higher values of SampEn describes more irregularities in the time series. It is more refined than ApEn. In order to evaluate sample entropy, continuous matching of points inside the radius ‘$r$’ are done as long as there is match exists. The variables $A(k)$ and $B(k)$ for all lengths $k$ up to ‘$e$’ keep track of all matching templates. It is given by:

$$SampEn(k, r, N) = -\ln \frac{A(k)}{B(k-1)}$$

for $k = 0, 1, ..., m-1$ with $B(0) = N$, the length of the HR signal, $r$ is taken as 0.2 and $m$ (maximum template length) is set to 2 [Song et al., 2010].

3.3.5. Detrended fluctuation analysis (DFA)

It assess the self-similar properties of short term HR signals [Peng et al., 1996]. The roughness of the signal is indicated by the factor ‘$\alpha$’. This value is close to 1 for normal subjects and may have unique ranges for various cardiac classes.

3.3.6. Correlation Dimension (D2)

D2 is a useful measure of self-similarity of a signal [Grassberger et al., 1983]. According to the algorithm [Grassberger et al., 1983], Correlation integral ($C(r)$) function is constructed first. It was performed by measuring the gap between $N$ pairs of data points and arranging the output $dr$ proportional to $r$. The gap between a pair of points is estimated by $s(i,j) = |X_i - X_j|$.

$C(r)$ is given by:

$$C(r) = \frac{1}{N^2} \sum_{x=1}^{N} \sum_{y=1, x \neq y}^{N} \Theta(r-|X_x - X_y|)$$

Where, $X_x$ and $X_y$: indicate phase space trajectory points,

$N$: total amount of phase space points,

$R$: radial length of a circular disc centered at $X_r$. 

---

N→ Amount of points in the phase space.

$r$→ radial length of a circular disc centered at the reference point $x_iX_i$.

$\Theta$→ Step function.
Correlation Dimension (D2) can be described by:

$$D2 = \lim_{r \to 0} \frac{\log[C(r)]}{\log(r)}$$  \hspace{1cm} (10)

D2 will have higher value, if the RR variations is more and vice versa.

3.3.7. Higher Order Spectrum (HOS)

It is a novel tool for evaluating non-Gaussian and non-stationary bio-signals. It identifies diversions from Gaussianity and phase correlations among frequency components of the signal [Chua et al., 2010]. HOS is more immune to noise and can retain the actual phase information of the signal. The 3rd order statistics is the bispectrum $B(f_1, f_2)$. It is the Fourier transform of the 3rd order correlation of a signal and is given by:

$$B(f_1, f_2) = E[X(f_1)X(f_2)X(f_1 + f_2)]$$  \hspace{1cm} (11)

Where, $X(f)$ is the Fourier transform of input $X(nT)$, $n$ is the variable, $T$ is the sampling period, and $E[.]$ is expectation operator. The normalized bispectrum ($B_{\text{norm}}(f_1, f_2)$) will have magnitude range 0 to 1 [Nikias et al., 1987; 1993a] and is defined as

$$B_{\text{norm}}(f_1, f_2) = \frac{E[X(f_1)X(f_2)X(f_1 + f_2)]}{\sqrt{P(f_1)P(f_2)P(f_1 + f_2)}}$$  \hspace{1cm} (12)

Where $P(f)$ is the power spectrum. In this paper, we have presented discriminating bicoherence and bispectrum plots for normal and CAD HR signals. Both bispectrum and bicoherence plots exhibit symmetry as they are products of three Fourier coefficients.

