FEATURE EXTRACTION, MODELING AND SYNTHESIS OF ECG FROM ARTERIAL BLOOD PRESSURE AND CENTRAL VENOUS PRESSURE SIGNALS BY SIGNAL PROCESSING TECHNIQUES

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR AWARD OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

AWADHESH PACHAURI

Registration Number 010 of 2013

**DEPARTMENT OF ELECTRONICS AND COMMUNICATION
ENGINEERING
SCHOOL OF ENGINEERING
TEZPUR UNIVERSITY
NAPAAM, TEZPUR, ASSAM
DECEMBER, 2014**

Dedicated to the sweet memory of my dear parents
Late Shri Mizazi Lal Pachauri & Late Smt. Siya Dulari

ABSTRACT

Over the past few decades, there has been an increasing demand for the development of sophisticated technologies in medical field to evaluate patient conditions. These technologies can assist the health care professionals to study patient condition in reasonable time and response to therapies more rapidly and with relative ease. Diseases in the human cardiovascular system are one of the main issues in modern health care. These cause the majority of deaths and also often impair people in their most productive age. According to the report of the World Health Organization (WHO), about 9.4 million deaths are caused every year in the world due to heart disease. Out of these deaths, 16.5% can be attributed to high blood pressure [1]. This includes 51% of deaths due to strokes and 45% of deaths due to coronary heart disease [2]. Some of the heart diseases require immediate attention when they are difficult to be identified at an early stage (such as premature ventricular contraction (PVC)) due to their occasional appearance. These problems faced by cardiologists have emerged as a great motivation to research in engineering. The study of relationship between electrical and mechanical events of cardiac cycle can greatly enhance the study of dynamics of heart beat [3].

The human physiological processes produce biomedical signals and phenomenon which describe their nature and actions. One of the most common biomedical signals is the Electrocardiogram (ECG) obtained in real time by non-invasive technique that reflects the cardiac activity. This signal comprises of repetitive sequence of P, Q, R, S and T waves, each representing a particular event or activity of the heart.

Although the electro-physiological phenomenon of the heart is well known and mostly assessed by cardiologists with the help of ECG, there are two other biomedical signals very closely associated with the blood circulatory system such as the Arterial Blood Pressure (ABP) and Central Venous Pressure (CVP) which provide a better understanding in critical and complex heart situations. Generally, blood pressure is measured with the help of a cuff sphygmomanometer which describe cardiac health in terms of systolic and diastolic values. Arterial blood pressure waveform is rich in the estimation of cardiovascular function as

compared to systolic and diastolic readings obtained by cuff sphygmomanometer. These readings may not be representative of the patient's usual blood pressure. Therefore, continuous waveform of blood pressure is useful to acquire more meaningful information about its range and variability [4]. Continuous arterial pressure monitoring not only provides information in regards to blood pressure, it also provides a means to assess the cardiovascular status by observing waveform characteristics since the waveform has close correlation with ECG .

ABP waveform is the combination of systolic peak, diastolic onset, dicrotic notch and dicrotic peak. If ABP and ECG waveforms are recorded simultaneously, systolic peak follows ECG R wave, diastolic onset follows ECG T wave and systolic peak to peak interval is equal to RR interval of ECG. There are physiological relations among the onset and offset of the above mentioned waves [5]. Abnormal condition of heart displayed in ECG signal is also present in ABP and CVP signals.

A combined investigation on ECG, ABP and CVP signals have prognostic significance as well as help to pinpoint the offending lesion when multiple obstructions are present and can thus be used to improve the treatment approach. Therefore, parallel analysis of ECG, ABP and CVP signals can be used to illustrate the cardiac signatures efficiently and leads to development of robust algorithms for the analysis of cardiac health.

Cardiologists look for life threatening disturbances in the intervals, amplitudes and areas of the waves recorded from the ECG. Long term ECG recording such as Holter ECG in ICU comprises of thousands of beats. Visual interpretation of the anomalies from the huge data by the experts is a tedious task and requires a lot of time in the interpretation of the ECG record. This requires development of signal processing techniques for feature extraction to reduce the workload of the cardiologists in detection of abnormalities such as cardiac arrhythmia etc. The errors occurred in the subjective interpretation of biomedical signals can be overcome by computer aided feature extraction techniques [6].

Although, ECG, ABP and CVP signals are believed to be faithful independent signatures of the cardiac system, many unseen and hidden correlating facts of the relevant signals can be revealed. The signals - ECG, ABP and CVP - comprise of a great deal of cardiac information that shows a high level of inter-relationship among these biomedical signals.

This inter-relationship motivates to generate one signal from other signal or from the combination of two or more simultaneously recorded signals available for a particular subject. When one of the physiological signals such as ECG under continuous observations is corrupted with noise, missing for certain duration or not possible to acquire (e.g surgical dressing of patients does not permit to position the sensors at the desired place for particular signal acquisition), the prediction of the desired signal using the combination of other available signals can suffice the requirement. Therefore, the ability of prediction of unavailable signals using computer based methods is an innovative idea in this field.

In the last decade, Various ECG feature extraction techniques have been developed so far which include wavelet transform [7-12] histogram and genetic algorithms [13], Artificial Neural Networks [14], moving average filter [15], differentiation and correlation [16] comprising of their own merits and demerits. Of these methods [7-16], the advanced signal processing methods using time-frequency analysis and filtering using wavelet transform have proved to be a very useful tool in determining the precise location of the QRS complex. Wavelet transform is popular because it satisfies energy conservation law and original signal can be reconstructed [8].

In these [7-12] wavelet based methods, ECG peak detection has been carried out by authors by selecting different detail signals. However, selection of detail signals in the above literature is not sufficiently justified by any of these authors. Most of the QRS detection algorithms mentioned above [15,11,12] are developed on ECG lead-II signals of MIT-BIH database as the QRS complexes are outstanding in lead II. In [17], author has included two ECG signals from leads V5/V2 but accuracy reported is 94% and 92% respectively.

Although a considerable amount of research has been carried out for feature extraction of ECG signals but there are very few algorithms reported for feature extraction of ABP signal and no algorithm is reported in literature for feature extraction of CVP signal. Most of the algorithms on ABP signal are developed on proprietary datasets from selected patients [19,20]. These methods are based on continuous independent assessment of refractory period (RP), analysis of signal by means of producing two moving averages [19], template matching [20], rank filter and decision logic [21], windowed and weighted Slope Sum Function (SSF) [22], peak and trough detection methods [23], heart rate, amplitude and

inter beat intervals [24] and combinatorial analysis of ABP waveforms and their derivatives [25].

The researchers have paid major attention on the detection of systolic peaks [24], onsets [18, 22], dicrotic notches [19] and three features (systolic peak, onset, dicrotic notch) together [25]. Dicrotic peak is also an important feature of ABP signal, however, no attempt has been made to detect dicrotic peak in ABP signal.

Similar to development of algorithm for feature extraction of ECG, ABP and CVP, development of techniques for modeling and synthesis of ECG is also a field of interest for researchers. Previous works on modeling and synthesis of ECG include a method for generation of RR-tachograms [26], a dynamical model that mimics the real ECG signal of a normal person [27], use of this model to generate realistic ECG, BP, respiration signals [28] and modeling an arbitrary ECG without in band noise [29] are also suggested. Further a three-dimensional dynamic model is proposed for ECG modeling that is generalized to model maternal and fetal ECGs [30]. Models for generation of multi-lead ECGs [31] and simulation of abnormal rhythms [32] are also reported in literature. Gaussian wave-based state space model is used for generating synthetic ECGs as well as separate characteristic waves (CWs) such as the atrial and ventricular complexes [33] and extended Kalman filter based dynamic algorithm is proposed for tracking the ECG characteristic waveforms [34]. Mathematical modeling of electrical activity of heart [35] and computer simulation of qualitative ECG are also suggested [36].

The inherent shape of Hermite Basis Functions (HBF) bearing resemblance to ECG signals is used for shape determination of ECG [37, 38]. In addition to this, piece wise modeling of ECG is also suggested [39]. A modified Zeeman model using radial basis network is proposed for ECG modeling but this model is able to generate single cycle of ECG [40]. Methods including Gaussian Combination Model [41], Hilbert transform [42], Hidden Markov models [43] and data flow graph method [44] are also known in literature for ECG modeling.

In physionet challenge 2010 [45], reconstruction of ECG, ABP, respiratory, fingertip plethysmogram (PLETH) signals using ANN [46-50] and wavelet based approach [51] has been suggested. ANN has been found to be a promising technique for ECG reconstruction

as suggested in [46-50], however the technique needs at least one ECG signal from other leads.

Although a considerable amount of research has been carried out for ECG modeling and synthesis, however, a parametric ECG model based on measured phenomenological cardiac data such as ABP and CVP has not been attempted so far. Synthesis of ECG from ABP and CVP signals is found to be a promising modeling approach in this research. We have used parameter estimation and system identification to develop the linear time invariant (LTI) model of the heart which can synthesize ECG from ABP and CVP.

Objectives of the proposed research

On the basis of the above literature review, the motivation towards development of technique for feature extraction and modeling of ECG, ABP and CVP is based on the following objectives –

1. Feature extraction of ECG signal by wavelet technique supported by signal energy, frequency spectrum and correlation analysis.

2. Feature extraction of ABP and CVP signals by wavelet technique supported by signal energy, frequency spectrum and correlation analysis.

3. Modeling and synthesis of ECG by system identification technique using measured physiological data of ABP and CVP signals.

4. Modeling and synthesis of ECG by Artificial Neural Network paradigm.

With the above quoted objectives, the research work is carried out as stated below –

Objective 1: ECG Feature Extraction

An algorithm for ECG feature extraction using wavelet (Daubechies) technique supported by signal energy, frequency spectrum and correlation analysis is proposed. Another algorithm for ECG peak detection using energy analysis technique is also proposed. The merit of both of these algorithms motivated us to apply these algorithms for the detection of life threatening heart disease - premature ventricular contraction (PVC) beats in ECG.

> **Methodology:** We present ECG peak detection by two methods - wavelet based method and energy analysis of ECG signal. The wavelet based approach described in

this objective is robust and simple to implement with no requirement of preprocessing. The selection of detail wavelet component has been justified by energy, frequency and correlation analysis. Since, there are wide variations in amplitudes of wavelet decomposed signals; a fixed threshold does not work for R peak detection. Therefore, we have adopted a 'window based threshold' where the threshold value is adjusted depending upon the signal amplitude over a certain duration. The selected detail signal is first thresholded then the maximum amplitude levels of all the peaks are detected. The signal is then filtered by applying a refractory period to select the R peaks. The R peaks detected by wavelet method are used for the detection of remaining features of ECG signal such as P, Q, S and T waves.

In the energy analysis technique for R wave detection, energy calculation of ECG signal under test has been performed by dividing the signal records into a number of windows. The techniques used include window shifted by window size and window shifted by one sample. Energy analysis of detail coefficients show that d4 signal containing highest energy content comprise of maximum information of QRS complexes. This concept motivates us to detect ECG peaks if the window based energy analysis of ECG signal is performed and the resultant energy signal is further analyzed using thresholding and refractory period concepts for detection of ECG peaks. Therefore, window based energy analysis of ECG signal may result in the higher energy amplitudes wherever ECG peaks exist.

The detected R peaks are applied to detect PVC beats in ECG. The method for detecting the abnormal PVC complexes is based on the calculation of RR interval of detected R peaks and energy analysis of ECG signal. We have proposed a combined method for PVC detection where, RR interval calculation by wavelet and energy is supported by intersection of energy analysis technique on the ECG signal. The algorithm proposed for PVC detection includes detection of R peaks using window based energy analysis of ECG signal using a window of 100 ms duration that incorporate window shift by one sample and further energy analysis of ECG signal using a window of 600 ms duration where window is shifted by window size.

Results:

Wavelet based ECG feature extraction method has been performed on 44 records of MIT-BIH arrhythmia database mostly from lead-II as well as from other leads each of 30 minutes duration. The overall sensitivity, positive predictivity and accuracy obtained are 99.62%, 99.87% and 99.48% respectively. The accuracies obtained for records 102 and 104 are 99.86% and 99.77% respectively which are higher than the result of [17] (94% and 92% respectively) where the author has included two ECG signals (records102 and 104) from other leads.

In energy based peak detection algorithms, five minute segments of each of the forty four records from MIT- BIH database have been tested for R peaks including records from other than lead II. The method achieves an accuracy, sensitivity and positive predictivity 98.17%, 98.82%, 99.36% respectively using window shifted by window size and 98.63%, 99.36%,99.28% respectively using window shifted by one sample.

The algorithm for PVC detection is applied on 37 records of the database. The accuracy, sensitivity and positive predictivity reported for PVC detection are 96.79%, 98.31% and 98.48% respectively.

Objective 2 : ABP and CVP Feature Extraction

This part of research enumerates (i) Extraction of all features of ABP and CVP signals and (ii) energy based approach for peak detection of ABP signal. The merit of wavelet based method applied for ECG feature extraction has motivated us to apply this technique for the analysis of ABP and CVP signal.

Methodology:

The wavelets used for decomposition of ABP signal are daubechies db4 and symmetric sym4. Selection of detail coefficients after wavelet decomposition has been justified by energy, frequency and correlation analysis of detail coefficients. Further, it is found that application of 'window based threshold' overcomes the setback of missing peaks due to large variations in the signal amplitude at any particular instant.

In this objective, an algorithm for the automatic detection of systolic peaks in the ABP signal using energy analysis is also proposed. The ABP signal under test undergoes window based energy analysis by selecting a window of 100 ms duration. Similar to ECG peak detection technique, the areas in ABP signal where systolic peaks are available appear as high energy zones. Window based amplitude threshold and interval threshold are applied to reject the unwanted peaks.

The detection of features in CVP signal is carried out using db4 wavelet and selection of relevant detail coefficient has been validated based on energy, frequency and correlation technique. Here, negative amplitude thresholding is used.

Results:

ABP feature extraction

The algorithm of ABP waveform delineation using wavelet technique has been applied on 1 minute segment of 22 signals of MGH/MF waveform database, 14 signals of Fantasia database, 15 signals of MIT-BIH polysmographic database and 1 signal of CSL database. The performance of the algorithm is given in terms of percentage of accuracy (Ac), sensitivity (Se) and positive predictivity (PP) and error.

In terms of type of wavelet used - sym4 wavelet is found to result better accuracy and positive predictivity as compared to db4 for all the four components of ABP signal, however the sensitivity in both cases are comparable.

In terms of the database used for the analysis-accuracy and sensitivity is found to be highest in MGH/MF database while MIT-BIH polysmographic database outperforms the other database in positive predictivity. Performance of both the wavelets for all the features on CSL database is found to be better than other three database since we have used only one signal from CSL database.

The energy based peak detection algorithm has been tested on 9 records of MGH/MF waveform database. The algorithm reported an accuracy of 99.53%, sensitivity of 99.98% and positive predictive value of 98.14%.

CVP Feature Extraction

CVP feature extraction has been performed on selected segments of two records in which all the features are available and can be annotated. The algorithm reported highest values in accuracy of 85.07%, sensitivity of 93.95% and positive predictivity of 91.70%.

Objective 3 : Modeling and synthesis of ECG using system identification technique

This part of research aims at system identification based modeling and synthesis of ECG for 7 records – mgh003, mgh004, mgh005, mgh007, mgh008, mgh029 and mgh031 of MGH/MF waveform database where three signals (ABP, CVP and ECG) are available for system identification. We attempt to model the cardiac system of both healthy subjects (i.e. normal) and having PVC and SVPB (Supraventricular premature beat i.e. abnormal) and then we simulate the models using cross-validation inputs (ABP, CVP) i.e.

   a) Normal ABP, CVP inputs to normal model

   b) Normal ABP, CVP inputs to abnormal model

   c) Abnormal ABP, CVP inputs to abnormal model

Methodology:

We present a system identification based approach for modeling of ECG using ABP and CVP signals in autoregressive models and state space models using prediction error minimization (PEM) and subspace algorithms. The inputs to the model are simultaneously acquired ABP and CVP signals and synthesized output is ECG signal. The input and output signals of the model are preprocessed before testing and validation of the model. One cycle of ABP, CVP and ECG from 7 records are used for estimating the state space model. The model generates two transfer functions related to each input signal. Initially, a higher order model is obtained which is reduced to a significant order.

The accuracy of the synthesized and original ECG is evaluated by a universally accepted metric given by 'best fit' percentage. When a model is developed with normal ABP/CVP of a normal heart, and validated using normal data, a better fit percentage is obtained, while the best fit percentage goes down if the model is validated by an abnormal data. An interesting inference can be drawn to relate the condition of heart (i.e. a normal or having abnormal PVC/ SVPB) to the stability of the model transfer function. A stable transfer function is obtained if the ECG cycles of the model are normal, whereas an unstable transfer function is obtained from an abnormal ECG cycle. Therefore, the generated transfer functions are analyzed for stability using pole zero plots and step responses.

Results

By using subspace method of identification in state space model, we have achieved a maximum best fit percentage of 80.4852% for normal ECG model validated by normal ABP & CVP data of mgh008 record, while a maximum fit percentage of 75.5584% was achieved when an abnormal ECG model (having SVPB) validated by ABP and CVP data having SVPB in mgh004 record.

Moreover, the state space modeling by PEM method, the technique reported a highest best fit percentage of 74.58% when a normal ECG model was validated by normal data in mgh029 signal.

In case of autoregressive modeling, we have achieved a maximum of 76.2032% fit percentage while validating a normal ECG model by normal data of mgh031 signal.

Further, comparative study of all the three models shows that subspace method of system identification results in higher fit percentage for all the seven records. Therefore, it is concluded that the system generates a stable transfer function when model is simulated using normal data, while an unstable transfer function is generated when simulated using abnormal (PVC and SVPB) beats.

Objective 4: Modeling and synthesis of ECG using artificial neural network (ANN)

The ANNs are found to be most successful modeling technique adopting the capability of non-linear learning and multi-dimensional mapping. Non-linear connectivity between the pressure signals (ABP & CVP) and the ECG signals is utilized in this objective to model and synthesize ECG using ANN.

Methodology

The proposed technique utilizes one of the most popular ANN algorithm – Radial basis function (RBF) to map the synchronously sampled pressure signals (ABP & CVP) as input and the corresponding ECG as the output. Normalization is one of the most important step in ANN modeling where the real engineering units are normalized in the scale of -1 to +1. We have done ANN modeling in the following two phases –

i)   On 16 records each comprising 20,000 samples of ABP, CVP and ECG signals of MGH/MF waveform database to model and synthesize the ECG from which R peaks are detected. The RBF structure used in this case has *spread constant* of 110 and number of neurons 250.

ii) On 16 records each comprising 1,00,000 (85,000 for training and 15,000 for testing) samples of MGH/MF waveform database to model and synthesize the ECG which was used for comparison with original ECG by 'cosine measure' and 'cross-correlation' techniques. In this approach we have used RBF structure of *spread constant* of 10 and number of neurons 500.

Results

In our first approach of ANN modeling with 20,000 samples, the numbers of peaks obtained in the synthesized ECG are compared with expert annotations for ECG peaks available from the database. The overall accuracy, sensitivity and positive predictivity for 16 records are 95.96%, 97.05% and 98.99% respectively.

In the second approach on 1,00,000 samples for modeling of ECG, the suynthesized ECG signals are tested by cosine measure and cross-correlation for resemblance with original ECG signals. The highest cosine measure and cross-correlation achieved are 0.9493 and 0.9498 in mgh006.

Thesis outline

Chapter 1 describes the background and introduction to the biomedical signals of cardiac system and their detection techniques and review of literature as the basis of objectives achieved in this research. Chapter 2 presents algorithms for ECG feature extraction using wavelet and energy analysis techniques with justification of selection of the wavelet detail coefficients. In this chapter application of wavelet and energy technique to PVC detection is also proposed. Chapter 3 illustrates the application of a wavelet and energy analysis based technique for ABP and CVP feature extraction. Chapter 4 presents system identification based approach for modeling and synthesis of ECG using measured ABP and CVP signals for normal and abnormal conditions of the heart. The stability analysis of the developed models is also explained using pole zero plots and step response. In chapter 5, a technique using ANN for modeling and synthesis of ECG using measured ABP and CVP signals has been described.

References

[1] Lim, S.S, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet*, **380**(9859), 2224–2260, 2012.

[2] World Health Organization, The global burden of disease: 2004 update, WHO Library Cataloguing-in-Publication Data Geneva,2008, http://www.who.int/ .

[3] Coblentz, B. et al. The relationship between electrical and mechanical events in the cardiac cycle of man, *J. Br Heart*, **11**, 1-22, 1949.

[4] Kain, H.K., et al. Arterial blood pressure measurements with a portable recorder in hypertensive patients: I. variability and correlation with casual pressures, *Circulation*, **30**, 882-892, 1964.

[5] Morgan, B.L., Hemodynamic waveforms interpretation, *Critical care concepts*, 2005.

[6] Rute Almeida, R., et al. Multilead ECG delineation using spatially projected leads from wavelet transform Loops, *IEEE Transactions on Biomedical Engineering*, **56**(8), 1996-2005, 2009

[7] Cuiwei, Li., et al. Detection of ECG characteristic points using wavelet transforms, *IEEE Trans. Biomed. Eng*, **42**(1), 1995.

[8] Sahambi, J.S., et al. Using wavelet transforms for ECG characterization, an online digital signal processing system, *IEEE Engineering in Medicine & Biology*, 77-83, 1997.

[9] Saxena, S.C., et al. QRS detection using new wavelets, *J. of Medical Engineering & Technology*, **26**, 7-15, 2002.

[10] Legarreta, I.R., et al. A comparison of continuous wavelet transform and modulus maxima analysis of characteristic ECG features, *IEEE Computers in Cardiology*, **32**, 755–758, 2005.

[11] Mahmoodabadi, S.Z., Ahamadian, A., Abolhasani, M.D., Eslami M. & Bidgoli, J.H. ECG feature extraction based on multiresolution wavelet transform, Proceedings of IEEE, Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 2005.

[12] Rudnicki, M. & Strumillo, P., A real-time adaptive wavelet transform - based QRS complex detector, ICANNGA 2007, Part II, LNCS 4432, Springer-Verlag Berlin Heidelberg, 281–289, 2007.

[13] Tu, C., et al., A new approach to detect QRS complexes based on a histogram and genetic algorithm, *J. Medical Engineering & Technology*, **29** (4), 176-180, 2005.

[14] Ghongade, R., & Ghatol, A.A. A brief performance evaluation of ECG feature extraction techniques for artificial neural network based classification, TENCON 2007, 1-4, doi:10.1109/TENCON.2007.4429096.

[15] Chen, H.C. & Chen, S.W. A moving average based filtering system with its application to real-time QRS detection, *Computers in Cardiology*, **30**: 585-588, 2003, doi:10.1109/CIC.2003.1291223.

[16] Wrublewski, T.A., Ying, S. & Beyer, J.A. Real-time early detection of R-waves of ECG Signals, ECG signal processing III, IEEE Engineering in Medicine & Biology Society, 11th Annual International Conference 1989.

[17]Lee, R.G., et al. A novel QRS detection algorithm applied to the analysis for heart rate variability of patients with sleep apnea, *Biomedical Engineering Applications, Basis & Communications*, **17**(5)258 – 262, 2005.

[18]Zong, W., et al. Effects of vasoactive drugs on the relationship between ECG-pulse wave delay time and arterial blood pressure in ICU patients, *Computers in Cardiology*, **25**, 673-676, 1998.

[19]Navakatikyan, M.A., et al. A real-time algorithm for the quantification of blood pressure waveforms, *IEEE Transactions on Biomedical Engineering*, **49** (7), 2002.

[20]Hocksel, S.A.A.P. et al. Detection of dicrotic notch in arterial pressure signals, *J. Clinical Monitoring*, **13**, 309-316, 1997.

[21]Aboy, M., Crespo, C., McNames, J., & Goldstein, B. Automatic detection algorithm for physiologic pressure signal components, *Proceedings of Second Joint EMBS/BMES Conference* (Houseton, TX USA, 2002), 196-197.

[22]Zong, W., et al. An open source algorithm to detect onset of arterial waveform pulses, *Computers in Cardiology*, **30**, 259-262, 2003.

[23]Palaniappan, R., & Krishnan, S.M., Detection of ectopic heart beats using ECG and blood pressure signals, *International Conference on Signal Processing & Communications* (SPCOM' 2004).

[24]Aboy, M., et al. An automatic beat detection algorithm for pressure signals, *IEEE Transactions on Biomedical Engineering*, **52** (10), 2005.

[25]Li, B.N., et al. On an automatic delineator for arterial blood pressure waveforms, *Biomedical Signal Processing and Control*, **5**, 76–81, 2010.

[26]McSharry, P.E., et al. Method for generating an artificial RR tachogram of a typical healthy human over 24-hours, *Computers in Cardiology*, **29**, 225– 228, 2002, doi: 10.1109/CIC.2002.1166748.

[27]McSharry, P.E., et al. A dynamical model for generating synthetic electrocardiogram signals, *IEEE Trans Biomed Eng* **50** (3), 289–294, 2003.

[28]Clifford, G.D., & McSharry, P.E. A realistic coupled nonlinear artificial ECG, BP, and respiratory signal generator for assessing noise performance of biomedical signal processing algorithms, *Proc of SPIE International Symposium on Fluctuations and Noise* (2004), **34**, 290–301.

[29]Clifford, G.D., et al. Model-based filtering, compression and classification of the ECG, *Int. J. Bioelectromagn,* 7, 158–161, 2005.

[30]Sameni, R., et al. Multichannel ECG and noise modeling: application to maternal and fetal ECG signals, *EURASIP J. Adv. Signal Process.* 1-14, 2007, doi:10.1155/2007/43407.

[31]Clifford G.D., et al. An artificial multi-channel model for generating abnormal electrocardiographic rhythms, *Computers in Cardiology,* 35, 773-776, 2008.

[32]Clifford G.D., et al. An artificial vector model for generating abnormal electrocardiographic rhythms, *Physiol. Meas.,* 31, 595–609, 2010.

[33]Sayadi, O., et al. Synthetic ECG generation and Bayesian filtering using a Gaussian wave-based dynamical model, *Physiological Measurement,* 31, 1309–1329, 2010.

[34]Sayadi, O., et al. Life-threatening arrhythmia verification in ICU patients using the joint cardiovascular dynamical model and a Bayesian filter, *IEEE Transactions On Biomedical Engineering,* 58 (10), 2748-2757, 2011.

[35]Plonsey, R. & Barr, R.C. Mathematical modeling of electrical activity of the heart, *J. Electrocardiology,* 20, 219—226, 1987.

[36]Wang, J.T. et al. A knowledge-based system for qualitative ECG simulation and ECG analysis, 773-776, 1992, IEEE.

[37]Rasiah, A.I, Togneri, R. & Attikiouzel, Y. Modeling 1-D signals using hermite basis functions, *IEE Proceedings on Vision, Image, and Signal Processing,* 144 (6), 345-354, 1997.

[38]Dixit, P., Segmented modeling of ECG signal by using hermite basis function, *International J. Engineering and Mathematical Sciences (IJEMS),* 1 (1), 40-46, 2012, www.ijems.org.

[39]Ahmadian, A., Karimifard, S., Sadoughi, H., & Abdoli, M. An efficient piecewise modeling of ECG signals based on Hermitian basis functions, *Proceedings of the 29th Annual International Conference of the IEEE EMBS (2007),* Lyon, France, 3180-3183.

[40]Ayatollahi, A., et al. A comprehensive model using modified Zeeman model for generating ECG signals, *Iranian J. Electrical & Electronic Engineering,* 1(2), 88-93, 2005.

[41]Parvaneh, S., & Pashna, M. Electrocardiogram synthesis using a Gaussian combination model (GCM), *Computers in Cardiology*, **34**, 621-624, 2007.

[42]Nunes, J.C. & Amine, N.A. Hilbert transform based ECG modeling, Biomedical Engineering, Springer, **39** (3), 133-137, 2005.

[43]Andreao, R.V. et al. "ECG signal analysis through Hidden Markov Models, IEEE *Transactions on Biomedical Engineering*, **53** (8), 1541-1549, 2006.

[44]Zheying, Li., & Minjie, M., ECG modeling with DFG, 27th Annual Conference of IEEE Engineering in Medicine and Biology, 2691-2694, 2005.

[45]Moody, G.B., The physionet/computing in cardiology challenge 2010: mind the gap, *Computing in Cardiology*, **37**, 305-309, 2010.

[46]Rodrigues, R., Filling in the gap: a general method using neural networks, *Computing in Cardiology*, **37**, 453-456, 2010.

[47]Thomas, C., et al. An approach to reconstruct lost cardiac signals using pattern matching and neural networks via related cardiac information, *Computing in Cardiology*, **37**, 441–444, 2010.

[48]Silva, L., et al. Medical multivariate signal reconstruction using recurrent neural network, Computing in Cardiology, **37**, 445–447, 2010.

[49]Sullivan, A.M., et al. Reconstruction of missing physiological signals using artificial neural networks, *Computing in Cardiology*, **37**, 317–320, 2010.

[50]McBride, J., et al. Reconstruction of physiological signals using iterative retraining and accumulated averaging of neural network models, *IOP Publishing Physiol. Meas.*, **32**, 661-675, 2011.

[51]Wei, W., The multi-parameter physiologic signal reconstruction by means of wavelet singularity detection and signal correlation, *Computing in Cardiology*, **37**, 457–459, 2010.

Contribution from this thesis:

[1] Pachauri, A., & Bhuyan, M. Robust detection of R-wave using wavelet technique, in International Conference on Signal Processing, Communication and Networking (ICSPCN' 2009), Singapore, 335-339, World Academy of Science, Engineering and Technology **56**, 2009, 901-905.

[2] Pachauri, A., & Bhuyan, M. Wavelet and energy based approach for PVC detection, in IEEE International Conference on Emerging Trends in Electronic and Photonic Devices & Systems (ELECTRO' 2009), Varanasi, India, 258-261.

[3] Pachauri, A., & Bhuyan, M. PVC detection by energy analysis", in 2nd International Conference on RF and Signal Processing Systems (RSPS' 2010), Guntur, India, 380-384.

[4] Pachauri, A., & Bhuyan, M. A new approach to ECG peak detection, in International Conference on Biomedical Engineering and Assistive Technologies (BEATS' 2010), Jalandhar, Punjab, India.

[5] Pachauri, A., & Bhuyan, M. ABP peak detection using energy analysis technique, IEEE International Conference on Multi-media, Signal Processing and Communication Technologies (IMPACT' 2011), AMU, Aligarh, India.

[6] Pachauri, A., & Bhuyan, M. Wavelet transform based arterial blood pressure waveform delineator, *International Journal of Biology and Biomedical Engineering*, 1 (6), 15-25, 2012.

[7] Pachauri, A., & Bhuyan, M. Modeling of ECG using ABP and CVP signals: A system identification based approach, *International Journal of Engineering, Science and Innovative Technology*, 2(6), 321 -330, 2013.

[8] Pachauri, A., & Bhuyan, M. Modeling of ECG from arterial blood pressure and central venous pressure using artificial neural network, in IEEE International conference on recent advances and innovations in engineering (ICRAIE' 2014), Jaipur, India.

Contribution under review:

[1] Pachauri, A., & Bhuyan, M. System identification based modeling and synthesis of electrocardiograms, *International Journal of Modeling and Simulation*.

DECLARATION

I do hereby declare that the thesis entitled, **"Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques"** being submitted to the Department of Electronics and Communication Engineering, School of Engineering, Tezpur (Central) University is a record of original research work carried out by me. All sources of assistance have been assigned due acknowledgement. I also declare that neither this work as a whole nor a part of it has been submitted to any other university or Institute for award of any other degree or diploma.

Place: Tezpur

Date: 17·12·14

(Awadhesh Pachauri)

TEZPUR UNIVERSITY

(A Central University established by an Act of Parliament)
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
SCHOOL OF ENGINEERING
NAPAAM, TEZPUR-784028
DISTRICT: SONITPUR: ASSAM: INDIA

CERTIFICATE BY THE SUPERVISOR

This is to certify that the thesis entitled **"Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques"** submitted to the School of Engineering, Tezpur University in partial fulfillment for the award of the degree of Doctor of Philosophy in **Electronics and Communication Engineering,** is a record of research work carried out by **Mr. Awadhesh Pachauri** under my supervision and guidance.

All help received by him from various sources have been duly acknowledged.

No part of this thesis has been submitted elsewhere for award of any other degree.

Date: 17-12-14

(M Bhuyan)
Professor
Supervisor

TEZPUR UNIVERSITY

(A Central University established by an Act of Parliament)
NAPAAM, TEZPUR-784028
DISTRICT: SONITPUR: ASSAM: INDIA

<u>CERTIFICATE</u>

This is to certify that the thesis entitled **"Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques"** submitted to Tezpur (Central) University in the Department of Electronics and Communication Engineering under the School of Engineering in partial fulfillment for award of the degree of Doctor of Philosophy by **Awadhesh Pachauri** has been examined by me on 0 2-01- 15 and found satisfactory.

The committee recommends for award of the degree of Doctor of Philosophy.

Supervisor

Date: 02-01-15

External Examiner

Date: 0 2.01.15

ACKNOWLEDGEMENTS

I take the advantage of this excellent opportunity to express my deep feelings and gratitude to those who stood with me all the time to help me and overcome the numerous difficulties during this doctoral research work.

My first sense of sincere gratitude goes to my supervisor Professor Manabendra Bhuyan for his invaluable guidance and moral support. It is my greatest fortune to have him as my supervisor. I am grateful for his insightful inspiration and constant encouragement. I am lucky to work under such a devoted educator and top-grade researcher. Each minute of his valuable time, pieces of advice imparted, his extensive help, a great belief in me and the liberty to pursue the work freely almost as I pleased, made me to come up with flying colours. It is his excellent guidance that makes the foundation of successful completion of research work in time. It will not be overstated to mention that I would not have been able to achieve my goal without his insightful inspiration, invaluable guidance and moral support. The concept of modeling was new to me and very overwhelming. However, Prof. M Bhuyan encouraged me all the through constantly to work hard to achieve my goal. I have learnt a lot of things from him. I shall ever remain grateful to him.

I am grateful to DRC chairman Prof. J.C. Dutta and DRC committee members - Prof. PP Sahu Dept. of ECE and Dr. S.M. Hazarika Dept. of CSE for their interest and comments during the entire duration of research. Their encouraging guidance provided me with better knowledge and sense of understanding.

I am thankful to faculty members - Prof S. Bhattacharya, Dr. S. Sarmah, Dr. S. Roy, Dr. B. Deka, Dr. V.K Nath, Dr. N.M Kakoty, Mr. Riku Chutia, Mrs. D. Hazarika, , Mr. B Mondol, Mr. R.K. Baruah, Mr. D Sonowal, Ms. Priyanka Kakoty, Ms. Ananya Bonjyotna and Lab supervisor Mr. Ankul C. Baishya for their valuable help and cooperation.

I am thankful to research scholars - Mr. Madhuriya Pratim Das, Mr. Bidyut Deka, Ms. Aradhana Dutta, Ms. Nimisha Dutta, Mr. Anup Kumar Bordoloi, Dr. Mahipal Singh, Dr. Mausumi Barthakur, Ms. Nilima Gogoi, Mr. Pranjal Borah, Ms. Shashikala Kalita, Mr. Hemanga Kumar Das, Mr. Abdul Barik, Ms. Soma Chakraborty, Ms. Rashmi Deka, Ms. Kuntala Boruah, Mr. Pranjal Barman, Mr. Anil Hazarika and Mr. Lachit Dutta for those enjoyable moments which

they have been so generous to share with me. I take pleasure to express my sincere thanks to my roommate Mr. Anup Kumar Bordoloi for supporting me most of the time and keeping a fun full environment during my stay in the University. I am thankful to Mr. Mukut Senapati, research fellow for his cooperation. My sincere thanks to Mr. Champak Talukdar, research fellow, as his presence adds a lot of pleasure to my life in Tezpur University.

I am thankful to Mr. Dwipendra Chandra Das, Mr. Durlav Mahanta, Mr. Noor Ali in ECE Dept office, Mr. Bubul Das (now in Dept of Chemical Science) and technical staff Mr. Kulendra Sharma, Mr. Rupak Sharma and Ms. Sayanika of Dept. of ECE for their consistent help.

I am thankful to Tezpur University library for the wealth of books and journals which I referred for invaluable support of literature review. I am also thankful to Physionet database, www.google.com and the authors whose publications are accessible from the web which make the successful completion of this research.

My special thanks to Mr. George B. Moody, Institute for Medical Engineering and Science Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, USA for helping me instantly and resolving my problems related to Physionet database used in this research.

I am thankful to my employer - Indian Air Force for providing me right atmosphere, which helped me achieving my goals. I am also thankful to my colleagues – Mr. Chandan Kumar Sah, Mr. Amit Pathak, Mr. Krishan Saurabh, Mr. Arun, Mr. Pawan Sen, Mr. Abhishek Kumar Gupta of my section from Air Force Station Halwara, Ludhiana, Punjab for their consistent help, encouragement and supporting me to give more time for the research.

I am grateful to Mr. Elder brother – Shri Ram Gopal Pachauri, Advocate who has always guided me to proceed in right direction since childhood. My ultimate thanks to my wife Mrs. Sapna and sons - Abhay and Anant for their love, patience and understanding provided during these years, I am thankful to them for their great support in difficult moments. They deserve more credit than I can ever give them for spending their time alone as most of the time, I am engaged either in duty or research work. Their contribution is crucial in achieving this doctoral research.

Last but not the least, my sincere thanks to all those who helped and guided me directly or indirectly.
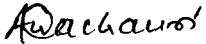
Date: 17·12·14   (Awadhesh Pachauri)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Figures

model and (G7, G8) of abnormal (SVPB) validation data of
mgh004 record

LIST OF ABBREVIATIONS

ABP	Arterial Blood Pressure
AF	Atrial fibrillation
ARMAX	Autoregressive moving average model with exogenous inputs
ARX	Autoregressive
AV	Atrioventricular
bpm	Beats per minute
CVP	Central venous pressure
CWT	Continuous wavelet transform
DBP	Diastolic blood pressure
DWT	Discrete wavelet transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electrooculogram
FFT	Fast Fourier Transform
ICP	Intra cranial pressure
LBBB	Left bundle branch block
LMS	Least mean square
LSE	Least square error
LTI	Linear time invariant

MA	Moving average
MAP	Mean arterial pressure
MI	Myocardial ischemia
PAC	Premature atrial contraction
PAP	Pulmonary artery pressure
PEM	Prediction error minimization
PLETH	Plethysmogram
POX	Pulse oxymetry
PVC	Premature ventricular contraction
QMF	Quadrature mirror filters
RBBB	Right bundle branch block
SA	Sino-atrial
SBP	Systolic blood pressure
SV	Stroke volume
SVPB	Supraventricular premature beat
VF	Ventricular fibrillation
VPB	Ventricular premature beat
VT	Ventricular tachycardia
WT	Wavelet transform

CHAPTER 1

Introduction

1.0. Introduction

Over the past few decades, there has been an increasing demand for the development of sophisticated technologies in medical field to evaluate patient conditions. These technologies can assist the health care professionals to study patient condition in reasonable time and response to therapies more rapidly and with relative ease. Diseases of the human cardiovascular system are one of the main issues in modern health care. These cause the majority of deaths and also often impair people in their most productive age. According to the report of World Health Organization (WHO), about 9.4 million people die every year due to heart disease. Of these deaths, 16.5% can be attributed to high blood pressure [1]. This includes 51% of deaths due to strokes and 45% of deaths due to coronary heart disease [2].

The first IEEE Life Sciences Grand Challenges Conference was held on October 4-5, 2012 at the National Academy of Sciences in Washington DC. The conference was sponsored by the National Science Foundation (NSF) at the Institute for Engineering and Medicine of the University of Minnesota, and was endorsed by the International Academy of Medical and Biological Engineering. The keynote lecture by Nobel laureate Dr. Phillip Sharp articulated the view that advances in information technology are required for the realtime acquisition, storage and processing of large, often massively parallel datasets, more complete understanding of normal and disease conditions through quantitative models, which in turn could lead to better informed and engaged patients [3].

Some of the heart diseases such as cardiac arrhythmia etc. can be silent killer and require immediate attention after considerable change is noticed in the signals associated to heart activity and symptoms of the patient. It puts forth a challenge to the cardiologists to cure

the patient at an advance stage when the primary episodes of arrhythmia is difficult to be identified at an early stage such as premature ventricular contraction (PVC)) due to their occasional appearance. These problems faced by cardiologists have emerged as a great motivation to research in engineering.

The prevalence of chronic diseases such as heart disease and stroke are the major cause of death in almost all countries, and such cardiovascular diseases present challenging problems in early diagnosis and treatment, repair or replacement. Engineering has played a vital role in understanding the cardiac electrodynamics, elastomechanics and in the development of diagnostic instruments, prosthetic valves, pacemakers, implantable cardioverters/ defibrillators (ICD) and automated external defibrillators (AED). Of all these diagnostic methods for cardiovascular diseases, the great need is for a comprehensive, computational and systematic approach to diagnose and describe the cardiovascular system [3].

These approaches comprise of the comprehensive information of many experts pertaining to a specific diagnosis. These methods can help to develop a robust and precise method for the diagnosis of a particular cardiovascular disease using cardiac signals along with the opinion of medical specialists.

The study of relationship between electrical and mechanical events of cardiac cycle can greatly enhance the study of dynamics of heart beat [4]. As a result, it is possible to carry out complete hemodynamic assessment with relative ease at the patient's bedside [5].

In intensive care units (ICU), where, continuously monitoring of physiological signals is required, there are various cardiac conditions and features which need to be detected on a large data of signal record and validation of abnormal ECG alarms using other cardiac signals. Situation may arise when actual ECG signal is missed or corrupted due to malfunctioning in sensors or due to external disruption. These disruptions in actual signals result in a great difficulty for precise diagnosis. At times, acquisition of ECG may not be possible due to surgical dressing of patients. So the synthesis of ECG using ABP and CVP signals can be used to supplement the information when actual ECG is either missing or corrupted. Also, a single lead ECG information need to be expanded to multilead

information using the information derived from other cardiovascular signals such as ABP, CVP, pulmonary artery pressure (PAP) etc.

1.1. Human cardiovascular system

The cardiovascular system is resulting from the motion of blood. It consists of the heart, blood and blood vessels. The function of the cardiovascular system is to circulate blood throughout the body's tissues so it can deliver certain substances to cells and remove other substances from them. The heart pumps the blood throughout the body through three types of flexible tubes called blood vessels viz arteries, capillaries and veins. The arteries pass the blood from the heart to all parts of the body. Then the blood circulates through the narrow blood vessels called capillaries. Many substances and some blood cells mixed with other substances pass into and out of the blood by moving through the thin porous capillary walls. The blood is then conceded through the veins, which return the blood to the heart.

1.2. The heart

The heart is the central organ of the cardiovascular system. It is a hollow, cone shaped muscular organ, little larger than fist size that contracts at regular intervals and thus forces the blood through the circulatory system. On an average a heart beats 100,000 times and pumps about 2,000 gallons of blood each day. The heart is enclosed in a protective membrane like sac called the pericardium, which surrounds the heart and secretes a fluid that reduces friction as the heart beats. The heart is divided in two sides by a wall, known as the septum. The left side of the heart pumps blood rich of oxygen and nutrients to all parts of the body and right side of the heart pumps blood depleted of oxygen and nutrients to lungs for oxygenation. The septum prevents the mixing of oxygenated and deoxygenated blood. The circulating blood brings oxygen and nutrients to all the body's organs and tissues, including the heart itself. It also picks up waste products from the body's cells. These waste products are removed as they are filtered through the kidneys, liver and lungs [6].

1.2.1. The heart's structure

The heart's structure as shown in Figure 1.1 consists of four chambers for pumping of the blood throughout the body. The upper two chambers are called left and right atria and lower

two chambers are called left and right ventricles. Unidirectional flow of blood is maintained with opening and closing actions of four valves of the heart. These valves are located within the openings leading from the atria to the ventricles and from the ventricles to the arteries. The positions of the valves in the heart are –

    i.   The tricuspid valve is between the right atrium and right ventricle.

    ii.   The pulmonary or pulmonic valve is between the right ventricle and the pulmonary artery.

    iii.   The mitral valve is between the left atrium and left ventricle.

    iv.   The aortic valve is between the left ventricle and the aorta.

Each valve has a set of flaps (also called leaflets or cusps). All the valves of the heart have three flaps except mitral valve which has two flaps. Under normal conditions, the valves allow blood flow in just one direction. Blood flow occurs due to building up of the pressure gradient across the valves that cause them to open.

The heart's wall is composed of three layers: the *endocardium*, *myocardium*, and *epicardium*. The inner lining of the heart is called *endocardium*. This layer is smooth and has no gaps that allow blood to contact the underlying collagen. The smoothness of this layer prevents blood from clotting and thus prevents the movement of clots into arteries and block blood flow. The middle thick layer is myocardium that constitutes most of the wall of the heart. The *myocardium* consists mostly of heart muscle (cardiac muscle), though it may also contain fat tissue and collagen fibers. The outer layer of the heart is the *epicardium* that contains some connective tissue coated with a smooth, slippery layer of epithelial cells. This coating allows the beating heart to move easily within the pericardial cavity.

1.2.2. Functioning of the heart

The heart pumps blood by a highly organized sequence of contractions of its four chambers. The right atrium receives partially deoxygenated and nutrient depleted blood from the veins. Venous blood is darker than arterial blood because of the difference in dissolved gases such as carbon dioxide (CO_2) which is absorbed from body's tissues. When the heart is relaxed, venous blood flows through the open tricuspid valve to fill the right ventricle. An electrical signal starts the heartbeat thus causing the atria to contract. This

contraction "tops off" the filling of the ventricle. Shortly after the atrium contracts, the right ventricle contracts. As this occurs, the tricuspid valve closes and the partially deoxygenated blood is pumped through the pulmonary valve, into the pulmonary artery and on to the lungs for oxygenation. This newly oxygenated blood is bright red. At the same time when the right atrium contracts, the left atrium also contracts, topping off the flow of oxygenated blood through the mitral valve and into the left ventricle. Then a split second later the left ventricle contracts, pumping the blood through the aortic valve, into the aorta and on to the body's tissues.