Various features can be estimated from the bispectrum and few of them are given below:

a) Normalized Bispectral squared entropy1 (P1) is given by

$$P_1 = -\sum q_i \log q_i$$  \hspace{1cm} (13)

where $q_n = \frac{|B(f_1, f_2)|^2}{\sum_{\Omega}|B(f_1, f_2)|^2}$ and $\Omega$ is the region where $f_1 > f_2$ and $f_1 + f_2 < 1$ is as shown in the plot below.
The region $\Omega$, is the principal domain used for the evaluation of the bispectrum.

b) Normalised Bispectrum entropy 2 is given by:

\[
P_2 = -\sum_i r_i \log r_i
\]

where $r_n = \frac{|B(f_1, f_2)|^3}{\sum|B(f_1, f_2)|^3}$

$\Omega = \text{is the principal domain shown in Figure 5.}$

c) The mean bispectrum magnitude is:

\[
M_{\text{ave}} = \frac{1}{L} \sum_{\Omega} |B(f_1, f_2)|
\]

$M_{\text{ave}}$ can be used to distinguish between two classes.

d) The bispectrum phase entropy is:

\[
P_p = \sum_{\Psi} p(\Psi_n) \log p(\Psi_n)
\]

Where:

\[
p(\Psi_n) = \frac{1}{L} \sum_{\alpha} 1(\Phi(b(f_1, f_2)) \in \Psi_n)
\]

And:

\[
\Psi_n = \{ \Phi / -\pi + 2\pi n / N \leq \Phi < -\pi + 2\pi(n+1) / N, n = 0,1,......N-1. \}
\]

Here, $L$ indicates the amount of points in the principal region and $\Phi$ is the bispectrum phase angle. The function $1(\cdot)$ yields 1 if $\phi$ (phase angle) falls inside bin $\Psi_n$.

e) The weighted center of bispectrum (WCOB) is defined as:
where, $i$ and $j$ are frequency bin index in the non-redundant region.

Figure 6(a) presents the contour plots of bispectrum for normal and CAD heart rate signal (Figure 6(b)) respectively. By visually examining these plots, we can distinguish between normal and CAD subjects clearly. In the bispectrum plot of the normal subject (Figure 6(a)), there are peaks concentrated in the centre, while for CAD subject (Figure 6(b)), the peaks are present throughout the plot and spread throughout the frequency spectrum.
3.3.8. Cumulant Computation

It is not easy to analyze the nonlinear and non-stationary behavior time series using 1st and 2nd order statistics [Nikias, 1993b]. So, third order cumulant which is a third order correlation derived from HOS can be used for HR signals. It has been successfully implemented to differentiate automatically control, ictal and interictal EEG signals [Acharya et al., 2011].

Let \( \{x_1, x_2, x_3, \ldots, x_k\} \) indicate a \( k \) dimensional random process of zero mean value. Its moments are given by [Nikias, 1993):

\[
\begin{align*}
    m_1^x &= E[x(n)] \\
    m_2^x(i) &= E[x(n)x(n+i)] \\
    m_3^x(i, j) &= E[x(n)x(n+i)x(n+j)] \\
    m_4^x(i, j, k) &= E[x(n)x(n+i)x(n+j)x(n+k)]
\end{align*}
\]

where \( m_1^x, m_2^x, m_3^x \) and \( m_4^x \) are 1st, 2nd, 3rd and 4th order moments, \( E[.] \) indicates the expectation operator, and time lag parameters are \( i, j \). Using moments, cumulants are evaluated as [Nikias, 1993]:

\[
\begin{align*}
    \gamma_1^x &= m_1^x \\
    \gamma_2^x(i) &= m_2^x(i) / m_1^x \\
    \gamma_3^x(i, j) &= m_3^x(i, j) / m_1^x \\
    \gamma_4^x(i, j, k) &= m_4^x(i, j, k) / m_1^x
\end{align*}
\]