The organized contraction of all the four chambers is governed by an electrical impulse called action potential. Movement of this action potential across a heart chamber results in its contraction. This action potential is generated in the *sinoatrial node (SA node)* which is a small bundle of highly specialized cells in the right atrium. The SA node is called the natural "pacemaker" that generates action potential at regular intervals.
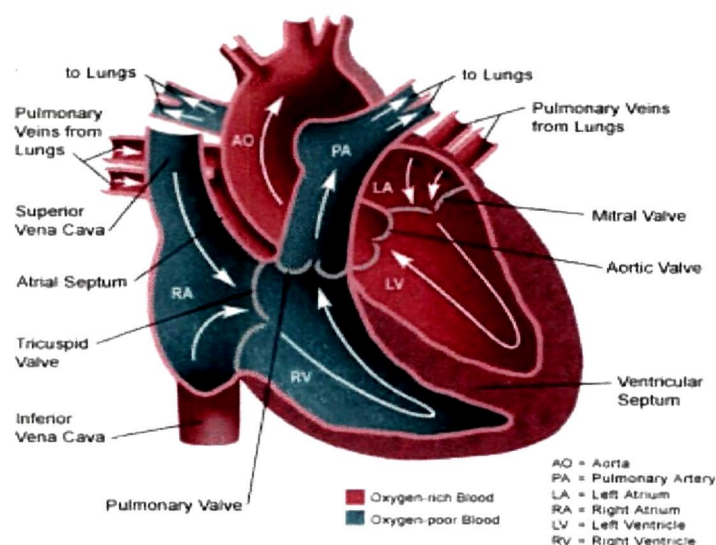
Figure 1.1 The heart's structure

The action potential generated by the SA node move throughout the right and left atrium, causing the muscle cells to contract. The atrial contraction is followed by the ventricles contraction. The ventricles contract together in a wringing motion, squeezing blood from

them. Thus, the movement of action potential results in the coordinated, sequential contraction of the heart's four chambers.

1.3. Electrical activity of the heart

The *SA node, atrioventricular node, atrioventricular bundle, right and left bundle branches*, and *Purkinje* fibers comprise the conduction system of the heart. All of these cells of the electrical conduction system are able to produce action potentials but at a slower rate as compared to SA node. The cardiac muscle of the SA node possesses the property of self-excitation. Under normal conditions, electrical activity is spontaneously generated by the SA node. The action potential generated by SA node is propagated throughout the right atrium, and through *Bachman's Bundle* to the left atrium, stimulating the atrial contraction.

The action potential propagates from SA node to atrioventricular (AV) node via specialized pathways known as internodal tracts. The AV node is situated between two sets of chambers of atria and ventricles. After a pause of about 1/10 of a second, the atrial depolarization wave reaches the AV node. This pause assists the atria to complete their contraction and empty their blood into the ventricles before the ventricles contract. Following their contraction, the ventricles begin to relax. After the ventricles have completely relaxed, another action potential originates in the SA node to begin the next cycle of contractions. The coordinated, sequential contraction of the heart's four chambers is governed by the movement of action potential throughout the cardiac muscle. This sequence of events is known as cardiac cycle.

1.4. The Cardiac cycle

Cardiac cycle is related to any of the events subjected to the flow of blood that occurs from the beginning of one heartbeat to the beginning of the next and spans an interval of approximately 0.8 second. During each cardiac cycle, all the four chambers of the heart go through a contraction and relaxation phase due to the action potential generated by the SA node. Ordinarily, this occurs about 60 –100 times each minute for a person above 18 years of age. Schematic representation of sequence of events in a cardiac cycle is shown in Figure 1.2.

The cardiac cycle has two basic components:

i. A contraction phase called systole during which blood is ejected from the heart's chambers.

ii. A relaxation phase called diastole during which the chambers of the heart are filled with blood.

In the first phase of the cycle both atria contract, the right atria contracts first followed almost instantly by the left atria. This atrial contraction results in filling of the relaxed ventricles with blood. Then both ventricles contract and force their blood to pulmonary artery and aorta respectively in a powerful surge. During ventricular contraction, the atria relax and are filled by blood once again by the veins. This cycle lasts, on the average 800 ms. The pressure and volume in atria, ventricles and arteries increase and decrease during each cardiac cycle [5]. These changes in the pressures and volumes at different locations (aorta, pulmonary artery, both atria and ventricles), thoracic vena cava
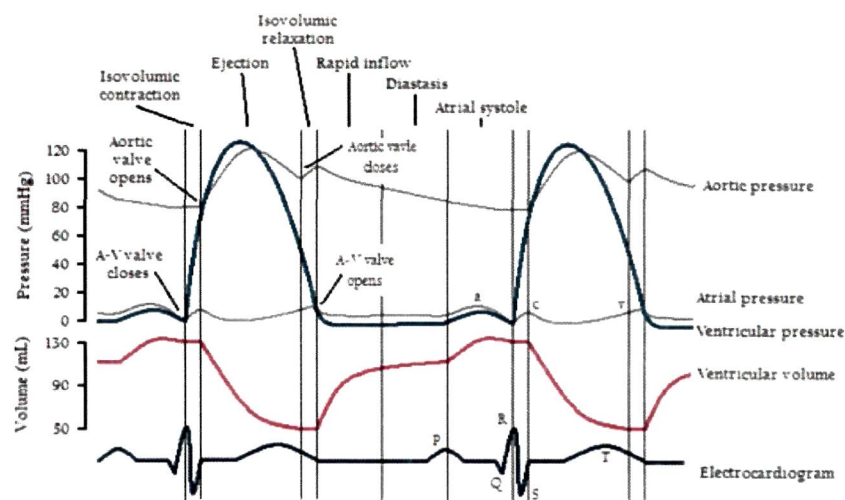
Figure 1.2 Schematic representation of cadiac cycle showing the changes in ventricular volume, ventricular pressure, aortic pressure, atrial pressure and ECG waveform

near the right atrium in the heart, are recorded by electronic device to assess the performance of cardiac function. The normal range of pressure signals is listed in Table 1.1.

Table 1.1 Normal range of pressure signals

Pressure Signal	Associated cardiac muscle	Normal diastolic pressure/ range (mmHg)	Normal systolic pressure/ range (mmHg)
Aotric pressure	Aorta	70 (60-90)	120 (100-140)
Right atrial pressure	Right atrium	2	8
Pulmonary artery pressure (PAP)	Pulmonary artery	< 10	< 25
Left ventricular pressure	Left ventricle	3-12	100-140
Central venous pressure	Thoracic vena cava near right atrium	3	8

In the normal heart, electrical activity produces the mechanical activity of systole and diastole. There exists a time difference between these electro-mechanical coupling, or the excitation-contraction phase. In the simultaneous recordings of the electrical and mechanical tracings, the changes in the activity of cardiac muscles are reflected first in electrical tracing before these changes are noticeable in pressure tracings [5].

1.5. Electrocardiogram

The spontaneous contraction and relaxation of atria and ventricles due to propagation of action potential produce electrical currents that can be measured at the surface of the body. Changes in the action potential of all the cardiac muscles are detected by the electrodes placed on the body surface attached to the recording device. The recorded tracing is called an electrocardiogram (ECG, or EKG). ECG or EKG is abbreviated from the German word "Elektrokardiogram". A normal ECG waveform consists of repetitive sequence of P, Q, R, S and T waves. ECG indicates the overall rhythm of the heart, and weaknesses in different parts of the heart muscle. Each heartbeat is the result of distinct cardiological events, represented by distinct features in the ECG waveform. ECG is the best way to measure and diagnose abnormal rhythms of the heart, particularly abnormal rhythms caused by damage to the conductive tissue that carries electrical signals, or abnormal rhythms caused by levels of salts, such as potassium, that are too high or low.

1.5.1. ECG waves

i. **P wave** – The spread of action potential throughout the atria is seen as the P wave on the ECG. P wave is a slow, low amplitude wave with an amplitude

---

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

of about 0.1-0.2 mV and a duration of about 60-80 ms.

ii. **QRS complex** – The ventricular depolarization results QRS complex in ECG tracing and rapid depolarization of the right and left ventricles. The ventricles have a large muscle mass compared to the atria, so the QRS complex usually has a much larger amplitude than the P-wave. It is of about 1 mV amplitude and 40-100 ms duration [7]. The QRS complex consists of three distinctive waves: Q, R and S.

iii. **T wave** –Ventricular repolarization (relaxation) is represented by a slow T wave which has an amplitude and duration of 0.1-0.3 mV and 120-160 ms respectively. Atrial repolarization wave is not seen in ECG as it appears during ventricular contraction.

iv. **U wave** – This wave is more often completely absent in ECG tracing. It normally has a low amplitude and always follows the T wave.

v. **J point** – It is the point at the end of QRS complex and beginning of ST segment. It is used to measure the degree of ST elevation or depression present.

1.5.2. ECG intervals

i. **PQ Segment** – This is an isoelectric segment of about 60-80 ms after the P-wave.

ii. **ST Segment** – Ventricular muscle cells possess a relatively long action potential of 300-350 ms. The Plateau portion of the action potential causes an iso-electric after the QRS complex, called ST segment which is about 80-120 ms.

iii. **RR interval** – It is interval between an R wave to the next R wave. It is generally 0.6 to 1.2 sec.

iv. **PR interval** – The PR segment represents the interval between beginning of P wave to beginning of QRS complex. It is mainly because of delay of action potential at AV node. It is generally 0.12 to 0.20 sec [7]. This appears flat on the ECG tracing. During this interval, the atria contract and begin to relax. At the

end of the PQ interval the ventricles begin to depolarize.The PR interval is more clinically relevant.

v.   QT interval – The QT interval is measured from the beginning of the QRS complex to the end of the T wave. It is less than 0.3 sec [8]. A prolonged QT interval is a risk factor for ventricular tachyarrhythmias and sudden death.

A normal ECG waveform with its constituent components is shown in Figure 1.3.
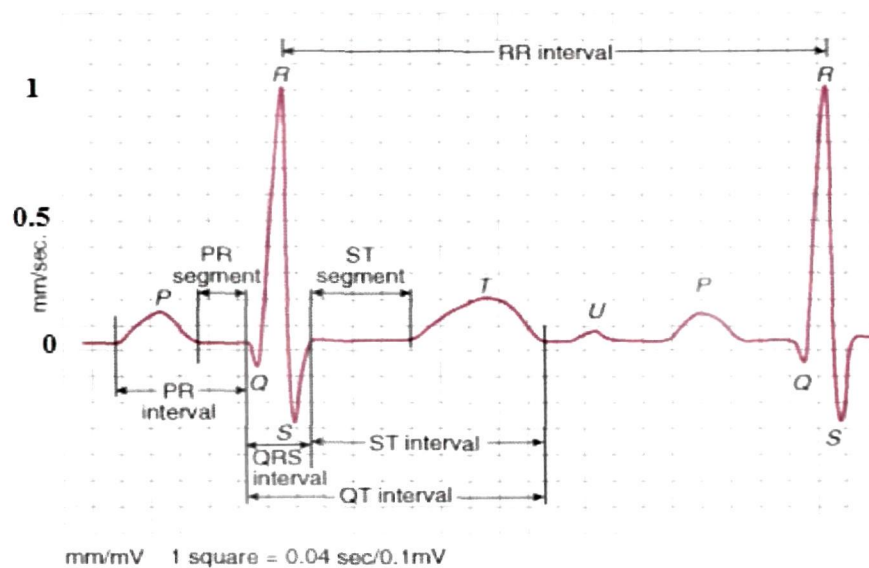
Figure 1.3 A normal ECG waveform

1.5.3. ECG leads

An electrode called 'lead' is used to record the electrical signals of the heart from a certain positional combination which are placed at specific points on the patient's body.

- Movement of depolarization wavefront (or mean electrical vector) towards a positive electrode is reflected as *positive* deflection on the ECG in the corresponding lead.

- Movement of depolarization wavefront away from a positive electrode is reflected as the *negative* deflection on the ECG in the corresponding lead.

- Perpendicular movement of depolarization wavefront to a positive electrode, appears an *equiphasic* (or isoelectric) complex on the ECG.

Modern ECG includes 12 leads I, II, III (limb leads), aV$_R$, aV$_L$, aV$_F$, (augmented limb leads) and V$_1$, V$_2$, V$_3$, V$_4$, V$_5$, V$_6$ (chest leads). Out of these 12 leads, the first six are derived from the same three measurement points. Therefore, any two of these six leads comprise of exactly the same information as the other four.

1.5.3.1. Limb leads

Leads I, II and III are called limb leads or standard leads of modern 12 lead ECG. They form the basis of Einthoven's triangle. The Wilson central terminal provides a virtual ground by comprising the sum of three lead electrodes connected to the central terminal through three resistors of 5KΩ. It is assumed that the heart lies as a point source at the centre of Einthoven's triangle [8].

The Einthoven limb leads (standard leads) are defined as -

$$\text{Lead I:} \quad V_I = V_L - V_R$$

$$\text{Lead II:} \quad V_{II} = V_F - V_R$$

$$\text{Lead III:} \quad V_{III} = V_F - V_L$$

Where, V_I, V_{II} and V_{III} are the voltage of lead I, lead II and lead III respectively and V_L, V_R and V_F is the potential at the left arm, right arm and left foot respectively. The layout of limb leads in Einthoven's triangle is shown in Figure 1.4.

1.5.3.2. Augmented limb leads

Leads aVR, aVL, and aVF are called augmented limb leads. They are obtained from limb leads I, II, and III. The negative electrode for these leads is a modification of Wilson's central terminal, which is derived by adding leads I, II, and III together and plugging them into the negative terminal of the ECG machine. These leads describe the electrical activity of heart from different angles (or vectors). This zeroes out the negative electrode and allows the positive electrode to become the "exploring electrode" or a unipolar lead. This is possible because Einthoven's Law states

$$I + III = II$$

Wilson's central terminal is the basis of the development of the augmented limb leads aVR, aVL, aVF and the precordial leads (V1, V2, V3, V4, V5, and V6).
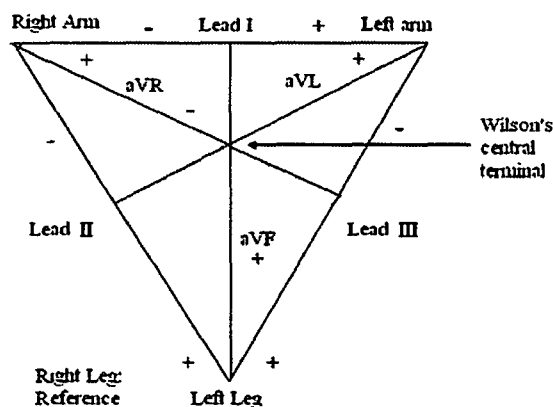
Figure 1.4 Einthoven's triangle and the axes of the six ECG leads formed by using four limb leads

1.5.3.3. Precordial leads

The precordial leads are used to measure the potential difference very close to the heart. Wilson's central terminal serves as negative electrode and these leads are considered to be unipolar. These leads record the heart's electrical activity in a horizontal plane. The heart's electrical axis in the horizontal plane is referred to as the Z axis.

Leads V1, V2, and V3 are referred to as the right precordial leads and V4, V5, and V6 are referred to as the left precordial leads. The position for placement of these precordial leads is shown in Figure 1.5. Modern four lead and 12 lead ECGs use an additional electrode (usually green). This is called the ground lead and is placed on the right leg by convention.

1.5.4. Clinical significance of ECG

The analysis of the ECG signal is the most readily available method for diagnosing cardiac rhythms. ECG effectively presents valuable clinical information regarding the rate, morphology and regularity of the heart while being a low cost and non-invasive test. Arrhythmias are one kind of dysfunctions or disturbances in the behavior of the heart. These disturbances produce abnormality in rate, rhythm and the site of impulse formation; factors that may in turn alter the normal sequence of atrial and ventricular activation. In ECG such arrhythmias manifest themselves as deformations in the observed waveform.

Deformation in the ECG waveform is a manifestation of the abnormality of the heart due to such arrhythmia. Some diseases which are diagnosed by the cardiologists in ECG tracing are listed in Table 1.2.
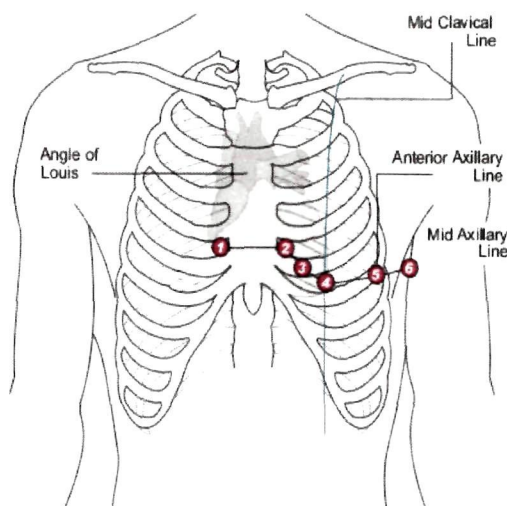
**Figure 1.5 Positions for placement of precordial (chest)
leads V1-V6 for ECG**

1.6. Arterial blood pressure

Blood pressure is an important cardiac variable and considered as a good indicator of status of cardiovascular system. Blood returning to the left atrium is at a low pressure, rising with contraction to 3 or 4 mm Hg. The left ventricle delivers blood to the body with considerable force. When the left ventricle pumps blood into the aorta, the aortic pressure rises to about 120 mm Hg with contraction, the same as the pressure in the arteries of the body. This maximum aortic pressure following ejection of blood into the aorta is termed as systolic pressure ($P_{systolic}$). As the left ventricle relaxes and refills, the pressure in the aorta falls to about 80 mm Hg. The lowest pressure in the aorta, which occurs just before the ventricle pumps blood into the aorta, is termed the diastolic pressure ($P_{diastolic}$) [9]. A typical "normal" blood pressure is 120/80.

Table 1.2 Heart diseases and corresponding change in ECG pattern
(Source : http://www.merckmanuals.com/)

Characteristics of ECG	Disease
Abnormal P wave	Left or right atrial hypertrophy, atrial escape (ectopic) beats
P wave is Absent	Atrial fibrillation, sinus node arrest or exit block, hyperkalemia
Variations in PP interval	Sinus arrhythmia
Long PR interval	First-degree AV block
Variations in PR interval	Multifocal atrial tachycardia
Wide QRS complexes	Ventricular flutter or fibrillation, hyperkalemia,
Short RR interval and wide QRS complex	Premature ventricular contraction
Long QT interval	MI (Myocardial ischemia), myocarditis, hypocalcemia, hypokalemia,
Short QT interval	Hypercalcemia, hypermagnesemia
Depression in ST segment	MI, ventricular hypertrophy, pulmonary embolism, left bundle branch block (LBBB), right bundle branch hyperventilation, hypokalemia
Elevation in ST segment	Myocardial ischemia, acute MI, LBBB, acute pericarditis, left ventricular hypertrophy, hyperkalemia, pulmonary embolism, hypothermia
Tall T waves	Hyperkalemia, acute MI, LBBB, stroke, ventricular hypertrophy
Non-specific change in T wave morphology	Pulsus paradoxus
Small, flattened, or inverted T wave	MI, myocarditis, left ventricular hypertrophy, pericarditis, pulmonary embolism, conduction disturbances (eg, Right bundle branch block (RBBB)), electrolyte disturbances (eg, hypokalemia)
Heart rate > 90 bpm	Sinus tachycardia
Heart rate > 50 bpm	Sinus bradycardia

1.6.1. Measurement of ABP

Commonly, arterial blood pressure is measured by doctors non-invasively with the help of a cuff sphygmomanometer which describe cardiac health in terms of systolic and diastolic values. The sphygmomanometer provides only single measurements of the systolic and diastolic levels [10]. ABP is most accurately measured invasively by placing a cannula needle in an artery or using a simple fluid filled pressure tubing-transducer. Direct monitoring of ABP is the only scientifically and clinically validated method for real time continuous monitoring of blood pressure. From continuous

measurements, cardiac performance can be further evaluated by calculating derived parameters [5].

1.6.2. ABP signal features

ABP waveform is rich in the estimation of cardiovascular function as compared to systolic and diastolic readings obtained by cuff sphygmomanometer. Therefore, continuous waveform of blood pressure is useful to acquire more meaningful information about its range and variability [11]. Continuous ABP monitoring not only provides information in regards to blood pressure, it also provides a means to assess the cardiovascular status by observing waveform characteristics since the waveform has close correlation with ECG. Analysis of ABP waveform provide useful diagnostic information, however physicians pay little attention to the morphology and detail of the ABP waveform and rely more on the diagnosis of ECG waveform. Morphologic features of individual ABP waveforms provide diagnostic clues to various pathologic conditions. Observation of arterial waveform patterns over consecutive heartbeats provides an additional set of diagnostic clues.

A normal ABP waveform comprises of four components such as systolic and diastolic pressures, dicrotic notch and dicrotic peak. The systolic peak and diastolic onset in ABP waveform depict maximum and minimum pressure in aorta. The appearance of dicrotic notch in ABP waveform is due to closing of aortic valve. Dicrotic peak in ABP waveform is the reflected impulse arising due to closing of aortic valve. A normal ABP waveform along with its features is shown in Figure 1.6. The recording is acquired from abp1 signal of CSL (Complex systems laboratory) database described later in this chapter [12].

1.6.3. Clinical significance of ABP waveform

Continuous monitoring of ABP waveform can be used to detect various diseases such as ischemic episodes, aortic stenosis, pulsus paradoxus etc. Analysis of ECG waveform can be used to diagnose heart diseases to a great extent, however, under certain conditions, variations in ABP waveform are reflected more clearly. The minimum pressure maintained by the coronary artery and tissue perfusion is called mean arterial pressure (MAP). This pressure can be calculated from the systolic and diastolic values obtained from electrical or manometric measurements using the following formula –

$$MAP = \frac{SBP + (2 \times DBP)}{3}$$

$$(1.1)$$

Where,

SBP = Systolic Blood Pressure (mm Hg)

DBP = Diastolic Blood Pressure (mm Hg)

MAP = Mean Arterial Pressure (mm Hg)

The normal value for MAP ranges from 70 to 105 mm Hg.

Characteristics of ABP waveform in different cardiac diseases are summarized in Table 1.3. Changes in ABP waveform in case of pulse paradoxus disease are shown in Figure 1.7 along with corresponding ECG waveform.

1.6.4. Relation between ABP and ECG

If ABP and ECG waveform are simultaneously recorded, the systolic components follow R wave in ECG and consist of a steep pressure upstroke, peak, and decline and correspond to the period of left ventricular systolic ejection. The downslope of the arterial pressure waveform is interrupted by the dicrotic notch, then continues its decline during diastole following T wave in ECG, reaching its lowest point at end-diastole. The difference between successive systolic peaks is equal to RR interval in ECG from which PR and QT intervals
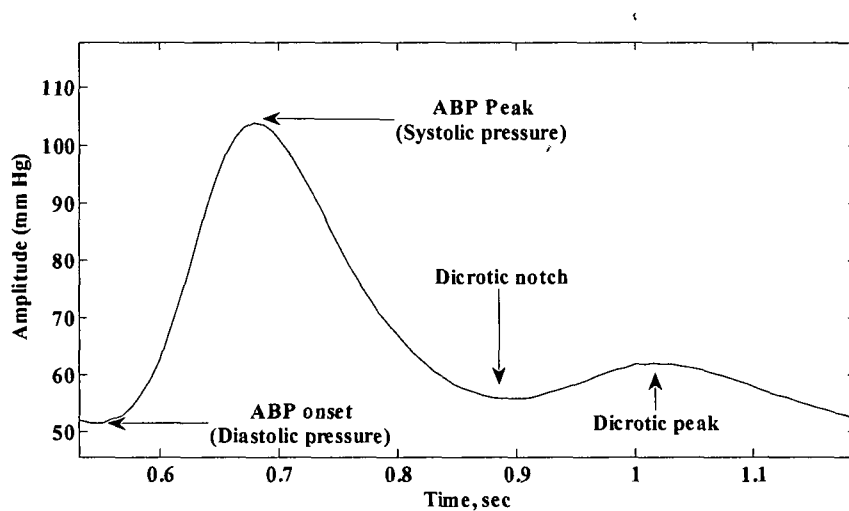
Figure 1.6 Normal ABP waveform and its features

Table 1.3 ABP waveform characteristics in case of cardiac diseases

Characteristics of ABP signal	Disease
A small pulse wave with a delayed peak systolic pressure	Aortic stenosis
Higher peak systolic pressure during the next systole.	Aortic insufficiency or aortic regurgitation
varying amplitudes in the arterial waveform	Atrial fibrillation
Decrease in peak systolic pressure	Premature ventricular contraction
Regular altering amplitudes of the peak systolic pressures	Pulses alternans
> 10 mm Hg difference in systolic pressures from inspiration to expiration	Pulsus Paradoxus

can be estimated [13, 14]. It is observed that any abnormality reflected in ABP waveform due to cardiac disease is also associated with ECG waveform. Diagnosis of cardiac diseases from ABP waveform and corresponding ECG waveform are summarized in Table 1.4. Simultaneously acquired ABP and ECG waveforms in case of pulsus paradoxus and artial fibrillation diseases are shown in Figure 1.7 and Figure 1.8.

Intra thoracic pressures are low during inspiration and higher during expiration. This pressure change is generally limited to 3-10 mmHg. Under abnormal conditions, this pressure changes more than 10 mmHg, this condition is referred as pulsus paradoxus as

Table 1.4 Disease diagnosis by ABP signal and corresponding ECG signal

ABP Characteristics	ECG Characteristics	Disease
Systolic ABP drops after PVC and PSC [34]	QRS complex is wider and RR interval is small [15, 16]	Premature ventricular contraction and premature supraventricular contraction
MAP < 60 mm Hg	Variations in ST-T segment [17]	Ischemic Episodes
Decreased systolic blood pressure and variation in systolic peak to peak interval	Change in RR interval, Absence of P wave [18] (Figure 1.8)	Cardiac arrhythmia (Atrial fibrillation)
Decreased systolic blood pressure	P waves are recorded after QRS complexes	AV dissociation

shown in Figure 1.7. ABP waveform reflects changes in systolic blood pressure (>10 mm Hg) during inspiration and expiration whereas ECG waveform displays non-specific

changes in T wave morphology. These changes are due to the fact that ventricles are unable to fill sufficiently due to this disease.
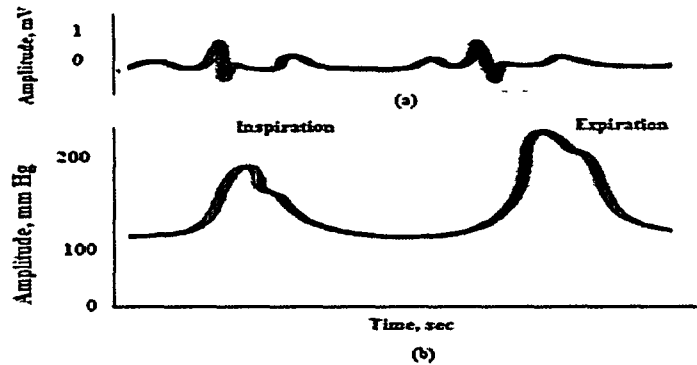
Figure 1.7 (a) ECG waveform and (b) corresponding ABP waveform in case of pulsus paradoxus disease [5]

1.7. Central venous pressure

CVP is the pressure of deoxygenated blood at the junction of thoracic vena cava and the right atrium of heart that reflects the driving force for filling the right atrium and ventricle. It reflects the suitability of blood volume owing to capacity of the venous system. It reflects the amount of blood returning to the right atrium and functional capacity of right ventricle
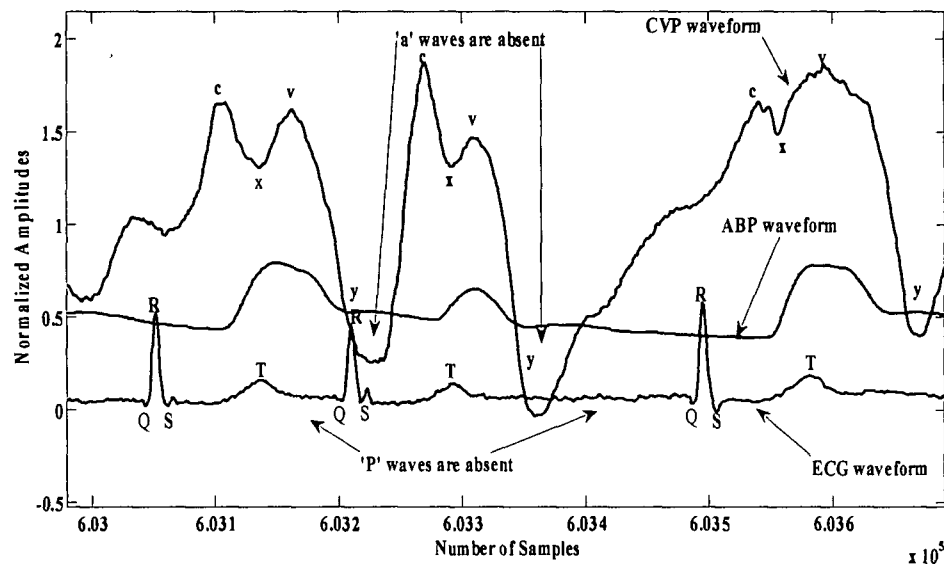
Figure 1.8 Changes in the features of ABP, CVP and ECG in atrial fibrillation

[19]. In the event of impairment of the contractility of right ventricle, ventricular stroke output is maintained by higher pressures in the right atrium.

Atrial activity also encompasses the systolic and diastolic phase of hemodynamic events like ventricles [5]. Blood returning from the right atrium through veins is under a relatively low pressure of about 1 or 2 mm Hg. The right ventricle increases the blood pressure to about 20 mm Hg during systole while the blood is transferred to lungs for oxygenation.

1.7.1. Measurement of CVP

CVP can also be measured using a simple fluid-filled mechanical manometer. However, in clinical practice, electronic pressure trnsducers are preferred because in manometric measurements continuous display is not possible. Also, CVP measurement by an open manometer may expose the patient to the risk of infection and venous air embolism. A CVP waveform comprises of plenty of cardiac information, therefore the continuous measurement of CVP is preferred using electronic pressure transducer and the waveforms are displayed on a bedside monitor in ICU.

1.7.2. CVP features

A CVP waveform reflects the events of cardiac contraction. CVP waveform exhibits the slight variations in pressure occurred during the cardiac cycle. CVP waveform has five phasic components which are recognized by three upward waves of low amplitudes and two downward waves appearing between two consecutive R-peaks of ECG. The ascending waves are characterized as 'a', 'c' and 'v' waves and descending waves as 'x' and 'y' waves. Increase in right atrial pressure during right atrial contraction results in 'a' wave and 'c' wave appears due to closing of AV valve but is not always evident. The increase in atrial pressure due to right atrial filling results in 'v' wave and rapid decline in pressure due to flow of blood from right atrium to right ventricle results in 'y' descent in CVP tracing. These waves have close correlation with ECG. A normal CVP waveform representing all its five features is shown in Figure 1.9.
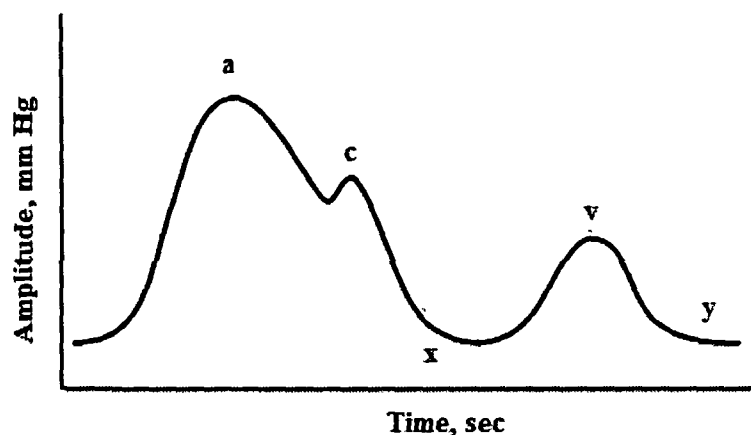
Figure 1.9 CVP waveform and its features

1.7.3. Clinical significance of CVP signal

Variations on the normal CVP waveform can provide information about cardiac pathology and can be used for the diagnosis of various pathophysiologic conditions of the heart. One of the most common applications includes the rapid diagnosis of cardiac arrhythmias. Few cardiac abnormalities are reflected in abnormal CVP tracing as listed in Table 1.5.

Table 1.5 Central venous pressure waveform abnormalities

(Source : http://www.healthsystem)

CVP characteristics	Disease
Loss of 'a' wave	Atrial fibrillation
Prominent c wave	
Cannon 'a' wave	Atrioventricular dissociation
Tall systolic 'c-v' wave	Tricuspid regurgitation
Loss of 'x' descent	
Tall 'a' wave	Tricuspid stenosis
Attenuation of 'y' descent	
Tall 'a' and 'v' waves	Right ventricular ischemia
Steep 'x' and 'y' descents	
Tall 'a' and 'v' waves	Pericardial constriction
Steep 'x' and 'y' descents	
Dominant 'x' descent	Cardiac tamponade
Attenuated 'y' descent	

1.7.4. Relation between ECG and CVP

The CVP waves arising due to cardiac activity have a close relationship with ECG. If simultaneously recorded CVP and ECG signals are available, then following relationship

between CVP signal features and ECG are observed as shown in Table 1.6. A normal CVP waveform along with ABP and ECG waveforms is shown in Figure 1.10.

The ABP, CVP and ECG signals also preserve the relationship among their features under abnormal cardiac conditions such in atrial fibrillation, 'a' wave obliterates, amplitude of 'c' wave increases but 'v' wave and 'y' descent are preserved in CVP signal. Corresponding

Table 1.6 Relation between CVP signal activities and ECG

CVP signal feature	ECG feature
'*a*' wave	Offset of P wave
'*c*' wave	End of QRS segment
'*x*' descent	Onset of T wave
'*v*' ascent	Offset of T wave
'*y*' descent	Onset of P wave
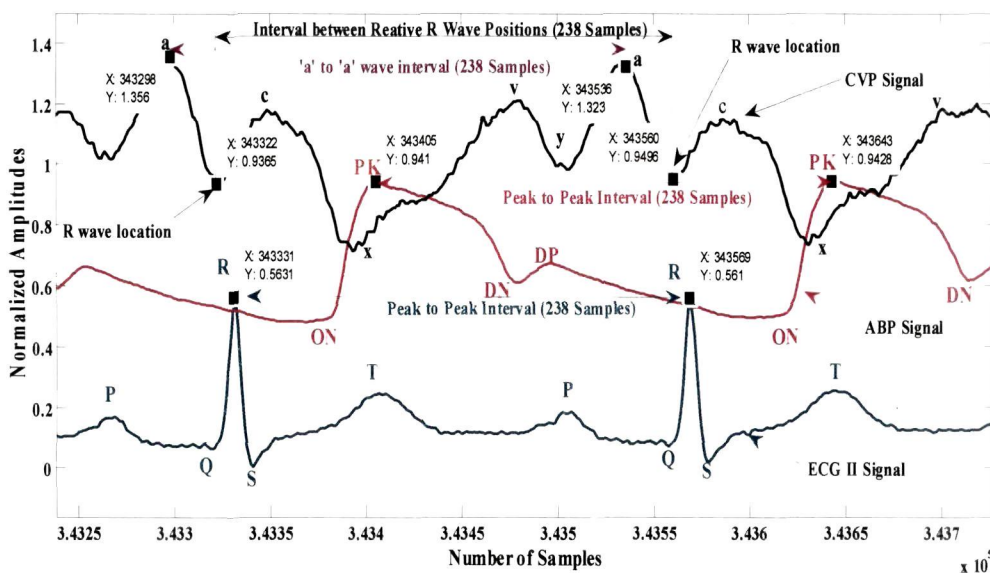
Figure 1.10 Relationship among the features of CVP, ABP and ECG signals

ECG shows absence of P wave and change in RR interval and ABP waveform displays change in systolic peak to peak interval and drop in systolic pressure. Changes in ABP, CVP and ECG waveform in atrial fibrillation are shown in Figure 1.8. The changes in ABP, CVP and ECG signals in certain abnormal conditions are listed in Table 1.7. Therefore, it is seen that CVP waveform is abundant in physiological of cardiovascular system. This signal can be used to detect various cardiac diseases like ECG signal.

Table 1.7 Disease diagnosis by pressure signals and corresponding ECG signal

ABP signal characteristics	CVP signal characteristics	ECG characteristics	Disease
Decreased systolic blood pressure and variation in systolic peak to peak interval (Figure 1.8)	Absence of 'a' wave and prominent 'c' wave (Figure 1.8)	Change in RR in interval, Absence of P wave [18] (Figure 1.8)	Cardiac arrhythmia (Atrial fibrillation)
Decreased systolic blood pressure	Early systolic cannon 'a' wave	P waves are recorded after QRS complexes	AV dissociation

1.8. Prior works on feature extraction techniques

Cardiologists look for any abnormal feature in the ECG waves and ECG segments as discussed in section 1.5.1. and 1.5.2. Although there are many cross-correlations between various ECG features in the context of disease diagnosis, few critically accepted conditions of feature values are listed in Table 1.2.

To take decision on cardiac abnormalities, the cardiologists manually calculate the features from the scale of the ECG trace. However, in case of large digitized data such as Holter ECG recording, it requires automated computer based techniques to perform the following tasks –

i) Extract the feature values from the digitized ECG records.

ii) Apply disease diagnosis algorithm to find abnormalities.

Therefore, ECG feature extraction has been identified as a major issue in the research of biomedical signal processing since long. From Table 1.2, it is clear that components of ECG waveform such as P, QRS, T waves and their intervals such as PR, RR and QT comprise of valuable clinical information.

1.8.1. ECG feature extraction techniques

In the last decade, Various ECG feature extraction techniques have been developed so far which include wavelet transform [20-25] histogram and genetic algorithms [26], Artificial Neural Networks [27], moving average filter [28], differentiation and correlation [29] comprising of their own merits and demerits. Out of these methods [20-29], the advanced

signal processing methods using time-frequency analysis and filtering using wavelet transform have proved to be a very useful tool in determining the precise location of the QRS-complex. Other features of ECG such as P, Q, S, T waves are detected in [20, 21, 24, 25] considering R peak as reference. Intervals such as PR, RR, QT, ST etc. can be estimated from all detected components of ECG waveform. Wavelet transform is popular because it satisfies energy conservation law and original signal can be reconstructed [21].

In these [20-25] wavelet based methods, ECG peak detection has been carried out by authors by selecting detail signals 'd1-d4' in [20,21], 'd3-d5' in [24, 25], 'd5-d6' in [23] and 'd4' signal in [22]. In [20], author has selected 'd1-d4' signals based on 3 db bandwidth. It is evident from above literature that the authors have selected detail signals out of 'd1- d6' signals for ECG peak detection, however, selection of detail signals in the above literature is not adequately justified by any of these authors [20-25]. Most of the QRS detection algorithms mentioned above [24,25,28] are developed on ECG lead-II signals of MIT-BIH database as the QRS complexes are outstanding in lead II. In [30], author has included two ECG signals from leads V5/V2 but accuracy reported is 94% and 92% respectively.

1.8.2. ABP feature extraction techniques

Blood pressure waveform analysis has been well recognized in cardiac physiology for the assessment of properties of arterial vessel wall [31], cardiac output monitoring [32], estimation of pressure pulse index [33] and cardiac arrhythmia detection [34,35]. Continuous measurement of mean arterial pressure (MAP) obtained from ABP waveform has shown that a MAP < 30 mmHg in infants is significantly associated with severe haemorrhage [36]. Thereby it is well assumed that analysis of arterial blood pressure waveform can provide better insight of heart in cardiac physiology. Moreover, mathematical modeling of non-invasive ABP waveform has been used to estimate various cardiac parameters such as cardiac output, arterial compliance and peripheral resistance [37,38]. In some cases, parallel analysis of ABP waveform along with ECG has resulted in reducing false alarms to a certain extent for critical arrhythmias detection [39]. Ramaswamy etal [34,35] suggested that combination of ABP waveform along with ECG signal gives better results for the detection of ectopic beats than ECG signals only.

Therefore, analysis of ABP waveform can be used to estimate the cardiac health to certain stage when ECG waveform is not available and can assist to enhance the comprehensive knowledge of heart when simultaneously acquired ECG and ABP waveforms are available. In addition to this, study of systolic pressure variation is used in the early diagnosis of hypovolemia.

Cardiac signatures due to variations in ABP waveform under normal as well as abnormal conditions can be better explained if algorithms for feature extraction of ABP signals may be available. Although a considerable amount of research has been carried out for feature extraction of ECG signals but there are very few algorithms reported for feature extraction of ABP signal. Most of the algorithms on ABP signal are developed on proprietary datasets acquired from rabbits [40] and dogs [41]. These methods are based on continuous independent assessment of refractory period (RP), analysis of signal by means of producing two moving averages [40], template matching [41], rank filter and decision logic [42], windowed and weighted Slope Sum Function (SSF) [43], peak and trough detection methods [35], heart rate, amplitude and inter beat intervals [44] and combinatorial analysis of ABP waveforms and their derivatives [45].

The researchers have paid major attention on the detection of systolic peaks [44], onsets [43,46] and dicrotic notches [41]. There is only one algorithm reported in literature for the detection of three features (systolic peak, onset, dicrotic notch) together [45]. Dicrotic peak is also an important feature of ABP signal, however, no attempt has been made to detect dicrotic peak in ABP signal.

1.8.3. CVP feature extraction techniques

Like ECG and ABP waveform, CVP waveform comprises of plenty of important cardiac information, therefore, it is of great interest for cardiologists from clinical point of view. However, no attempt is made for the development of feature extraction technique of CVP signal.

1.9. Signal processing techniques

The techniques of signal processing involve recovery of information from physical signals. In the context of biomedical signal processing, the derivation of the information means –

amplitude, frequency and time duration of the signals which carries signatures of the health of the subject. As discussed in section 1.1 to 1.8, the cardiac conditions of a subject get reflected in the recorded ECG and cardiac pressure (ABP and CVP) signals and their features.

One of the aims of this research is extraction of features of the cardiac signals (ECG, ABP and CVP) by applications of either newly developed techniques or modified algorithm based on existing techniques. Deriving the features from cardiac signals involves various techniques such as wavelet decomposition, Fast Fourier Transform (FFT), energy analysis, correlation analysis, maxima/ minima detection, thresholding etc.

Another objective of this research in modeling and synthesis of ECG where, we have applied system identification and ANN techniques.

Preprocessing is a vital step in applying the signal and data to the signal processing algorithms. Filtering, normalization and offset removal are the most common preprocessing techniques that we have adopted in application of the algorithms. This section explains the overview of the signal processing techniques which are enumerated in details in the corresponding chapters of the work.

1.9.1. Wavelet transform

The wavelet transform is defined as the projection of a signal on the set of basis functions, referred to as wavelets that are derived from a basis function (i.e. mother wavelet) by dilation and contraction operations. Wavelets have been used for the illustration and analysis of many physiologic signals such as ECG, EMG because of their compact support (finite length). Physiologic signals can be reasonably characterized as isolated pulses or as sequences of pulses. Wavelet transform of a signal results in the concentration of signal energy in a relatively small number of coefficients that makes wavelet-based techniques potentially powerful tool for signal processing algorithms [47]. Noise generally encountered in the clinical environment is automatically eliminated due to inherent characteristics of wavelet technique, therefore, this technique has been a powerful tool to enhance signal quality by removal of noise and interference in the physiological signals and images [48-51]. Wavelet based techniques are also employed for data compression of

biomedical signals and images for telemedicine applications [52-55]. The wavelet analysis has also emerged as a very promising technique for feature extraction of ECG signals [20-25].

1.9.1.1. Wavelets

Unlike Fourier transform, wavelet transforms have an infinite set of possible basis functions. Thus, wavelet analysis provides immediate access to information that can be obscured by other time-frequency methods such as Fourier analysis. Figure 1.11 shows the coverage in the time-frequency plane with one wavelet function, the Daubechies wavelet (db1). There are several wavelet families like Harr, Daubechies, Biorthogonal, Coiflets, Symlets, Morlet, Mexican Hat, Meyer etc. and several other Real and Complex wavelets. Within each family of wavelets, there are wavelet subclasses distinguished by the number of coefficients and by the level of iteration. Wavelets are classified within a family most often by the number of vanishing moments. The names of the Daubechies family wavelets are written dbN, where N is the order, which represents the number of vanishing moments and db is the "surname" of the wavelet. The db1 wavelet, as mentioned above, is same as Haar wavelet. The wavelet functions ψ of the next nine members of the db family are shown in Figure 1.12.
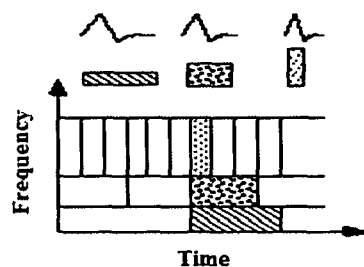
Figure 1.11 Daubechies wavelet basis functions, time-frequency tiles, and coverage of the time-frequency plane

The wavelet coefficient matrix is applied to the data vector in a hierarchical algorithm, sometimes called a pyramidal algorithm. The wavelet coefficients are arranged so that odd rows contain an ordering of wavelet coefficients that act as the smoothing filter, and the even rows contain an ordering of wavelet coefficient with different signs that act to bring out the details of data. The matrix is first applied to the original, full-length vector. Then
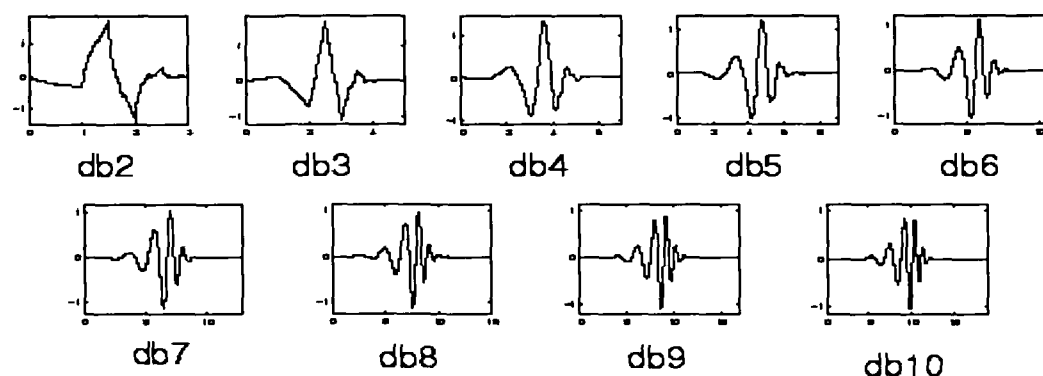
Figure 1.12 Daubechies wavelets

the vector is smoothed and decimated by half and the matrix is applied again. Then the smoothed, halved vector is smoothed, and halved again, and the matrix applied once more. This process continues until a trivial number of "smooth-smooth-smooth..." data remain. That is, each matrix application brings out a higher resolution of the data while at the same time smoothing the remaining data.