Figure 6 Typical Bispectrum and its contour plots for (a) Normal and (b) CAD subjects.
\[ C_1^i = m_i^i \]  \hspace{2cm} (22)

\[ C_2^i = m_i^2(i) \]  \hspace{2cm} (23)

\[ C_3^i = m_i^3(i,j) \]  \hspace{2cm} (24)

\[ C_4^i = m_i^4(i,j,k) - m_i^2(i)m_i^2(j-k) - m_i^2(k-i) - m_i^2(k)m_i^2(i-j) \]  \hspace{2cm} (25)

where \(C_1^i, C_2^i, C_3^i\) and \(C_4^i\) are the 1st, 2nd, 3rd and 4th order cumulants respectively. In the current study the third order cumulant is used for the analysis of HR signals. Figure 7(a) shows the 3rd order cumulant plot and its contour plot for normal HR signal and Figure 7(b) for CAD HR signal.
3.3.8 Empirical Mode Decomposition (EMD)

It is a direct, adaptive and data dependent model for nonlinear signal analysis. It does not assume linearity and stationarity conditions [Huang et al., 1998]. Any complicated signal can be decomposed into a group of Intrinsic Mode Functions (IMFs) which are AM and FM modulated waveforms. The decomposition is based on local time and scale of the signal. Martis et al. (2012) applied EMD for the analysis for EEG signals of control, preictal and ictal classes. Figure 8 presents eight IMFs of typical normal (Figure 8(a)) and CAD (Figure 8(b)) HR signal. Various important features can extracted from these IMFs to classify the normal and CAD HR signals.

Figure 7 Typical third cumulant and its contour plots: (a) normal and (b) CAD subject.
Figure 8 Typical IMFs extracted from EMD decomposition for HR signal: (a) normal and (b) CAD.
4. Results

Results of time domain, frequency domain and nonlinear techniques are presented in this section. Table 1 shows the feature values (Mean ± Standard Deviation(SD)) of the time domain parameters of normal and CAD HR signals. In this work, four time domain features were found to be clinically significant (p<0.05). They are mean HR, RMSSD, NN50 and pNN50 (listed in Table 1).

<table>
<thead>
<tr>
<th>Features</th>
<th>Normal (Mean±SD)</th>
<th>CAD (Mean±SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean HR</td>
<td>52.9±6.62</td>
<td>45.6±16.1</td>
<td>0.0008</td>
</tr>
<tr>
<td>RMSSD</td>
<td>44.5 ±16.0</td>
<td>72.7 ± 99.0</td>
<td>0.021</td>
</tr>
<tr>
<td>NN50</td>
<td>187±145</td>
<td>68.3 ± 103</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>pNN50</td>
<td>21.5 ± 15.7</td>
<td>7.31 ± 11.5</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

The clinically significant features like NN50 and pNN50, have lower values for the CAD subjects with respect to the normal. The difference is more than order 2. The next significant parameter, mean HR, is lower for CAD signals than for normal subjects. RMSSD is higher for CAD than for normal subjects.

In the frequency domain analysis, we have also obtained four clinically significant features for LF to HF. The ratio LF to HF indicates sympathetic parasympathetic balance of heart. Table 2 shows frequency domain analysis results for CAD and normal heart rate signals.

<table>
<thead>
<tr>
<th>Features</th>
<th>Normal (Mean ± SD)</th>
<th>CAD (Mean ± SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF/HF</td>
<td>2.93±2.46</td>
<td>943±3.109E+03</td>
<td>0.013</td>
</tr>
</tbody>
</table>

In our work, all the four frequency domain features have higher values for the CAD than for normal subjects. We have used a variety of nonlinear parameters for analysis. Table 3 gives the summary of these nonlinear features.
### Table 3 Results of nonlinear analysis.