1.9.1.2. Continuous wavelet transform

This transform is based on the convolution of the signal with a dilated filter [56]. Wavelet analysis divides the signal into different frequency components and depends upon choosing a mother wavelet. The signal under study is represented as a linear combination of dilation and translation parameters of this selected mother wavelet. If the scale is continuous then the transform is called continuous wavelet transform (CWT). If the scale is discrete, the transform can be either orthogonal or non-orthogonal [57].

The mother wavelet is chosen to serve as a prototype for all windows in the process. All the windows that are used are the dilated (or compressed) and shifted versions of the mother wavelet.

We can write a general transformation equation [22] as follows:

$$x(a,b) = \int_{-\infty}^{\infty} x(t)\, \psi_{a,b}(t)\, dt$$

(1.2)

Where $x(t)$ is the given signal to be processed. For wavelet transform, the function $\psi(t)$ is given by –

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

(1.3)

Where, $\psi_{a,b}(t)$ is a window of finite length, 'b' is a real number known as window translation parameter and 'a' is a positive real number called as dilation or contraction parameter.

Thus, CWT of the signal $x(t)$ can be written as –

$$X_W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\, \psi*\left(\frac{t-b}{a}\right) dt$$

(1.4)

Where, * denotes the complex conjugation [58]. In other words, it can be viewed as a measure of similarity between the signal and wavelet.

The admissible conditions for ψ (t) as mother wavelet are as follows [59] –

- It must have finite energy i.e.

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

(1.5)

- If $\hat{\psi}(f)$ is the Fourier Transform of $\psi(t)$ i.e.

$$\hat{\psi}(f) = \int_{-\infty}^{\infty} \psi(t)\, e^{-i(2\pi f)} \, dt$$

(1.6)

Then, the following condition must hold –

$$C_g = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(f)|^2}{f} \, df < \infty$$

(1.7)

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

Fourier transform must be real and vanish for negative frequencies. It is convenient to analyze the signal if the wavelet is dilated and contracted because the time-frequency plane can be conveniently covered for the dilation and contraction [60].

Therefore, CWT shows that the wavelet analysis is a measure of similarity between the basis functions (wavelets) and the signal itself. Here, the similarity is in the sense of similar frequency content. The calculated CWT coefficients refer to the closeness of the signal to the wavelet at the current scale.

1.9.1.3. Discrete wavelet transform

If the wavelet $\psi(t)$ is the derivative of a smoothing function $\theta(t)$, it can be shown that the wavelet transform of a signal $x(t)$ at a scale a is

$$W_a x(b) = -a \frac{d}{db} \int_{-\infty}^{\infty} x(t)\, \theta_a (t - b)\, dt$$

(1.8)

Where $\theta_a (t - b) = \frac{1}{\sqrt{a}} \theta\left(\frac{t}{a}\right)$ is the scaled version of the smoothing function [61]. It is evident from the above equation that the wavelet transform at scale 'a' is proportional to the derivative of the filtered signal with a smoothing impulse response at scale 'a'. Hence zero crossings of wavelet transform at different scales will result in local maxima or minima and maximum slopes in the filtered signal will occur at maximum absolute values of wavelet transform [62].

A dyadic wavelet transform is implemented using the set of high-pass and low-pass filters that are derived from coefficient wavelet referred as mother wavelet. These filters are called analytical filters. These high-pass and low-pass filters are related to each other by the following relation-

$$g[L-1-n] = (-1)^n \times h[n]$$

(1.9)

Where, g[n] and h[n] denote the high-pass and low-pass filter transfer functions respectively, and L is the filter length expressed in number of points [63]. Both filters are odd index alternated inversed versions of each other. Low-pass to high-pass conversion is made by the $(-1)^n$ term. Filters satisfying this condition are commonly used in signal

processing, and they are known as the Quadrature Mirror Filters (QMF). The two filtering and subsampling operations can be expressed by –

$$y_{high}[k] = \sum_{n} x[n].g[-n+2k]$$
(1.10)

$$y_{low}[k] = \sum_{n} x[n].h[-n+2k]$$
(1.11)

Detail signal (y_{high})and average signal (y_{low}) are the outputs of high-pass and low-pass filters respectively. These generated signals consist of lower scale and upper scale information of the original signal. The low-pass filter coefficient undergoes subsampling to generate another new detail signal and average signal. Thus the dyadic discrete wavelet transform (DWT) is the composition of dilated and translated form of mother wavelet. This process of decomposition of the signal may be continued until the average signal reaches the length of a single sample or a length that is not applicable for further application of the analysis filter pair [64]. Every decomposition results in decreasing the time resolution by a factor of 2 whereas the frequency resolution is doubled. The schematic representation of wavelet decomposition stages is shown in Figure 1.13.
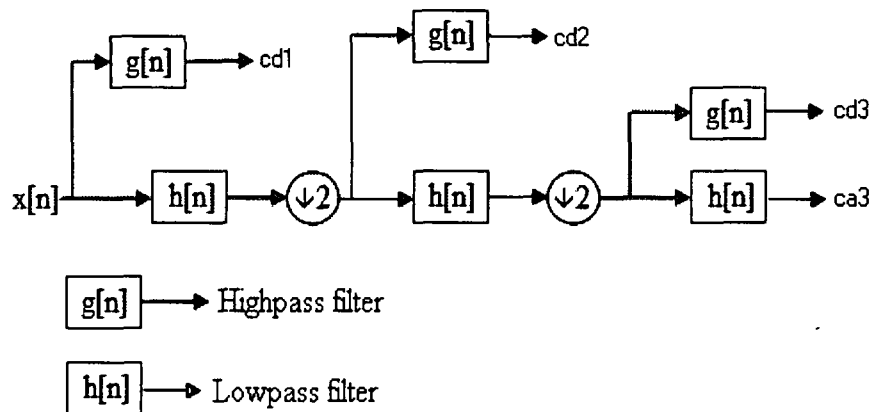
Figure 1.13 Schematic representation of wavelet decomposition stages

Perfect reconstruction of the signal is possible only with the ideal half band filters such as Daubechies set of wavelets. For reconstruction purpose, the decomposition process is followed in reverse order. The wavelet coefficients obtained at each level are upsampled by two, and passed through synthesis filters (high- pass and low-pass) and are added.

Therefore, the signal reconstruction is referred as the inverse DWT and formula for reconstruction is given by [66] –

$$x[n] = \sum_{k=-\infty}^{\infty} \left(y_{high}[k].g[-n+2k] \right) + \left(y_{low}[k].h[-n+2k] \right) \tag{1.12}$$

The wavelet transform satisfies the energy conservation principle and the original signal can be faithfully reproduced [21]. Figure1.14 shows an example of wavelet decomposed signals of ECG signal. The wavelet technique is applied in the feature extraction of ECG, ABP and CVP signals in Chapter 2 and Chapter 3. This method has been implemented in the removal of power-line interference in section 2.1.1.4 of Chapter 2.
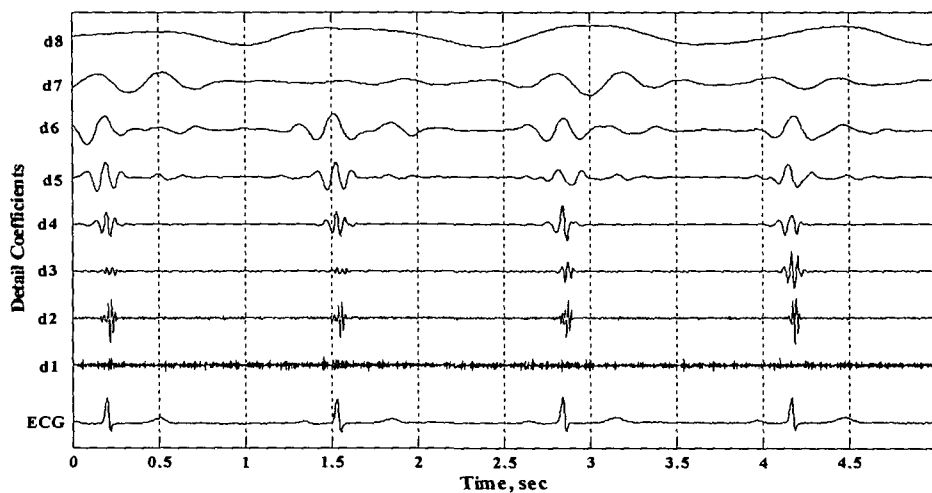
Figure 1.14 ECG signal and its detail signals using db6 wavelet

1.9.2. Signal energy analysis

Signals in time domain carries information of the system which is reflected in the signal amplitude and the time scale it is spread to. Moreover, signals of different amplitudes but with different spreading time may have similar strengths due to the fact that they carry equal energies. Such situation arises in biomedical application where the bio-generated signals are disrupted on their way by some resistance of the nerves. In such cases, the nerve cells recruit the same energy, however the amplitude gets lowered over a wider time duration. For biomedical signal processing, therefore, energy analysis gives a promising clue to analyze the signal strength for detection of some features of the signal.

In many applications, the signals are directly related to physical quantities capturing power and energy in a physical system. A signal can be defined as a function of varying amplitude over a time, so signal strength can be measured by the area under the curve. This area may have a negative part, but the negative area also contributes to the total signal energy. Therefore, signal energy is calculated by squaring the absolute value, then finding the area under that curve [63].

For a real signal energy will be given by –

The total energy for any continuous time signal $x(t)$ over the time interval $t_1 \leq t \leq t_2$ is defined as

$$E = \int_{t_1}^{t_2} |x(t)|^2 \, dt \qquad (1.13)$$

where $|x|$ denotes the magnitude of the signal x(t).

Similarly, the total energy in a discrete time signal $x[n]$ over the time interval $n_1 \leq n \leq n_2$ is defined as

$$E = \sum_{n=n_1}^{n_2} |x[n]|^2 \qquad (1.14)$$

The signal energy should be finite. In practice, all time limited signals are called energy signals. It is essential for a signal to have finite energy that the amplitude $x(t)$ or $x[n]$ must tend to zero otherwise the signal energy will be infinite. A signal $f(t)$ with its energy distribution is shown in Figure 1.15.
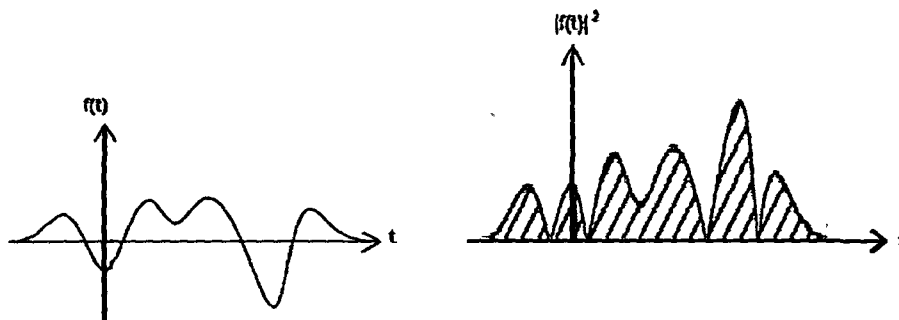
Figure 1.15 Signal with energy distribution in the shaded region

1.9.3. Fast Fourier Transform

Discrete Fourier Transform (DFT) corresponds to the computation of N samples of the Fourier transform of the N equally spaced frequencies $\left(w_k = 2\pi \frac{k}{N} \right)$. When DFT is computed by efficient algorithms, it is called Fast Fourier Transform (FFT). To achieve highest efficiency, FFT must compute all N point DFT. The DFT of a finite length sequence of length N is given by –

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \tag{1.15}$$

Where, $W_N = e^{-k(2\pi/N)}$ and $k = 0,1,2,\ldots(N-1)$

This equation involves N complex multiplications (i.e. N^2 multiplications and additions) to get each value of $X(k)$. FFT takes the computations in the order of $N \log_2 N$ computations instead of N^2.

Dividing these summations into even and odd values of n we can write –

$$X(k) = \sum_{r=0}^{N/2-1} x(2r) W_{N/2}^{rk} + W_N^k \sum_{r=0}^{N/2-1} x(2r+1) W_{N/2}^{rk} \tag{1.16}$$

Where, $x(2r)$ and $x(2r+1)$ are the sequences given by –

$x_e(r) = x(2r)$ and $x_0(r) = x(2r+1)$, $r = 0,1,2,\ldots N/2-1$

Replacing r by n, above equation (1.x) can be written as –

$$X(k) = \sum_{r=0}^{N/2-1} x_e(n) W_{N/2}^{nk} + W_N^k \sum_{r=0}^{N/2-1} x_0(n) W_{N/2}^{rk} \tag{1.17}$$

Therefore, N point DFT of a signal has been expressed as a sum of two DFTs each comprising of N/2 points [66].

FFT algorithms are used in spectral analysis, filter banks, data compression, solving partial differential equations, polynomial multiplications, convolution etc. We have used in this

research for frequency analysis of wavelet detail and approximation coefficients after decomposition and ECG/ ABP signals affected by baseline drift which are discussed in Chapter 2 and 3.

1.9.4. Cross-correlation analysis

It often requires that two signal components have to be compared for testing resemblance in time domain. Cross-correlation is the efficient technique that signifies the degree of similarity between two signals in time domain.

Let us consider two time domain signals $x(t)$ and $y(t)$ which may or may not be periodic and not restricted to a finite interval. Then the cross-correlation between these two signals $x(t)$ and $y(t)$ is defined as

$$C_{1,2}(\tau) = Lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} x(t)\, y*(t)\, dt \tag{1.18}$$

Where, time (t) is a dummy variable. For real valued signals, the conjugate symbol '*' may be neglected.

The correlation coefficient is given by the following relation –

$$\text{Correlation coefficient (r)} = \frac{SS_{XY}}{\sqrt{(SS_{XX})(SS_{YY})}} \tag{1.19}$$

The sum of squares (SS) for a variable 'X' is

$$SS_{XX} = \sum (x_i - \bar{x})^2 \tag{1.20}$$

Similarly, the sum of squares for variable (Y) is given by –

$$SS_{YY} = \sum (y_i - \bar{y})^2 \tag{1.21}$$

The sum of cross-products SS_{XY} is given by

$$SS_{XY} = \sum (x_i - \bar{x})(y_i - \bar{y}) \tag{1.22}$$

Positive or negative correlation is determined by the sign of the correlation coefficient. The

value of correlation coefficient ranges from +1 to −1. The strength of correlation is dependent on the magnitude of correlation coefficient. However, following guidelines are suggested for finding out the strength of correlation.

$$0 < |r| < 0.3 \qquad \text{weak correlation}$$

$$0.3 < |r| < 0.7 \qquad \text{moderate correlation}$$

$$|r| > 0.7 \qquad \text{Strong correlation}$$

We have determined cross-correlation coefficients for finding the similarity between ECG/ ABP/CVP signals and their wavelet decomposed detail coefficients described in Chapter 2 and 3 respectively. We have also used to cross-correlation for similarity measures of ECG (synthesized using ANN) with original ECG which is described in chapter 5.

1.9.5. Thresholding

Thresholding is a technique used for signal and image denoising. In biomedical applications, it is used for medical image analysis to differentiate structures such as organs and tumors, so that tissue volumes can be measured and computer-guided surgery can be performed. Similar methods find their applications for white blood cell classification [67], recover a function from noisy sampled data [68], image denoising and segmentation [69,70], compression and filtering of biomedical signals [71].

In thresholding, small coefficients are suppressed by introduction of a threshold. Such a procedure is called wavelet thresholding. There exist various thresholding procedures such as soft thresholding and hard thresholding. Wavelet shrinkage is usually performed using one of two predominant thresholding schemes known as hard and soft thresholding.

Hard threshold filter *Hh* removes coefficients below a threshold value, determined by the noise variance. This is also referred to as the "keep or kill" method. It has shown that hard thresholding provides an improved signal to noise ratio. The soft threshold filter *Hs* shrinks the wavelet coefficients above and below the threshold. Soft thresholding reduces coefficients toward zero .The process of thresholding is lossy, therefore the obtained signal is irreversibly different than the original signal.

Soft threshold filter is used to achieve smooth signal after thresholdng. However, the hard threshold filter performs better. It is difficult to choose a threshold value. A small threshold value creates a noisy result near the input, while a large threshold value introduces bias.

Hard thresholding sets any coefficient less than or equal to the threshold to zero. If x is the signal which is to be thresholded and th is the threshold then in case of hard thresholding –

$$if \; x(i) \le th$$
$$x(i) = 0$$

(1.23)

In Soft thresholding the threshold is subtracted from any coefficient that is greater than the threshold. This moves the time series toward zero.

$$if \quad x(i) \le th$$
$$x(i) = 0$$
$$else \; x(i) = x(i) - th$$

(1.24)

1.9.6. Maxima and minima of signals

In signal processing, detection of maxima and minima is desired when there is a need to see the locations in the signal where there are sharp variations in signal amplitudes. A maxima is the highest point in the valley whereas minima is the point where signal amplitude is lowest. Maxima and minima in a signal are shown in Figure 1.16.
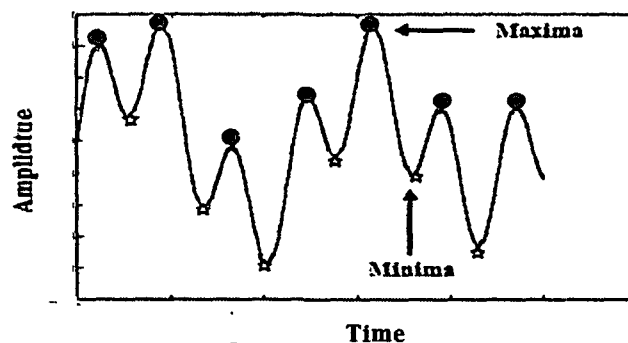
Figure 1.16 Maxima and minima of a signal

1.9.7. Modeling of dynamic systems

One of the applications of sophisticated monitoring of cardiac signals in ICU is uninterrupted generation of ECG while ECG is missing or corrupted due to malfunction of the ECG generation system. Synthesis of ECG can be achieved by modeling of the ECG from other measured data such as pressure signals. System identification is an efficient technique by which an approximate model of ECG can be developed. Further, the developed model can be used to synthesize ECG faithfully for monitoring.

System Identification is the field of building mathematical models of dynamic systems based on measured data [72]. It is achieved by adjusting the parameters of a developed model until the model output best matches with the measured output. As compared to mathematical models, system identification based models are less common and comprise of very small physical insight, however these are easier to construct as compared to mathematical models. Following are the basic models used for modeling of dynamic systems –

    a) Autoregressive (ARX) model

    b) Autoregressive moving average (ARMAX) model

    c) Transfer function model

    d) State space (SS) model

    a) Autoregressive (ARX) model

It is the simplest model that incorporates the stimulus signal. An autoregressive model can be defined as one of a group of linear prediction formulae that tries to predict an output of a system based on the previous outputs. The term autoregressive is used as the output is a function of its past values [73]. For parameter estimation of an ARX model, there are two methods for parameter estimation - the least squares (LS) method and the instrumental variable (IV) method. The parameters of the other model structures are estimated by the use of a prediction error method.

The LS method fits a model by minimizing the sum of square errors for estimating parameters. LS method is the special case of the prediction error minimization (PEM)

method described later in this chapter. A general ARX model structure is shown in Figure 1.17.

$$e(t)$$

$$u(t) \longrightarrow \boxed{B(q)} \longrightarrow \oplus \longrightarrow \boxed{\frac{1}{A(q)}} \longrightarrow y(t)$$
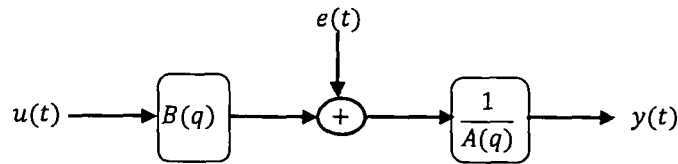
Figure 1.17 ARX Model Structure

In general Autoregressive Model structure may be written as

$$y(t) + a_1 y(t-1) + a_2 y(t-2)....a_{na} y(t-na) = b_1 u(t-nk) + ...b_{nb} \, u(t-nk-nb+1) + e(t)$$

$$(1.25)$$

The parameters *na* and *nb* are the orders of the ARX model and *nk* is the delay.

Where,

$y(t)$ – Output at time t

na – Number of poles

nb – Number of zeros plus 1

$n(k)$ – Number of input samples that occur before the input affects the output, also called the dead time in the system. For discrete systems with no dead time, there is a minimum 1–sample delay because the output depends on the previous input and $nk = 1$

The difference equation (1.x) can be written as –

$$A(q)y(t) = B(q) \, u(t-nk) + e(t) \tag{1.26}$$

'*q*' is s the shift operator. Specifically,

$$na: \qquad A(q) = 1 + a_1 \, q^{-1} .... a_{na} \, q^{-na} \tag{1.27}$$

$$nb: \qquad B(q) = b_1 + b_2 \, q^{-1} .... b_{nb} \, q^{-nb+1} \tag{1.28}$$

Where, $A(q)$ and $B(q)$ are unknown polynomials and $e(t)$ denotes the error. AR models do not support multiple-output continuous-time models. For multiple output continuous time models, state space models are used. The input *u* is the exogenous i.e. no extra input to the

system is available. The transfer function with input-output relationship obtained from ARX model may be employed for characterization of the dynamic performance of the system [73]. Polynomials estimation method in ARX model is the result of solving linear regression equations in analytical form. Hence, ARX models are recommended when model order is high. The disadvantage of ARX model is that system dynamics involve disturbances which can be reduced if a good signal to noise ratio is retained.

The equation of estimated output is given by [74]-

$$\hat{y}(t \mid t-1) = B(q)u(t) + [1 - A(q)]y(t) \tag{1.29}$$

b) Autoregressive moving average model with exogenous inputs (ARMAX)

Autoregressive moving average models with exogenous inputs (ARMAX) models include disturbance dynamics unlike AR models. These models are used for the estimation of the order and structure of system model using all relevant information such as the measurable input/output variables, internal variables of the system, measurable disturbance, and even the phenomenological information of the system. There are various estimating methods of ARMAX models such as pseudo linear regressive, correlation methods, subspace methods, etc. [75].

A general ARMAX model structure (Figure 1.18) is given by-

$$y(t) + a_1 y(t-1) + a_2 y(t-2)....a_{na}y(t-na) = b_1 u(t-nk) + ...b_{nb} u(t-nk-nb+1) +$$
$$e(t) + c_1 e(t-1) + ...c_{nc}e(t-nc)$$

$$\tag{1.30}$$
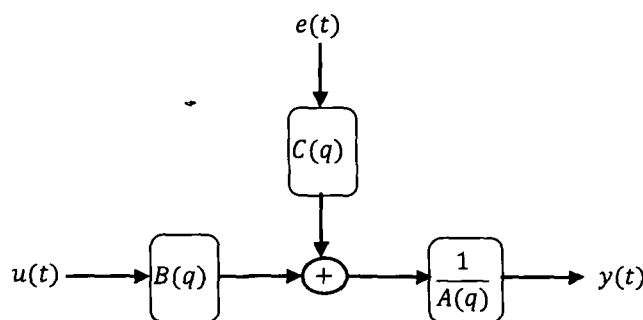
Figure 1.18 ARMAX model structure

Which may be written as –

$$A(q)y(t) = B(q) u(t - nk) + C(q)e(t) \tag{1.31}$$

Where,

$$na: \quad A(q) = 1 + a_1 q^{-1} \dots\dots\dots a_{na} q^{-na} \tag{1.32}$$

$$nb: \quad B(q) = b_1 + b_2 q^{-1} \dots\dots\dots b_{nb} q^{-nb+1} \tag{1.33}$$

$$nc: \quad C(q) = 1 + c_1 q^{-1} \dots\dots\dots c_{nc} q^{-nc} \tag{1.34}$$

The equation of estimated output is given by [74] –

$$\bar{y}(t \mid t - 1)C(q) = B(q)u(t) + [C(q) - A(q)]y(t) \tag{1.35}$$

c) Transfer function model

The linear polynomial models such as ARX model, ARMAX models etc are commonly used in control engineering as these models describe deterministic and stochastic part of the system separately. But in classical control engineering, deterministic part of the system is more preferred than the stochastic part of the system. Transfer function models can illustrate only the deterministic part of the system. These models can be used to depict both continuous time and discrete time systems. Transfer function models are good choice for single input single output (SISO) or multiple input and single output (MISO) physical systems. For multiple input and multiple output physical systems, state space models are well suited. These models can also be estimated in the frequency domain with frequency response function data.