<table>
<thead>
<tr>
<th>Features</th>
<th>Normal (Mean±SD)</th>
<th>CAD (Mean±SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD1</td>
<td>31.5 ± 11.3</td>
<td>52.5 ±69.4</td>
<td>0.014</td>
</tr>
<tr>
<td>L&lt;sub&gt;mean&lt;/sub&gt;</td>
<td>14.3±5.57</td>
<td>39.1± 45.5</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Max line length (L&lt;sub&gt;max&lt;/sub&gt;)</td>
<td>378± 229</td>
<td>513±290</td>
<td>0.0044</td>
</tr>
<tr>
<td>Recurrence rate (REC)</td>
<td>38.5 ±10.5</td>
<td>55.7±18.4</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Determinism (DET)</td>
<td>98.4 ± 1.10</td>
<td>99.4 ± 0.772</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>ApEn</td>
<td>1.33 ± 0.121</td>
<td>1.05 ±0.288</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>SampEn</td>
<td>1.47 ±0.225</td>
<td>1.04 ±0.390</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>DFA (α1)</td>
<td>1.15 ± 0.209</td>
<td>0.933±0.407</td>
<td>0.0002</td>
</tr>
<tr>
<td>Correlation dimension (D2)</td>
<td>3.41 ± 1.27</td>
<td>1.07 ± 1.16</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

SD1 measures the short term variability of the heart signal. This value (SD1) for CAD signals is higher than for normal signals. Thus, SD1 reflects the fast variations brought by CAD on heartbeat. The next four parameters, L<sub>mean</sub>, L<sub>max</sub>, REC and DET, belong to the RQA analysis. The values of these four parameters are high for the CAD group. For the first three values, the increase was significant while for the last parameter, CAD group showed only a slight increase compared to the normal group. The higher RQA parameters indicate more order or less variation in the signal. Hence, higher values of RQA parameters correctly indicate that the variation in CAD is less compared to normal subjects.

Entropy parameters (ApEn and SampEn) showed higher values for normal HR signal compared to CAD. ApEn value will be small for cardiac impairment cases. It is evident from Table 3 that for the CAD signal, ApEn takes a value much less than normal subjects. SampEn parameter also showed low value for CAD. In general, the results showed a reduction in entropy-based parameters for CAD. That means the entropy is reduced due to the reduction in HRV for CAD.

The DFA parameter takes a large value as the input time series signal is more rhythmic. Accordingly, for normal subjects, we obtained larger values for DFA compared to CAD subjects.
D2 is a quantitative measurement which indicates the nature of the path in a phase space. D2 decreases as the beat-to-beat variation decreases [Acharya et al., 2004a; 2004b]. The D2 value obtained for the CAD class is about one-third of that for the normal class.

Table 4 gives details of HOS parameters which were extracted, and the corresponding p-value. Except phase entropy ($P_e$), all HOS parameters showed higher values for the CAD class. CAD, with its high beat-to-beat variability, results in higher values for HOS parameters. The disorder in the HR signals of a CAD subject shows itself as an increase in the information content in the higher harmonics of the HR signal. CAD brings short term fast beat-to-beat variability, thus causing the signal to contain extra information in higher harmonics compared to normal HR signals.

Table 4 Results of HOS analysis.

<table>
<thead>
<tr>
<th>Features</th>
<th>Normal (Mean ± SD)</th>
<th>CAD (Mean ± SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>0.427 ±0.151</td>
<td>0.541 ±0.310</td>
<td>0.0074</td>
</tr>
<tr>
<td>$P_2$</td>
<td>0.246±0.126</td>
<td>0.405 ± 0.300</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>$M_{avg}$</td>
<td>0.392 ±0.481</td>
<td>134± 393</td>
<td>0.0052</td>
</tr>
<tr>
<td>$P_e$</td>
<td>3.55 ± 5.488E-02</td>
<td>3.12 ± 0.867</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Wcob1</td>
<td>26.3 ± 18.0</td>
<td>39.5± 29.3</td>
<td>0.0023</td>
</tr>
<tr>
<td>Wcob3</td>
<td>34.7 ± 10.6</td>
<td>43.2 ±21.4</td>
<td>0.0038</td>
</tr>
<tr>
<td>Wcob4</td>
<td>10.3 ±2.87</td>
<td>12.8 ± 7.31</td>
<td>0.0086</td>
</tr>
</tbody>
</table>