Transfer function model for a continuous time and discrete time model is denoted by –

$$y(t) = G(s)u(t) \tag{1.36}$$

$$y(k) = G(z)u(k) \tag{1.37}$$

Where, $G(s)$ denotes the transfer function of continuous time system with input $u(t)$ and output $y(t)$ and $G(z)$ denotes the transfer function of discrete time system with input $u(k)$ and output $y(k)$.

d) State space model

State space models are used to describe the complex systems which are of higher order and consist of several parameters as input and output and a large number of measurements. State space structure is the powerful way to represent a system. The state-space representation is the most reliable Linear Time Invariant (LTI) model used for computer analysis.

In a linear time invariant sense, a discrete time state space model appears as the first order finite difference model [76] −

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1.38}$$

$$y(t) = Cx(t) + Du(t) \tag{1.39}$$

The system of first-order differential equations (equation 1.38) is known as the state equation of the system and equation 1.39 is called output equation. Here, $'u(t)'$ denotes input vector or control vector, $y(t)$ and $x(t)$ are output vector and state vector respectively. Matrices A, B, C are state matrix, input matrix, output matrix and matrix D is feed through or feed forward matrix. D is often selected as zero matrix for the purpose of simplicity. In other words, system is chosen not to have direct feed through.

The discrete state-space model is given by:

$$x(KT + T) = Ax(KT) + Bu(KT) \tag{1.40}$$

$$y(KT + T) = Cx(KT) + Du(KT) \tag{1.41}$$

Where, K is the sampling instant and T is the sampling interval.

Replacing $x(KT)$ by $x(k)$ above equations (1.40) and (1.41) reduce to,

$$x(k + 1) = Ax(k) + Bu(k) \tag{1.42}$$

$$y(k + 1) = Cx(k) + Du(k) \tag{1.43}$$

This equation can also be written as −

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_{n-1}(k+1) \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix} + \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix} [u(k)] \tag{1.44}$$

$$y(k) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x1(k) \\ x2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} \tag{1.45}$$

Where, $x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix}$ $A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix}$ $B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix}$

And $C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$

Here, $u(k)$ and $y(k)$ represent the system input and output at k^{th} sampling instant.

The transfer function of a LTI system in continuous time is given by –

$$y(t) = Hu(t) \tag{1.46}$$

The transfer function H in Laplace domain in terms of state matrices is given by –

$$\widehat{H}(s) = C(sI - A)^{-1}B + D \tag{1.47}$$

Where, I is the identity matrix, this equation can also be written as –

$$\widehat{H}(s) = \frac{b_0 s^m + b_0 s^{m-1} + \cdots + b_m}{s^n + a_1 s^{n-1} + \cdots + a_n} \tag{1.48}$$

And Z transform of this system is given by –

$$H(z) = \frac{b_0 z^m + b_0 z^{m-1} + \cdots + b_m}{z^n + a_1 z^{n-1} + \cdots + a_n} \tag{1.49}$$

The state space equation of multiple input and single output system is represented as the matrix of single input of output transfer function.

1.9.8. System identification

System identification techniques can exemplify the physiologic mechanisms by studying measured input – output data of a system, therefore this approach is an important aspect in the study of cardiovascular system [77]. System Identification studies have extensively used in the diverse fields of signal processing [78,79]. System identification has been used in the study of biomechanics to construct the models of joint dynamics [80], modeling of respiratory acoustics [81], processing of noisy biomedical signals [82], construction of functional biomedical images [83]. Subspace methods are employed to identify the state space models from short transients of ankle joint stiffness experiment using ensemble data [84]. A flow chart of the iterative system identification cycle is shown in Figure 1.19. System identification based approach deals with the selection of best parameters using the technique such as least-squares error, prediction error minimization etc.. If the obtained model is supposed to fulfill the intended application, predictions are made using the resultant model else the model class is revised and process is repeated [85].

Depending upon input-output relation, the identification of systems can be divided into two groups:

a) Static system identification

In this type of identification the output at any instant depends upon the input at that instant. These systems are described by the algebraic equations. The system is memory less and mathematically it is represented as-

$$y(n) = f[x(n)] \tag{1.50}$$

Where $y(n)$ is the output at the nth instant corresponding to the input $x(n)$.

b) Dynamic system identification

In this type of identification the output at any instant depends upon the input at that instant as well as the past inputs and outputs. Dynamic systems are described by the difference or differential equations. These systems have memory to store past values and mathematically represented as-

$$y(n) = f[x(n), x(n-1), x(n-2)\cdots\cdots\cdots y(n-1), y(n-2),\cdots\cdots] \tag{1.51}$$

Where $y(n)$ is the output at the nth instant corresponding to the input $x(n)$. The process of system identification deals with the determination of the complete characteristics of an unknown system from its output with the input being known. The characteristic of the unknown system can be determined in terms of its impulse response.
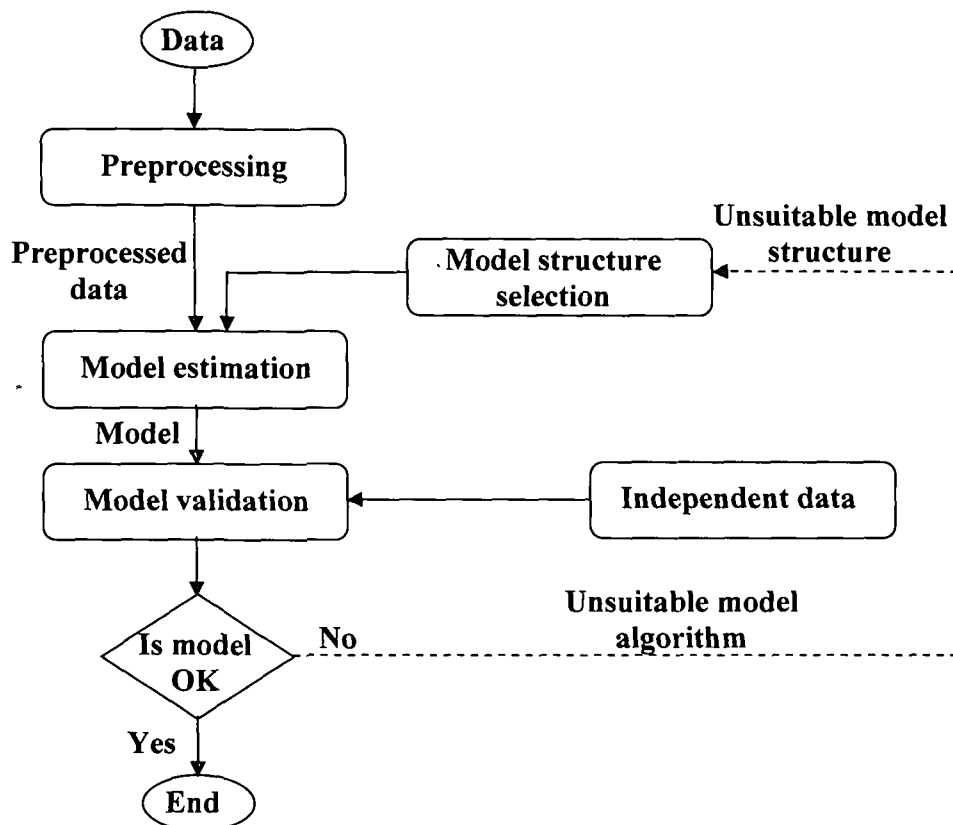
Figure 1.19 Flow chart of system identification cycle

1.9.9. Methods of system identification

Parameter estimation in dynamic systems is regarded as subpart of system identification to develop dynamic models from the measured data. There are the following methods used in system identification for estimation of model parameters –

a) Prediction error minimization (PEM) method

The objective of prediction error minimization (PEM) method is to build a predictor and compare its predictions with available data using some suitable measure [86]. The PEM

algorithm uses the measured data and reduces the prediction error to a large extent. The prediction error approach is shown in Figure 1.20.

The general properties of PEM method are –

a) This method requires prior information regarding the model structure such as type of model and orders of each term.

b) Biased parameter estimates are obtained owing to structural inconsistency. For different input excitations, different bias is shown.

c) Trial and error procedure with different orders is generally followed to find an economical model.

d) Generally, nonlinear equation should be solved to find an estimate.

Thus the input–output data to be used for identification are assumed to be generated in the following way [74]-

$$y(k) = G(q)u(k) + H(q)e(k) \qquad (1.52)$$

Where $e(k)$ is independent from $u(k)$. The terms $G(q)$ and $H(q)$ represents- deterministic and stochastic part of the system. If the system matrices of this algorithm depends on the parameter vector 'θ' we can define the model as - Given a finite number of samples of the input signal $u(k)$ and the output signal $y(k)$ and the order of the predictor

$$\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + K(y(k) - C\hat{x}(k) - Du(k)) \qquad (1.53)$$

$$\hat{y}(k) = C\hat{x}(k) + Du(k) \qquad (1.54)$$

The system matrices A, B, C, D, and constant K in this predictor are to be determined such that the output $y(k)$ approximates the output of 1.52.

b) Least square error (LSE)

LSE is used to fit a model by minimizing the sum of square errors for estimating parameters between the observed data and expected data [87].

Let us consider that a system with input $u(t)$ and output $y(t)$ at time 't' with N number of measurements such that $t = 0, 1, 2, \ldots\ldots N$. The output of the model $y(t)$ is given by –
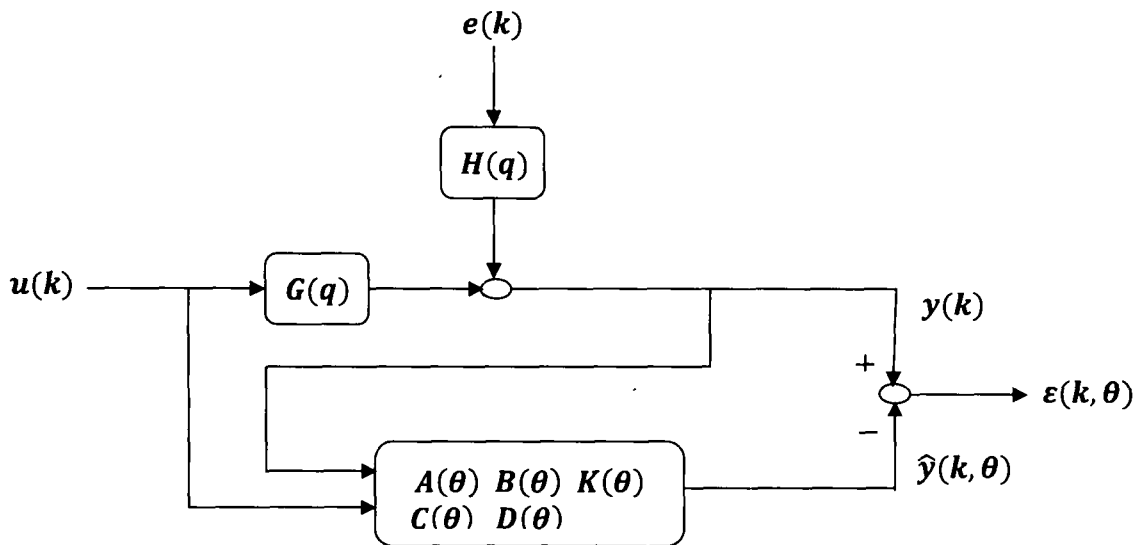
Figure 1.20 Model estimation by PEM method

$$y(t) = h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + \cdots\cdots + h_n u(t-n) \qquad (1.55)$$

The notations $h_0, h_1, h_2, \cdots\cdots h_n$ are the parameters of the model. The model represented is a moving average model with 'n' delays.

The error (E) between predicted and observed output given by the relation —

$$E = \left(\sum_{t=n}^{N} (y(t) - \hat{y}(t))^2 \right)^{\frac{1}{2}} \qquad (1.56)$$

$$E = \left(\sum_{t=n}^{N} (h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + \cdots\cdots + h_n u(t-n)) - \hat{y}(t))^2 \right)^{\frac{1}{2}} \quad (1.57)$$

The model in selected in such a way that the error E is minimized. The method of least squares is computationally suitable measure of fit.

c) Subspace system identification

System identification algorithms that estimate state space models based on subspace approximations are called subspace system identification. Subspace algorithms are useful for both time domain as well as frequency domain data. In subspace approximation, an iterative parameter optimization is not required as in classical prediction error minimization

techniques and also identification of multivariable systems is dealt in the simpler ways as scalar systems [88]. Subspace identification technologies are used to estimate the state space model of a dynamical system directly from input and output measurements as state space realizations. These methods are fast and can be used to estimate the system proficiently even in presence of noise [89]. These methods can estimate a state space model without any previous information of the system [90]. These models are used for the analysis, simulation, prediction, detection of abnormalities, training and validation of developed algorithms.

Mathematically, the subspace models are represented by following difference equations [91]

$$x_{k+1} = Ax_k + Bu_k + w_k \qquad \cdot \qquad (1.58)$$

$$y_k = Cx_k + Du_k + v_k \qquad (1.59)$$

With

$$E\begin{bmatrix} w_p \\ v_p \end{bmatrix} (w_q^T \ v_q^T) = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{pq} \geq 0 \qquad (1.60)$$

The matrices $Q \varepsilon R^{n \times n}, S \varepsilon R^{n \times l}$ and $R \varepsilon R^{l \times l}$ are the covariance matrices of the noise sequences w_k and v_k. The vectors $u_k \varepsilon R^m$ and $y_k \varepsilon R^l$ are the measurements at time instant 'k' of the process for 'm' inputs and 'l' outputs respectively. The vector x_k is the state vector of the process at discrete time instant 'k' and comprise of numerical values of n states. $v_k \varepsilon R^l$ and $w_k \varepsilon R^m$ are unobserved vector signals, v_k is called the measurement noise and w_k is called the process noise. The vectors v_k and w_k are assumed as zero mean, stationary, white noise vector sequences.

$A \varepsilon R^{n \times m}$ is the system matrix that describes the dynamics of the system, $B \varepsilon R^{n \times m}$ is the input matrix, $C \varepsilon R^{l \times n}$ is the output matrix while $D \varepsilon R^{l \times m}$ is the direct feed-through matrix. For continuous time systems D is often considered as zero but in case of discrete time systems, it is not zero due to sampling.

Since for deterministic systems, the process noise and measurement noise are identically zero, therefore, we can rewrite equation (1.58) and (1.59) as

$$x_{k+1} = Ax_k + Bu_k \tag{1.61}$$

$$y_k = Cx_k + Du_k \tag{1.62}$$

The matrix pair {A, C} is considered to be assumable which means that all modes in the system can be observed in the output y_k and thus can be identified.

In subspace system identification method, Kalman filter states are obtained from input – output data using QR factorization and singular value decomposition (SVD) tools of linear algebra. After these states are known, the identification method is converted to linear least square problem. The salient features of subspace algorithm are -

i.   The subspace algorithms include advantage over prediction error method as these methods have conceptual and algorithmic simplicity.

ii.   Subspace algorithms are faster than prediction error methods because these do not include iterative approach. These algorithms are robust as these include well known algorithms from linear algebra.

iii. A reduced order model can be obtained directly using these methods without having the need to get a higher order model first and then reducing it to a lower order model as done in classical prediction error approach.
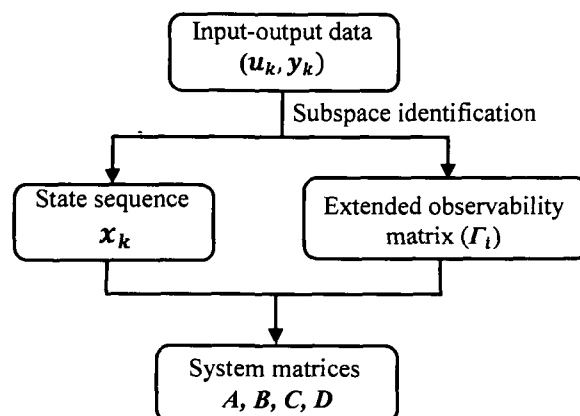
Figure 1.21 Overview of deterministic subspace identification procedure [92]

An overview of deterministic subspace system identification algorithm is shown in Figure 1.21.

Following the theorem of [93], the linear state space equations (1.61) and (1.62) can be written in matrix form as below –

$$X_f = A^i X_p + \Delta_i U_p \tag{1.63}$$

$$Y_f = \Gamma_i X_f + H_i U_f \tag{1.64}$$

Where, $U_p, U_f, Y_f, H_i, \Delta_i$ are past and future inputs, Y_f is future output, H_i is the lower block triangular Toeplitz matrix and Δ_i is the reverse extended observability matrix. X_p and X_f denote past and future state sequences.

Subspace algorithm involves the following steps [92] –

Step 1: Calculate the oblique projections:

$$Oi = \frac{Y_f}{U_f} W_p \tag{1.65}$$

$$Oi-1 = \frac{Y_f-}{U_f-} W_{p+} \tag{1.66}$$

Where, U_f and Y_f are the future input –output matrices and W_p is Hankel matries containing past values.

Step 2: Calculate the singular value decomposition (SVD) of the weighted oblique projection:

Let us consider that two weighing matrices $W_1 \varepsilon R^{li \times li}$ and $W_2 \varepsilon R^{j \times j}$ such that W_1 is of full rank and W_2 obeys that $rank\,(W_p) = rank\,(W_p, W_2)$ where W_p block Hankel Matrix containing past inputs and outputs.

SVD is obtained from the following relation

$$W_1 O_i W_2 = USV^T \tag{1.67}$$

Where, $T \varepsilon R^{n \times n}$ is an arbitrary non-singular matrix representing a similarity transformation. S may be partitioned into U_1 and S_1. The number of singular values in S defines the order of the system.

Step 3: Now we define extended observability matrices Γ_i as below –

$$\Gamma_i = W_1^{-1} U_1 S_1^{1/2} \tag{1.68}$$

Where, $\qquad \Gamma_i \overset{def}{=} \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{pmatrix} \quad \varepsilon \, R^{n \times j} \tag{1.69}$

The subscript 'i' denotes the number of block columns.

And $\quad \Gamma_{i-1} = \underline{\Gamma_i}$

Where, $\underline{\Gamma_i}$ is the extended observability matrix without last 'l' rows.

Step 4: The state sequences X_i and X_{i+1} are defined as

$$X_i = \Gamma_i^+ O_i \tag{1.70}$$

And $\qquad X_{i+1} = \Gamma_{i-1}^+ O_{i-1} \tag{1.71}$

Step 5: The state variables A, B, C and D are obtained by solving the following set of linear equations

$$\begin{pmatrix} X_{i+1} \\ Y_{i/i} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} X_i \\ U_{i/i} \end{pmatrix} \tag{1.72}$$

1.9.10. Stability analysis of transfer function models

Stability of a system implies that a small change in input does not result in large change in system behaviour. If a system is applied with bounded input and produces an unbounded response, the system is said to be unstable. In terms of measured input-output cardiac model, stability indicates change in the behaviour of the heart when obstructions to flow of

affect the pumping action of the heart which is displayed as abrupt change in characteristic waves of ECG, ABP and CVP signals. These changes can be reflected in terms of stable/ unstable response and location of poles of transfer functions of the model.

a) Step response and pole zero plot

The step response describes the transient response of a system. It describes the system behaviour with respect to disturbances and qualitatively gives an idea about the system stability.

Let us consider the step function

$$
\begin{aligned}
f(t) &= 0 \quad for \ t < 0 \\
&= A \quad for \ t > 0
\end{aligned}
$$
(1.73)

and let $y(t)$ be the corresponding output, which is called the step response. Step function whose height is unity is called unit step function.

The response of an LTI system to a unit step is given by [Roberts, M. J, 2007] –

$$
y(t) = h(t) * u(t) = \int_{-\infty}^{\infty} h(\tau) \, u(t - \tau) \, d\tau = \int_{-\infty}^{\infty} h(\tau) d\tau
$$
(1.74)

Therefore, when an LTI system is stimulated by a step input, the response of the system at any time t is defined as the integral of impulse response.

The response of a system to a complex exponential e^{st} is given by –

$$
y(t) = h(t) * e^{st} = \int_{-\infty}^{\infty} h(\tau) \, e^{s(t-\tau)} \, d\tau = \int_{-\infty}^{\infty} h(\tau) e^{-st} d\tau
$$
(1.75)

Where s is any complex constant. The integral $\int_{-\infty}^{\infty} h(t) e^{-st} dt$ is called the Laplace transform of $h(t)$.

b) Pole-zero plot

The poles of the transfer function of a system provide an insight into the natural response of the system. The stability of a linear system can be determined from the location of poles in s-plane.

From equation (1.53), solving by the Laplace Transform method we get,

$$Y(s) = \frac{K(s+z_1)(s+z_2).....(s+z_m)}{(s+p_1)(s+p_2)....(s+p_n)}$$
(1.76)

where $m < n$ (for limited response at high frequency).

The constants z_i are called the zeros of the transfer function or signal, and p_i are the poles. In complex plane, when viewed in the complex plane, the magnitude of $Y(s)$ will go to zero at the zeros, and to infinity at the poles.

The location of poles is viewed in complex plane to determine the system stability. A system is stable if and only if all of the poles occur in the left half of the complex plane. Similarly, the equilibrium state of a discrete-system with constant input is stable if and only if all poles of $H(z)$ have absolute values less than one; that is, the poles are all inside the unit circle of the z-plane. Figure1 22 (a-b) shows the pole-zero plots of a stable and unstable system respectively. If any of the poles lie in the right half of s-plane or outside the unit

(a)   (b)
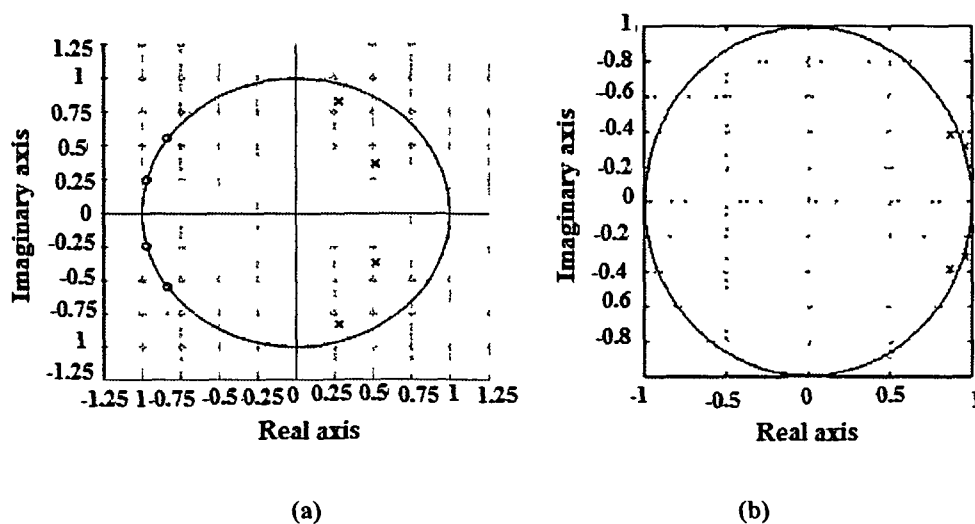
Figure 1.22 Pole-zero plots of (a) a stable system and (b) an unstable system

circle in z-plane, then with increasing time, they give rise to the dominant mode, that results in increase in transient response monotonically or oscillations with increasing amplitude will occur. This represents an unstable system. Therefore, closed-loop poles in the right-half of the s-plane are not permissible in the usual linear control system. If the closed loop poles lie to the left-half of the $j\omega$ axis, any transient axis, any transient response eventually reaches equilibrium [94]. The system stability criteria for stable, unstable, marginally stable system using pole zero plots and system response are shown in Table 1.8. It is clear from Table 1.8, that the system is stable if all of the poles have negative real parts and lie in the left half of the complex plane.

c) Routh's stability criterion

The stability of ECG model can also be analyzed by Routh's stability criterion. This approach is applicable to polynomial with only a finite number of terms. According to Routh's stability criterion, coefficients of the characteristic equation can be used to provide the information regarding the stability of the system [94]. Most of systems have transfer function of the form –

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \ldots\ldots\ldots + b_{m-1} s + b_m}{a_0 s^n + a_1 s^{n-1} + \ldots\ldots\ldots + a_{n-1} s + a_n} = \frac{B(s)}{A(s)} \tag{1.77}$$

Where, $a_0, a_1, \ldots\ldots a_n$ and $b_0, b_1, \ldots\ldots b_m$ are coefficients and $m \leq n$.

The procedure in Routh's stability criterion is as follows –

i. The s-domain transfer function is written in the following form-
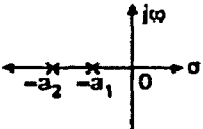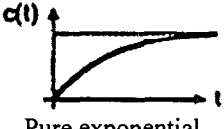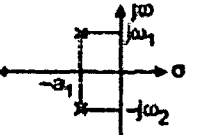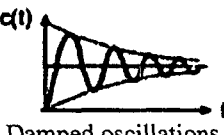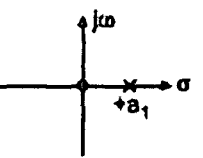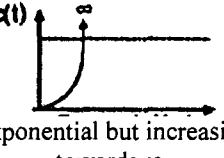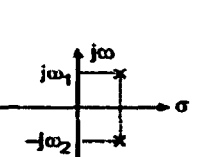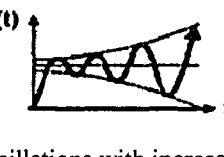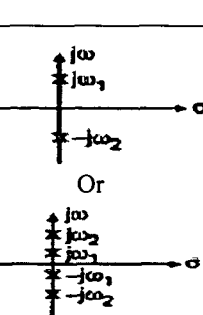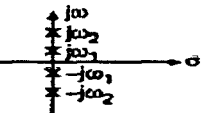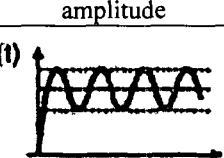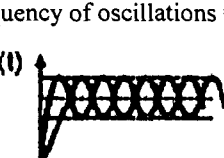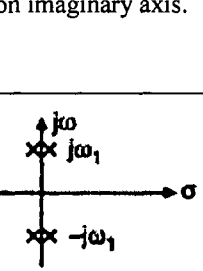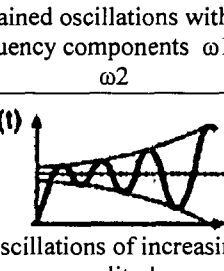
$$a_0 s^n + a_1 s^{n-1} + \ldots\ldots\ldots + a_{n-1} s + a_n = 0$$

Where, the coefficients are real quantities and $a_n \neq 0$, i.e. any zero root has been removed.

ii. If any of the coefficients are zero or negative in presence of at least one positive coefficient, there is a root or roots that are imaginary or that have positive real part. Therefore, in such a case the system is not stable.

---

iii.   For absolute stability, all the coefficients must be positive and all the coefficients must be present.

Table1.8 Closed loop poles, step response and stability [95]

Nature of closed loop poles	Location of closed loop poles in s-plane	Step response	Stability condition
Real –ve i.e. in L.H.S of s-plane		Pure exponential	Absolutely stable
Complex conjugate with –ve real part in L.H.S. of s-plane		Damped oscillations	Absolutely stable
Real positive i.e. in R.H.S. of s-plane (any one closed loop pole in R.H.S of s-plane irrespective of number of poles in left half of s-plane)		Exponential but increasing towards ∞.	Unstable
Complex conjugate with positive real part i.e. in R.H.S. of s-plane		Oscillations with increasing amplitude	Unstable
Non-repeated pair on imaginary axis without any pole in R.H.S. of s-plane	Or Two non-repeated pairs on imaginary axis.	Frequency of oscillations = ω1 Sustained oscillations with two frequency components ω1 and ω2	Marginally or critically stable system.
Repeated pair on imaginary axis without any pole in R.H.S. of s-plane		Oscillations of increasing amplitude	Unstable

1.9.11. Artificial neural network

Artificial Neural Networks (ANNs) are non-linear mapping structures inspired by the function of the human brain. ANNs can identify and learn correlated patterns between input data sets and corresponding target data. These can be used to predict the result of new independent data after training. ANN mimic the learning process of human brain, therefore these can be used to resolve the problems of non-linear and complex data. They are powerful tools for modeling, especially when the underlying data relationship is unknown These networks perform well with multiple inputs and multiple outputs. This capability of ANNs is well suited for modeling of biomedical signals which are non-linear and affected by noise most of the time.

ANN has been used extensively in biomedical applications, such as signal reconstruction [96-99], modeling of electrical activity of human heart [100] and EEG [101]. These are also used in the diagnosis of breast cancer [102, 103], prediction of the occurrence of coronary artery disease [104] and diagnosis of Down's Syndrome in unborn babies [105].

Different ANN paradigms such as MLP and RBF (multilayer perceptron and radial basis function network have been employed for the diagnosis of pathological conditions in serum electrophoresis [106], classification of PVC beats in ECG signals [16, 107], segmentation and classification of multi-spectral MRI images[108]. In addition, RBF have been used for the diagnosis of myocardial infarction [109]. Application of ANN in modeling of ECG using ABP and CVP signals will be discussed in chapter 5.

ANNs consist of a large number of simple, highly interconnected processing elements called neurons. An artificial neuron computes a weighted sum of its N input nodes, x_i, where, $i=1, 2, 3\ldots\ldots$N and generates an output depending upon the activation function. The function of a neuron is shown in Figure 1.23. Mathematically, it can be represented as

$$net = \sum_{i=1}^{N} w_i \ x_i - \mu_i \qquad (1.78)$$

$$y = f(net)$$

Where, y is the output, x_i is the ith input to the node, and w_i is the weight associated with input x_i. The terms μ is the threshold or bias of the node and f is the transfer function called activation function. Output of the neuron depends on these activation functions. There are various activation functions such as step, sigmoid, linear, Gaussian etc. Some commonly used activation functions are shown in Figure 1.24. These activation functions decide the output of a neuron. The learning capability of an artificial neuron depends upon the adjustment of weights according to learning algorithm.
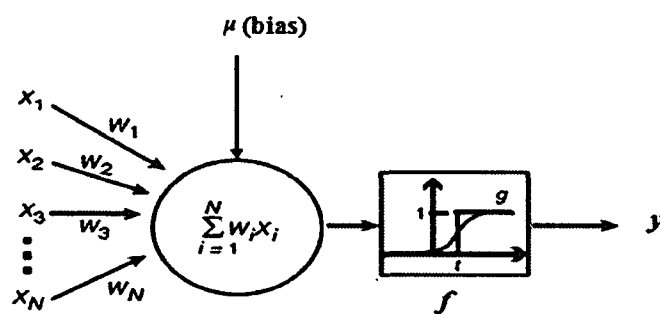
Figure 1.23 Function of an artificial neuron

These neurons are organized into a sequence of layers with full or random connections between the layers. Such an arrangement is called artificial neural network as shown in Figure 1.25. An artificial neural network consists of three layers – input layer, one or more hidden layers and output layer. The input layer is a buffer that presents the data to the network. The top layer is called output layer which gives the output response to the given input. Hidden layers are the real classifiers that work for the classification algorithm of ANN.

$$a^{linear} = a$$

(a) Linear

$$a^{sigmoid} = \frac{1}{1 + e^{-n}}$$

(b) Sigmoid

$$a^{radbas} = e^{-n^2}$$

(c) Radial basis function
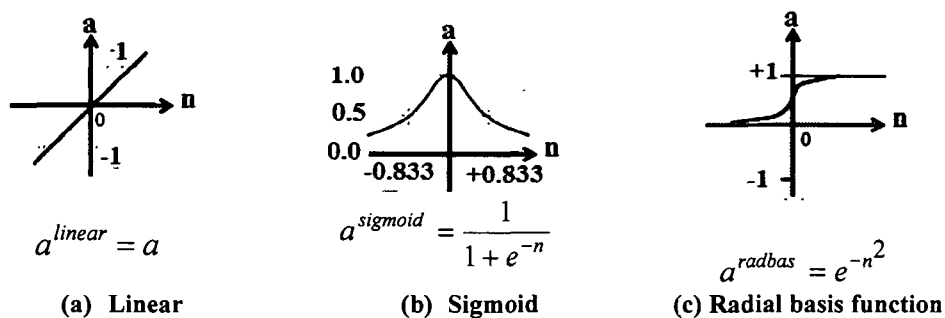
Figure 1.24 Activation functions
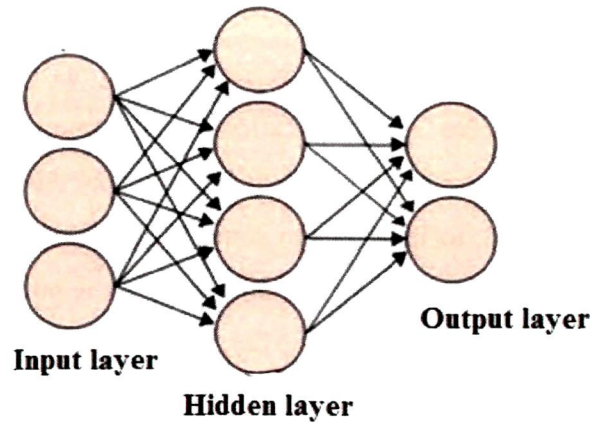
Figure 1.25 Architecture of a neural network

1.9.11.1.   Neural network architectures

There are various types of architectures of neural networks. Following two types of neural network architectures are most commonly used –

a) Feed forward networks

These networks employ unidirectional flow of information along connecting pathways i.e. information flows from input layer to hidden layer and then to output layer and there is no feedback from input to output. In these networks, the output of any layer does not affect the previous layer output. A multilayered feed forward neural network is shown in Figure 1.26.
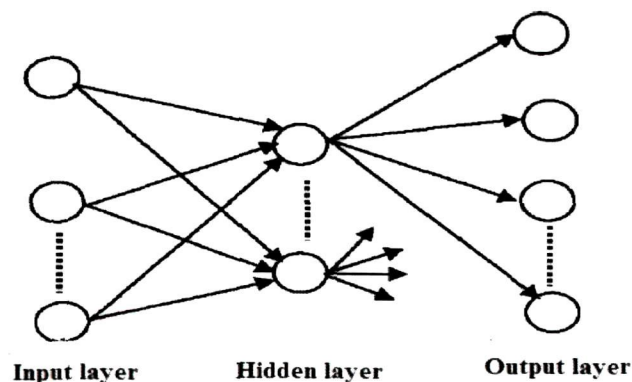
Figure 1.26 A multilayered feed forward neural network

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

b) Recurrent neural networks

Recurrent neural networks have different network architectures than feed forward neural networks. In these networks, there is a feedback loop from output to input. Therefore, these networks may comprise of one or more layers with feedback connections (Figure 1.27) or output of a neuron is feedback to its input. In some applications, the activation values of neurons acquire equilibrium state or the activation values of the output neurons may change such that the dynamic behaviour constitutes the output of the network.

Input layer   Hidden layer   Output layer
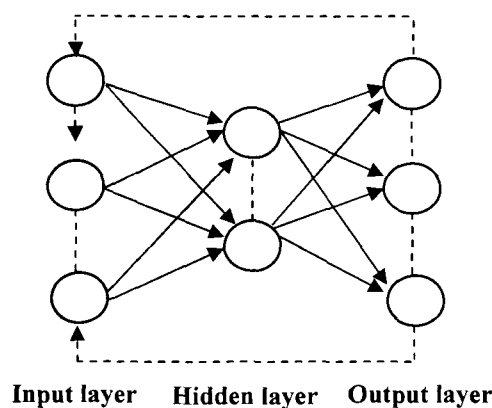
Figure 1.27 Recurrent neural network

ANNs are capable of disease diagnosis, modeling [100, 101] and reconstruction [96-99] of physiological time series data, because, once a network is trained for a particular pattern of the signal, it generates a pattern on testing with independent data which is more close to the input to which it was learnt.

1.9.11.2.   Learning paradigm

Learning in a neural network is accomplished through an adaptive procedure, known as *'learning rule'*. The weights of the network are incrementally adjusted so as to improve the performance over time. The basic learning rules are the supervised learning, unsupervised learning and reinforcement learning.

a) Supervised learning

In supervised learning, an input vector is presented at the inputs together with a set of desired responses, one for each node, at the output layer. It involves the comparison of

actual response to the desired response. If the actual response differs from the desired response, an error signal is generated and network weights are adjusted so as to achieve the desired response.

b) Unsupervised training

In unsupervised learning, there are no explicit targets associated with each input and there is no feedback path. In this type of learning, neural network organizes the received input in to some categories during training. When an external independent input is applied to the network, it generates an output depending upon the class to which the input belongs. In general, unsupervised learning performs the same task as an auto-associated network by compressing the information from the inputs.

c) Reinforcement leaning

In this type of learning, the complete detailed information is not available as the target output. This method of learning differs from supervised learning, where there is a target value for each input value. It is one kind of learning in which some feedback from the environment is given but the feedback signal is only evaluative not instructive.

1.9.11.3.   Radial basis function network

There are various types of neural networks and their architectures such as multi-layer perceptron, radial basis networks, learning vector quantization etc. Out of these ANN structures, RBF is capable of fast learning than back propagation networks and MLP. RBF networks are least affected with non-stationary inputs and these networks have been successfully used in biomedical applications [106-109] and in the prediction of time series data [110,111]. We have used RBF network for modeling and synthesis of ECG from ABP and CVP signals as discussed in Chapter 5.

The Radial Basis Function (RBF) network is a three-layer feed-forward network that uses non-linear transfer function for hidden layer and a linear transfer function for the output units [112]. The RBF networks require more neurons compared to standard feed forward networks. The performance of these networks is better when trained with a large amount of data. These network employ supervised learning.

A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, so that $\varphi(x) = \varphi(\|x\|)$ or alternatively on the distance from some other point μ called a center, so that $\varphi(x,\mu) = \varphi(\|x - \mu\|)$. Any function '$\varphi$' that satisfies the property $\varphi(x) = \varphi(\|x\|)$ is a radial function.

A radial basis function is typically used to build up function approximations of the form

$$z(x) = \varphi(\|x - \mu\|)$$ (1.79)

Where, x is an n dimensional vector, 'μ' is an n-dimensional vector called the centre of the radial basis function. The norm $(\|.\|)$ denotes Euclidean distance, and is a univariate function also referred as profile function, defined for positive input values. The model is built up as a linear combination of 'N' radial basis functions with 'N' distant centers. Given an input vector 'x', the output of the RBF network is the activity vector '$\hat{y}$' given by

$$\hat{y}(x) = \sum_{j=1}^{N} w_j \, \varphi\left(\|x - \mu_j\|\right)$$ (1.80)

$$\hat{y}(x) = \sum_{j=1}^{N} w_j \, z_j \, (x)$$ (1.81)

Where, 'w_j' is the weight associated with jth radial basis function centered at 'μ_j' and $z(j) = \varphi(\|x - \mu_j\|)$. The output '$\hat{y}$' approximates a target set of values denoted by 'y' as a sum of N-radial basis functions, each associated with a different center 'μ_j' and weighted by an appropriate coefficient 'w_j'. This type of approximation schemes are mainly used in time series prediction and control of nonlinear systems those display sufficiently simple chaotic behaviour. The sum can also be elucidated as a rather simple single-layer type of artificial neural network called a radial basis function network, with the radial basis functions taking on the role of the activation functions of the network.

A commonly used the transfer function for radial basis networks is Gaussian function which in case of scalar input is given –

$$h(x) = \exp^{-}\left(\frac{(x-\mu)^2}{r^2}\right)$$

(1.82)

Where, 'μ' is the centre of radial basis function and 'r' is the radius. The Gaussian RBF monotonically decreases with its distance from centre. A multiquadric RBF in the case of scalar input is given by –

$$h(x) = \frac{\sqrt{r^2 + (x-\mu)^2}}{r}$$

(1.83)

A radial basis function network is shown in Figure 1.28. In RBF networks, the hidden layer has radial basis neurons, and calculates its weighted inputs with distance, and its net input. The output layer has linear neurons and calculates its weighted input and its net inputs. Both layers have biases. Initially the radial basis layer has no neuron. The following steps are repeated until the network mean squared error falls below the predefined goal or the maximum number of neurons are reached –

a)   The network is simulated.

b)   The input vector with the greatest error is found.

c)   A radial basis neuron is added with weights equal to that vector.

d)   The linear layer weights are redesigned to minimize the error.
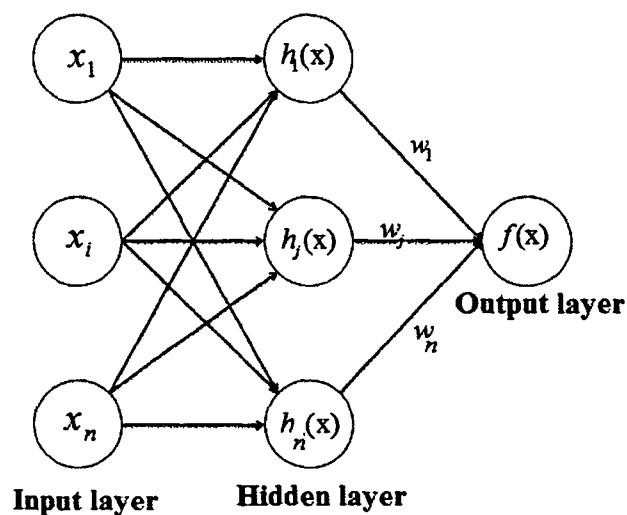
Figure 1.28 Radial basis function network

If $(x_1, x_2, x_3 \dots \dots x_n)$ are the inputs and $(w_1, w_2, w_3 \dots \dots w_n)$ are associated weights, then the output of an RBF with single hidden layer and Gaussian activation function (Figure 1.28) is given by –

$$f(x) = \sum_{j=1}^{n} w_j h_j(x)$$

(1.84)

1.9.12. Prior works on modeling and synthesis of ECG

Previous works on modeling and synthesis of ECG include a method for generation of RR-tachograms [113], a dynamical model that mimics the real ECG signal of a normal person [114], use of this model to generate realistic ECG, BP, respiration signals [115] and modeling an arbitrary ECG without in band noise [116] are also suggested. Further a three-dimensional dynamic model is proposed for ECG modeling that is generalized to model maternal and fetal ECGs [117]. Models for generation of multi-lead ECGs [118] and simulation of abnormal rhythms [119] are also reported in literature. Gaussian wave-based state space model is used for generating synthetic ECGs as well as separate characteristic waves (CWs) such as the atrial and ventricular complexes [120] and extended Kalman filter based dynamic algorithm is proposed for tracking the ECG characteristic waveforms [121]. Mathematical modeling of electrical activity of heart [122] and computer simulation of qualitative ECG are also suggested [123].

The inherent shape of Hermite Basis Functions (HBF) bearing resemblance to ECG signals is used for shape determination of ECG [124-125]. In addition to this, piece wise modeling of ECG is also suggested [126]. A modified Zeeman model using radial basis network is proposed for ECG modeling but this model is able to generate single cycle of ECG [127]. Methods including Gaussian Combination Model [128], Hilbert transform [129], Hidden Markov models [130] and data flow graph method [131] are also known in literature for ECG modeling.

Although a considerable amount of research has been carried out for ECG modeling and synthesis [113-131], however, a parametric ECG model based on measured phenomenological cardiac data such as ABP and CVP has not been attempted so far.

In physionet challenge 2010 [132], reconstruction of ECG, ABP, respiratory, fingertip plethysmogram (PLETH) signals using Artificial neural network (ANN) [96-99, 134] and wavelet based approach [138] has been suggested. ANN has been found to be a promising technique for ECG reconstruction as suggested in [96-99, 134], however the technique needs at least one ECG signal from other leads.

1.10. Preprocessing techniques

The ECG electrodes as well as the pressure transducers pick up various types of noise from the electronic circuit, power-line and environment which corrupts the actual signal to a great extent. Signals available in the database are raw signals; therefore the signals must be filtered before processing. We have used two types of filters for this application – moving average filter for remove high frequency noise and a Butterworth high-pass filter to remove low frequency noise.

1.10.1. Moving average filter

It is a smoothing filter to remove random noise in the signal which attenuates the high frequency components. It involves the computation of temporal statistics filtered using samples of signals during a certain window length. The window is moved to obtain filtered output at different time locations.

General equation of a moving average filter can be written as –

$$y(n) = \sum_{k=0}^{N} b_k \, x(n-k)$$

(1.85)

Where, x and y are the input and output signals respectively and N is the number of points used for moving average, called the order of the filter. The b_k values are the filter coefficients. A 20 point moving average filter is implemented in removing the high frequency components present in 7 records of ECG, ABP and CVP signals in section 4.1. of Chapter 4.

1.10.2. Butterworth high-pass filter

The Butterworth filter is perhaps the most commonly used frequency-domain filter due to its simplicity and the property of a maximally flat amplitude response in the pass band. A

Butterworth filter is specified by its cut-off frequency and order. As the order increases, the filter response becomes more flat in the pass-band, and the transition to the stop-band becomes faster or sharper. Selection of cut-off frequency (*fc*), order (*N*) of the filter is crucial because an improper value of *fc* and *N* causes distortion of the signal.

The magnitude squared function for a continuous time Butterworth high-pass filter is given by –

$$|H(\omega)|^2 = \frac{1}{1 + \left(\dfrac{\omega_c}{\omega}\right)^{2N}}$$

(1.86)

with ω normalized to the range (0, 2 π) for sampled or discrete time signals; in such a case, the equation is valid only for the range (0, π). The cut-off frequency ω_c should be specified in the range (0, π). As the order N increases, the filter characteristics become sharper i.e. remain close to unity over the pass band.