5. Discussion

Goldberger et al. (1987) showed that under normal conditions our heart is not a periodic oscillator. Since then, several nonlinear methods were proposed to quantitatively measure the heart rate variations [Goldberger et al., 1987; Pincus 1991]. Nonlinear parameters like recurrence percentage, fractal dimension, etc. were significantly different for normal and CAD subjects of the ECG signals [Antanavicius et al., (2008)]. Manis et al. (2007), Laitio et al. (2004) and Qtsuka et al. (2009) used correlation dimension and entropy features on heart rate signals to diagnose CAD. Karamanos et al. (2006) analyzed HR signals using DFA and showed that the self similarity nature of heart rate signals decreased in CAD subjects.
Table 5 Summary of studies conducted in automated detection of CAD and normal classes.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Base signal/Techniques</th>
<th>Classifiers</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karimi et al. (2005)</td>
<td>Heart sound, Wavelet analysis</td>
<td>Neural network</td>
<td>85%</td>
</tr>
<tr>
<td>Arafat et al. (2005)</td>
<td>ECG Stress Signals with Probabilistic Neural Networks</td>
<td>Fuzzy Inference Systems</td>
<td>80%</td>
</tr>
<tr>
<td>Lee et al. (2007)</td>
<td>HRV, Linear and Nonlinear Parameters</td>
<td>SVM Classifier</td>
<td>90%</td>
</tr>
<tr>
<td>Kim et al. (2007)</td>
<td>HRV, Multiple Discriminant Analysis with linear and nonlinear feature</td>
<td>Multiple Discriminant Analysis</td>
<td>75%</td>
</tr>
<tr>
<td>Zhao et al. (2008)</td>
<td>Diastolic murmurs, EMD-Teager Energy Operator</td>
<td>Back Propagation Neural Network</td>
<td>85%</td>
</tr>
<tr>
<td>Lee et al. (2008)</td>
<td>HRV, carotid arterial wall thickness</td>
<td>CPAR and SVM</td>
<td>85 - 90%</td>
</tr>
<tr>
<td>Babaoglu et al. (2010a)</td>
<td>EST-ECG, PSO+GA</td>
<td>SVM</td>
<td>81.46%</td>
</tr>
<tr>
<td>Babaoglu et al. (2010b)</td>
<td>EST-ECG, PCA</td>
<td>SVM</td>
<td>79.71%</td>
</tr>
<tr>
<td>Dua et al. (2012)</td>
<td>Nonlinear features +PCA</td>
<td>MLP</td>
<td>89.5%</td>
</tr>
<tr>
<td>Giri et al. (2012)</td>
<td>HR signals, ICA</td>
<td>GMM</td>
<td>96.8%</td>
</tr>
<tr>
<td>This work</td>
<td>HRV</td>
<td>No classification done</td>
<td></td>
</tr>
</tbody>
</table>
Table 5 shows the summary of the studies conducted for the automated diagnosis of CAD and normal HR signals. Karimi et al. (2005) presented classification result of 85% for CAD identification using the combination of artificial neural networks and wavelet features extracted from heart sounds. ECG stress signals combined with fuzzy and probabilistic methods were effectively used to detect CAD with an accuracy of 80% [Arafat et al., 2005]. Various linear and non-linear features were derived from heart rate signals in the left lateral, supine, and right lateral position [Lee et al., 2007]. In their work SVM yielded the highest accuracy of 90% compared to Bayesian classifiers, CMAR, and C4.5. The same group used the HRV features of different postures and carotid arterial wall thickness as features and classified the normal and CAD subjects with an accuracy of 85% to 90% using CAPAR and SVM classifier [Lee et al., 2008]. Classification was performed into control, angina pectoris and acute coronary syndrome using linear and nonlinear features of HR signals [Kim et al., 2007]. They reported an accuracy of 75%, and classified angina pectoris group with a sensitivity of 72.5% and specificity of 81.8%. Their system was able to classify people suffering from acute coronary syndrome with a sensitivity of 84.6% and specificity of 91.5%. Features extracted from heart murmurs using EMD – Teager energy operator automatically diagnosed normal and CAD subjects with an accuracy of 85% [Zhao et al., (2008)].

Binary Particle Swarm Optimization coupled with genetic algorithm applied on exercise data to detect the CAD yielded an accuracy of 81.4% using SVM classifier using twenty three features [Babaoglu et al., 2010a]. Same group reduced the twenty three features of the exercise stress test data to eighteen features and obtained an accuracy of 79.71% using SVM classifier (Babaoglu et al. (2010b)). Recently, Giri et al. (2012) classified normal and CAD classes using HR as base signal. Discrete wavelet transform (DWT) coefficients were subjected to data reduction using Independent Component Analysis (ICA). These ICA coefficients were classified using Gaussian Mixture Model (GMM) with an accuracy of 96.8%. The nonlinear features extracted from the HR signals were fed to
principal component analysis (PCA) for data reduction [Dua et al., 2012]. These PCA coefficients coupled with multilayer perceptron (MLP) method resulted in the highest classification accuracy (89.5%) to classify normal and CAD heart rate signals.

Unique ranges have been proposed for time, frequency and nonlinear features for normal and CAD HR signals. These extracted parameters can be used for automated detection of CAD using HR signals. The practical relevance of this study can be improved by using more diverse data from a wider range of subjects. It is risky to obtain the ECG signals during exercise from CAD affected subjects. Hence, signals like heart murmur, ECG stress signals, and HRV signals are more preferred to detect the normal and CAD classes.

The time domain analysis is not robust, due to the influence of artifacts and noise. The temporal information of the frequency content cannot be provided by the Fourier transform. The HR signal is a nonlinear signal and the information content in the higher harmonics of the signal can only be completely captured by nonlinear analysis methods. Hence, in this work, we evaluated the ranges of several nonlinear features extracted from normal and CAD affected subjects. We found that the RQA parameters, such as $L_{\text{mean}}$, $L_{\text{max}}$, REC and DET, were higher for CAD subjects indicating more rhythm. Since the activity of CAD subjects is less, similar signal patterns repeat or recur more frequently compared to the normal subjects. Hence, the parameter REC has higher value for CAD subjects. Similarly, the value of the determinism parameter or DET is higher for CAD subjects. This is again due to the fact that CAD subjects are less active than normal subjects. The same processes occur very frequently and thus it is easier to determine the HR signal. The entropy based parameters, ApEn and SampEn, are lower for CAD subjects indicating lower entropy (less activity due to impairment) for CAD. Almost all HOS parameters showed higher values for the CAD group, indicating the presence of higher frequency content in the CAD signals.

6. Conclusion

CAD is one of the prime reasons for the majority of cardiac deaths worldwide. In this work, we analyzed HR signals which were obtained from ECG data recorded from normal and CAD subjects. In our work, we have made an attempt to analyze both normal and CAD heart rate signals in time, frequency and non-linear domain. Our results show that HR signals are less variable in CAD subjects, compared to the normal subjects. We have
proposed unique ranges for features in various domains. Highly discriminative recurrence, Poincare, HOS plots have been presented to differentiate normal and CAD heart rate signals. These ranges of features and unique plots can be used in future to identify these two classes.

**Acknowledgements:** Authors thank Ms Ratna Yanti for running the codes and compiling the results and Thanjuuddin Ahmad for providing the data. HRV analysis Software, Biomedical Signal Analysis Group, University of Kuopio, Finland for providing the software.

**References**

5. U.R. Acharya, N. Kannathal, O. Sing, L. Ping, T. Chua, Heart rate analysis in normal subjects of various age groups, Biomedical Engineering Online 3 (1) (2004b) 24.


69. R. Wariar, C. Eswaran, Integer coefficient bandpass filter for the simultaneous removal of baseline wander, 50 and 100 Hz interference from the ECG, Medical & Biological Engineering & Computing 29(3) (1991) 333-336.
73. J.P. Zbilut, N. Thomasson, C. L. Webber, Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals, Medical Engineering & Physics 24 (2002) 53–60