If the discrete Fourier transform (DFT) is used to compute the Fourier transforms of the signals being filtered, above equation 1.86 may be modified to –

$$|H(k)|^2 = \frac{1}{1 + \left(\dfrac{k_c}{k}\right)^{2N}}$$

(1.87)

Where k is index of DFT array standing for discretized frequency. With K being number of points in the DFT array, k_c is the array index corresponding to the cut-off frequency ω_c i.e.

$$k_c = K \frac{\omega_c}{\omega_s}$$

(1.88)

The equation above is valid for $k = 0, 1, 2.........K/2$ with the second half over $(K/2 + 1, K - 1)$ being a reflection of first half. We have used Butterworth highpass filter to remove low frequency noise such as base line drift in ECG as well as ABP signals in record 100 in section 2.1.2.2 and mgh001 in section 3.1.2.1 of Chapter 2 and Chapter 3 respectively.

1.10.3. Detrending

A non-linear system may be linearized to obtain a linear model in the selected operating range at equilibrium point. The resultant linear model may be used to estimate the non-linear system near to the point of linearization in the selected operating range. Therefore,

detrending or offset nullification is defined as a method of removal of constant levels from a dataset to bring the signal to the equilibrium

If $\bar{u}$, $\bar{y}$ and $\bar{x}$ represent the input, output, and state of a model at the equilibrium point and $\tilde{x}$ is the deviation of the state sequence from the equilibrium value x, the linearized state space model can be denoted by –

$$\tilde{x}(k+1) = A\tilde{x}(k) + B(u(k) - \bar{u}) \qquad (1.89)$$

$$y(k) = C\tilde{x}(k) + \bar{y} \qquad (1.90)$$

The unknown offsets $\bar{u}$, and $\bar{y}$ can be dealt by the following methods in system identification [92] –

The offsets $\bar{u}$, and $\bar{y}$ are estimated as the mean of the measured sequences and given by the following relation –

$$\bar{u} = \frac{1}{N}\sum_{k=1}^{N}u(k) \qquad (1.91)$$

$$\bar{y} = \frac{1}{N}\sum_{k=1}^{N}y(k) \qquad (1.92)$$

These offsets values of u and y at equilibrium are subtracted from the input and output sequences u and y before starting the system identification process.

1.11. Database of cardiovascular signals

Analysis, modeling and synthesis of cardiovascular signals require a large amount of data obtained from the subjects under normal and abnormal conditions. This stored data enables the researchers to develop and validate their algorithms.

1.11.1. PhysioBank database

PhysioBank [135] is the collection of digital recordings of complex physiological signals. It comprises of several multi-parameter database such as MIT-BIH arrhythmia database, Fantasia database, MIT-BIH Polysmographic database, MGH/MF waveform database etc. which include recordings of cardiovascular, cardiopulmonary and other physiological

signals. These signals are acquired from patients under various pathological conditions which include arrhythmia, neurological disorders etc. The cardiovascular signals consist of ECG signals from multiple leads, arterial blood pressure, central venous pressure, respiratory, pulmonary artery pressure etc. It is revealed that these signals comprise of valuable information about the cardiovascular function which may not be possible to extract using conventional methods of analysis such as forecasting of sudden cardiac death in ambulatory patients. These cardiovascular signals are freely available and therefore, it is a source of attractions for the researchers in biomedical field to develop new algorithms for study of different heart diseases [135]. ECG signals in all these databases are annotated by two or more cardiologists independently and computer readable annotations were prepared after resolving the disagreements among the cardiologists. Annotations are the labels that describe the locations of the event in an ECG record. This facilitates the researchers to validate their algorithms developed using normal as well as abnormal ECG signals.

1.11.1.1.   Physiobank annotations

Physiobank database signals have standard set of annotation codes defined for ECG recordings. These annotation codes are defined for beat as well as non-beat annotations. Both types of annotation codes are listed in Table 1.9.

1.11.1.2.   MIT-BIH arrhythmia database

The MIT-BIH Arrhythmia Database is the collection of 48 records extracted from two channel ambulatory ECG recordings. Each recording is of 30 minutes duration. These recordings are digitized at 360 samples per second.

ECG signals are described by- a text header file (.hea), a binary file (.dat) and a binary annotation file (.atr). Header file consists of detailed information such as number of samples, sampling frequency, format of ECG signal, type of ECG leads and number of ECG leads, patients history and the detailed clinical information. In binary data file, the signal is stored in 212 format which means each sample requires number of leads times 12 bits to be stored and the binary annotation file consists of approved beat annotations. In most records, the upper signal is from modified limb lead II (MLII) and the lower signal is usually from modified lead V1 (occasionally V2 or V5, and in one instance V4). This

configuration was routinely used by the BIH Arrhythmia Laboratory as normal QRS complexes are usually prominent in the upper signal. Records 102 and 104 do not have lead II signal because signal acquisition was not possible for this lead due to surgical dressings on the patients, hence modified lead V5 was used for the upper signal in these records [136].

ECG signals from upper lead of 44 records of this database are used for the development and testing the performance of ECG feature extraction and PVC detection algorithms described in chapter 2.

1.11.1.3. Fantasia database

Fantasia database is the collection of 40 recordings obtained from 20 young (21-34 years of age) and 20 elderly (68-85 years of age) rigorously screened healthy subjects [137]. The records from old subjects are named as *f1o01, f1o02........f1o10* and *f2o01, f2o02........f2o10* while the recordings from young subjects are labeled as *f1y01, f1y02........f1y10* and *f2y01, f2y02........f2y10*. Each record consists of ECG and respiratory signals. The record *f2o01, f2o02........f2o10* and *f2y01, f2y02........f2y10* consists of invasive arterial blood pressures signals along with ECG and respiratory signals. Each signal is of 120 minutes duration. These signals were recorded from subjects in resting state in sinus rhythm while watching the movie Fantasia to keep the subjects in awake state. The continuous ECG, respiration, and blood pressure signals were sampled at 250 Hz. Each ECG beat is annotated by visual inspection. This database also has three files, .dat, header and annotation files. The dat file comprise of signals. The header file has the information about age, sex, length of signal, sampling frequency, type of signals and calibration units required for conversion of signals from raw units to physical units.

ABP signals from 14 records of this database are used for testing the performance of ABP feature extraction described in chapter 3.

1.11.1.4. MIT-BIH polysmographic database

The MIT-BIH Polysomnographic Database consists of 18 recordings of multiple physiologic signals during sleep [138]. The data are collected from patients in Beth Israel

Table 1.9 Beat and non-beat annotation codes for ECG signal [135]

Beat annotations		Non-beat annotations	
Code	Description	Code	Description
N	Normal beat	[	Start of ventricular flutter/fibrillation
L	Left bundle branch block beat	!	Ventricular flutter wave
R	Right bundle branch block beat	]	End of ventricular flutter/fibrillation
B	Bundle branch block beat (unspecified)	x	Non-conducted P-wave (blocked APC)
A	Atrial premature beat	(	Waveform onset
a	Aberrated atrial premature beat	)	Waveform end
J	Nodal (junctional) premature beat	p	Peak of P-wave
S	Supraventricular premature or ectopic beat (atrial or nodal)	t	Peak of T-wave
V	Premature ventricular contraction	u	Peak of U-wave
r	R-on-T premature ventricular contraction	`	PQ junction
F	Fusion of ventricular and normal beat	'	J-point
e	Atrial escape beat	^	(Non-captured) pacemaker artifact
j	Nodal (junctional) escape beat	|	Isolated QRS-like artifact
n	Supraventricular escape beat (atrial or nodal)	~	Change in signal quality
E	Ventricular escape beat	+	Rhythm change
/	Paced beat	s	ST segment change
f	Fusion of paced and normal beat	T	T-wave change
Q	Unclassifiable beat	*	Systole
?	Beat not classified during learning	D	Diastole
		=	Measurement annotation
		"	Comment annotation
		@	Link to external data

Hospital for evaluation of chronic obstructive sleep apnea syndrome. Each record consists of ECG signal, an invasive blood pressure signal (measured using a catheter in the radial artery), an EEG signal, and a respiration signal. ECG signal in each record is annotated while EEG and respiration signals are annotated with respect to sleep stages and apnea. Some records consists of six or seven channel recordings which include respiration signals, EOG (Electrooculogram) , EMG (Electromyogram), cardiac stroke volume (SV) signal and earlobe oximeter (SO2) signal. Each record consists of four files - .dat file, header file and two annotation files. The dat file comprise of signals, header file has information about age, gender, length of recording, weight of subject (in Kg), types of signals. The header file also consists calibration constants which are required for conversion of signals from raw units to physical units. The annotation files contain beat and sleep apnea annotations. ABP signals from 15 records are used for testing the performance of ABP feature extraction algorithm described in chapter 3.

1.11.1.5. MGH/MF waveform database

The Massachusetts General Hospital/Marquette Foundation (MGH/MF) Waveform Database is a comprehensive collection of electronic recordings of hemodynamic and electrocardiographic waveforms [139]. The signals in the database comprise of a broad spectrum of physiologic and pathophysiologic states. The signals in the database are sampled at 360 samples/second. Each record consists of .dat file, .hea file and .ari file. Header file consists of information regarding the types of signals and type of leads in case of ECG signal, sampling interval, sampling frequency, duration of the signal and units. It comprises of the information to convert the recorded signals from raw units to physical units. The ari file consists of ECG annotations whereas dat file comprise of 8 signals. The recorded signals vary in duration of 12 to 86 minutes whereas the certain recordings are one hour long. The recorded physiological signals contain ECG signals from three leads, ABP, CVP, PAP, respiratory impedance, and airway $CO2$ waveforms. Some recordings include intra-cranial, left atrial, ventricular and/or intra-aortic-balloon pressure waveforms.

1.11.2. CSL database

Complex systems laboratory (CSL) database is the collection of recording of ABP, ICP (intracranial pressure) and pulse oximetry signals acquired from pediatric intensive care unit. Each recording consists of two signals from two different patients (e.g. abp1, abp2 for ABP record). This database is now the part of biomedical signal processing laboratory [12]. The database has three sets of annotations for ABP, ICP and pulse oximetry signals. Two sets of annotations are created by the medical experts by visual inspection and annotations from beat detector proposed in [44] are also available. The signals in the database are sampled at 125 Hz. The abp1 signal of this database is used for development and testing of ABP feature extraction algorithm described in chapter 3.

1.12. Objectives of the research

On the basis of the above background and literature review, the motivation towards development of technique for feature extraction and modeling of ECG, ABP and CVP is based on the following objectives –

i. Feature extraction of ECG signal by wavelet technique supported by signal energy, frequency spectrum and correlation analysis.

ii. Feature extraction of ABP and CVP signals by wavelet technique supported by signal energy, frequency spectrum and correlation analysis.

iii. Modeling and synthesis of ECG by system identification technique using measured physiological data of ABP and CVP signals.

iv. Modeling and synthesis of ECG by Artificial Neural Network paradigm.

1.13. Thesis outline

Chapter 1 describes the background and introduction to the biomedical signals of cardiac system and their detection techniques and review of literature as the basis of objectives achieved in this research. Chapter 2 presents algorithms for ECG feature extraction using wavelet and energy analysis techniques with justification of selection of the wavelet detail coefficients. In this chapter application of wavelet and energy technique to PVC detection is also proposed. Chapter 3 illustrates the application of a wavelet and energy analysis

based technique for ABP and CVP feature extraction. Chapter 4 presents system identification based approach for modeling and synthesis of ECG using measured ABP and CVP signals for normal and abnormal conditions of the heart. The stability analysis of the developed models is also explained using pole zero plots and step response. In chapter 5, a technique using ANN for modeling and synthesis of ECG using measured ABP and CVP signals has been described.

CHAPTER

ECG Feature Extraction and its Application to PVC Detection

2

2.0. Introduction

In recent years, there has been an increasing demand for early diagnosis of cardiac abnormalities from biomedical signals originating from the heart. The advancement of information technology has accelerated the growth of the signal processing techniques. Modern era of medical science is supported by computer aided feature extraction and disease diagnostics in which various signal processing techniques have been employed in extracting features from the biomedical signals and analysis of these features. The objective of computer aided digital signal processing of ECG signal is to assist the cardiologists by reducing the time in interpreting the results. Most clinical diagnosis of ECG is based on limb leads and chest leads. The errors occurred in the subjective interpretation by visual and manual examination can be overcome by computer aided ECG feature extraction techniques incorporated along with ECG instruments.

QRS complex is the most prominent feature in electrocardiogram because of its specific shape, therefore it is taken as a reference in ECG feature extraction. R-wave detectors are also very useful tools in analyzing ECG features thus form the basis of ECG feature extraction. The other features of ECG are extracted by considering R-wave as the reference point. Therefore, ECG analyzers require the reliable and precise R-wave detection maintaining accuracy as well as giving the results promptly. The detection of QRS complex in ECG requires special and complex algorithms due to the varying morphologies of normal and abnormal complexes and because the ECG signal experiences different types of disturbances with complex origin.

Signal processing techniques can extract appropriate and useful information from biomedical signals such as ECG, ENG, EMG signals etc. that is not usually at hand in the raw signal [23]. For example, frequency analysis of EEG signal can be done to identify δ, θ, α and β waves depending upon the different frequency components associated with each wave. Therefore, signal processing has emerged as an important division in the biomedical signal analysis. The advanced signal processing methods using time-frequency analysis and filtering using wavelet transform has proved to be a very useful tool in determining the underlying properties of physiological signals such as the precise location of the QRS-complex.

Wavelet transform divides a continuous-time signal into different scale components [58]. A time-scale representation of the signal is obtained using digital filtering. In DWT, filters of different cut-off frequencies are used to analyze the signal at different scales. The signal is passed through a series of highpass and lowpass filters to analyze the high and low frequency components of the signal respectively. In using traditional analog and digital filters, the ECG signal may still contain noises like residual power-line interference. Although power-line interference is removed by traditional analog and digital filters, the desired components of ECG such as Q and R waves with frequency matching with power-line may be suppressed [140]. Wavelet transform has been effectively used for removal of power-line interference [141, 142] and base line drift [143] in ECG signals without suppressing the desired components. In case of wavelet transform, the resulting detail coefficients are band pass filtered. By selecting a detail coefficient at a certain level depending upon frequency range of ECG signal, an ECG signal free from both high as well as low frequency noise is obtained. Wavelet transform is popular because it satisfies energy conservation law and original signal can be reconstructed [21].

Among the wavelet based techniques [20-25], the detection of R wave in ECG in [20], involves the decomposition of ECG signal using quadratic spline wavelet and selection of detail coefficients d1-d4 using 3 db band width and application of detection rules (thresholds) to the modulus maxima of d1-d4. In [21], decomposition of ECG signal by first derivative of Gaussian smoothing function and application of thresholds to the detail coefficients d1-d4 is suggested for R peak detection. Newly constructed wavelets are suggested for ECG signal decomposition and detection of R peak is done by applying

threshold to d4 signal in [22]. In [24], authors suggest decomposition of ECG signal using db4, db6 wavelet and selection of d3-d5 signals using detection rules. In [23], mexican hat wavelet is used for decomposition and detail coefficients d5-d6 after thresholding are used for detection of R peaks. In [25], Haar, db4 and db6 wavelets are used for decomposition and d3-d5 signals are analyzed using decision rules for detection of R peak. The remaining features P.Q,S,T waves are detected considering R peaks positions as reference in [20-21, 23-25]. The authors in [24,25] proposed that the performance of db6 is better than other wavelets. Also, most of the QRS detection algorithms mentioned above [28,24,25] are developed on ECG lead-II signals on MIT-BIH database as the QRS complexes are outstanding in lead II. Some of the authors included ECG signals 102, 104 from leads V5 and V2 respectively but accuracy reported for these records was 94 % and 92 % respectively [30]. In all these wavelet methods [20-25], authors have selected detail coefficients out of d1-d6 from the wavelet decomposition stages but the selection of relevant detail coefficient is not adequately justified by the authors. Sometimes, situation may arise such that ECG signals from lead-II may not be available due to surgical dressing of patients or some other reason as in case of record 102 and 104 as stated in section 1.11.1.2. This may happen at any time in medical observations; therefore the detector must be capable of detecting the R-peaks from all type of leads.

In this chapter, we propose an algorithm for ECG feature extraction using wavelet (Daubechies) technique supported by signal energy, frequency spectrum and cross-correlation analysis. Another algorithm for ECG peak detection using energy analysis technique is also proposed. Both these algorithms are equally applicable for detection of R peaks from ECG signals of lead II as well as other leads such as V2, V5. The merit of both of these algorithms motivated us to apply these algorithms for the detection of life threatening heart disease - premature ventricular contraction (PVC) beats in ECG.

We present ECG peak detection by two methods - wavelet based method and energy analysis of ECG signal. The wavelet based approach described in this objective is robust and simple to implement with no requirement of preprocessing. The selection of detail wavelet component has been justified by energy, frequency and cross-correlation analysis. Since, there are wide variations in amplitudes of wavelet decomposed signals; a fixed threshold does not work for R peak detection. Therefore, we have adopted a 'window based

threshold' where the threshold value is adjusted depending upon the signal amplitude over a certain duration. The selected detail signal is first thresholded then the maximum amplitude levels of all the peaks are detected. The signal is then filtered by applying a refractory period to select the R peaks. The R peaks detected by wavelet method are used for the detection of remaining features of ECG signal such as P, Q, S and T waves.

In the energy analysis technique for R wave detection, energy calculation of ECG signal under test has been performed by dividing the signal records into a number of windows. The techniques used include window shifted by window size and window shifted by one sample. Energy analysis of detail coefficients show that d4 signal containing highest energy content comprises of maximum information of QRS complexes. This concept motivates us to detect ECG peaks if the window based energy analysis of ECG signal is performed and the resultant energy signal is further analyzed using thresholding and refractory period concepts for detection of ECG peaks. Therefore, window based energy analysis of ECG signal may result in the higher energy amplitudes wherever ECG peaks exist.

The detected R-peaks are applied to detect PVC beats in ECG. PVC is the premature beat that occurs in ventricular region of heart. The method for detecting the abnormal PVC complexes is based on the calculation of RR interval of detected R peaks and energy analysis of ECG signal. We have proposed a combined method for PVC detection where, RR interval calculation by wavelet and energy is supported by intersection of energy analysis technique on the ECG signal. The algorithm proposed for PVC detection includes detection of R peaks using window based energy analysis of ECG signal using a window of 100 ms duration that incorporate window shift by one sample and further energy analysis of ECG signal using a window of 600 ms duration where window is shifted by window size.

2.1. ECG feature extraction

Feature extraction of ECG signal involves detection of constituent components of ECG waveform such as P, Q, R, S, T waves. Detection of ECG peak is carried out using wavelet technique and energy analysis. Further, remaining features (P, Q, S and T) of ECG are detected using wavelet technique considering locations of detected R peaks as reference as described in the following sections –

2.1.1. ECG peak detection using wavelet technique

ECG signals (.dat files) downloaded from MIT-BIH database are first converted in to Matlab readable format (.mat files). The signals from both leads now become readable separately. Then the signals from upper lead as stated in section 1.11.1.2 in the record are taken for analysis. Detection process is performed on 44 out of total 48 downloadable records from the database. The block diagram of ECG feature extraction using db6 wavelet is shown in Figure 2.1.
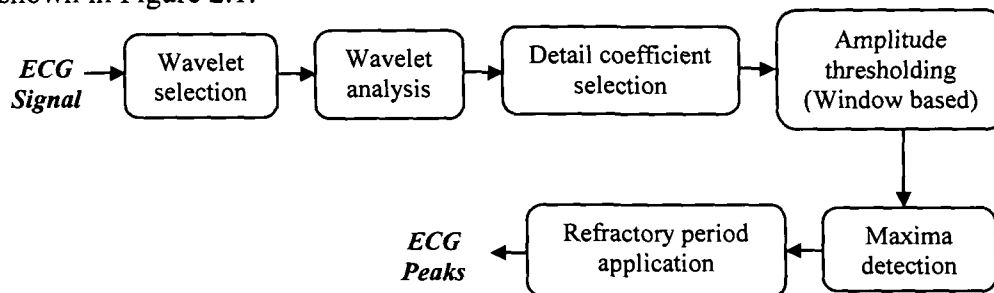
Figure 2.1 Block diagram of ECG peak detection using Daubechies (db6) wavelet

2.1.1.1. Wavelet selection

The selection of relevant wavelet is an important task before starting the detection process. But there is no universal method suggested to select a particular wavelet. The choice of wavelet selection is accorded on the type of signal to be analyzed. The wavelet having similar look to the signal being analyzed is usually chosen [22]. There are several wavelet families like Harr, Daubechies, Biorthogonal, Coiflets, Symlets, Morlet, Mexican Hat, Meyer etc. and several other Real and Complex wavelets. However, Daubechies (db6) family of wavelets has been found to give details more accurately than others [28,24,25]. Moreover, this wavelet show similarity with QRS complexes and energy spectrum is
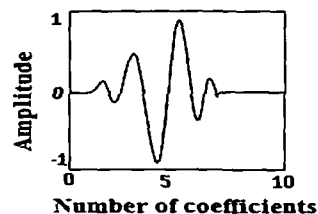
Figure 2.2 Wavelet function (ψ) of Daubechies (db6) wavelet

concentrated around low frequencies [22]. Therefore, we have chosen the above mentioned wavelet for extracting ECG features in our application. Wavelet function of Daubechies (db6) wavelet is shown in Figure 2.2.

2.1.1.2. ECG signal decomposition

The first step in ECG feature extraction is the detection of R wave which being the highest amplitude wave in ECG signal forms the basis of ECG feature extraction. R wave detection in ECG involves the following steps –

The signal under test is decomposed up to a required level depending upon dominant frequency components in the signal. The choice of required level of decomposition depends on the frequency components of interest available in the wavelet coefficient at a particular level. The maximum number of decomposition levels depends upon the total number of samples present in the signal. The relationship can be expressed as -

$$2^n = N \qquad\qquad (2.1)$$

Where, n = total number of levels of decomposition,

           N = total number of samples in the signal to be expressed as power of 2 for full decomposition of the signal.

As an example, ECG signal description of record 100.dat is shown in Table 2.1. The wavelet decomposition structure of 100.dat and 123.dat using db6 wavelet is shown in Figure 2.3 (a-b). The waveform shows signals decomposed up to 8 levels only. Although the original signal is of 30 minutes duration, for better illustration, details are scaled and signal is shown for 5 seconds duration only. The approximation coefficients at level 8 of ECG signals from record 100 and record 123 are shown in Figure 2.3(c). It is clear from Figure 2.3(a-b) and Figure 2.3(c) that none of the features of ECG signals such as P, Q, R, S and T waves resemble in approximation coefficients at level 8 of records 100 and 123.

2.1.1.3. Selection of detail coefficient by energy analysis

It has already been discussed earlier that selection of detail coefficient has not been adequately justified in the literature. It is assumed that most of the researchers select the detail coefficient when there is a similarity in the waveform of the detail coefficient with the original ECG. In this chapter, we propose a new method of selecting the detail coefficient based on energy content, the frequency analysis and cross-correlation coefficient. This technique proves to be a robust justification for selecting the detail

coefficient instead of visual similarity [22] and 3 db bandwidth [20]. The technique is described below –

Table 2.1 Details of record 100 (MLII, V5) and record 123 (MLII, V5)

Record No.	Sex	Age (Years)	Beats	Before 5:00 min	After 5:00 min	Total
100	Male	69	Normal	367	1872	2239
			APC	4	29	33
			PVC	-	1	1
			Total	371	1902	2273
123	Female	63	Normal	249	1266	1515
			PVC	-	3	3
			Total	249	1269	1518
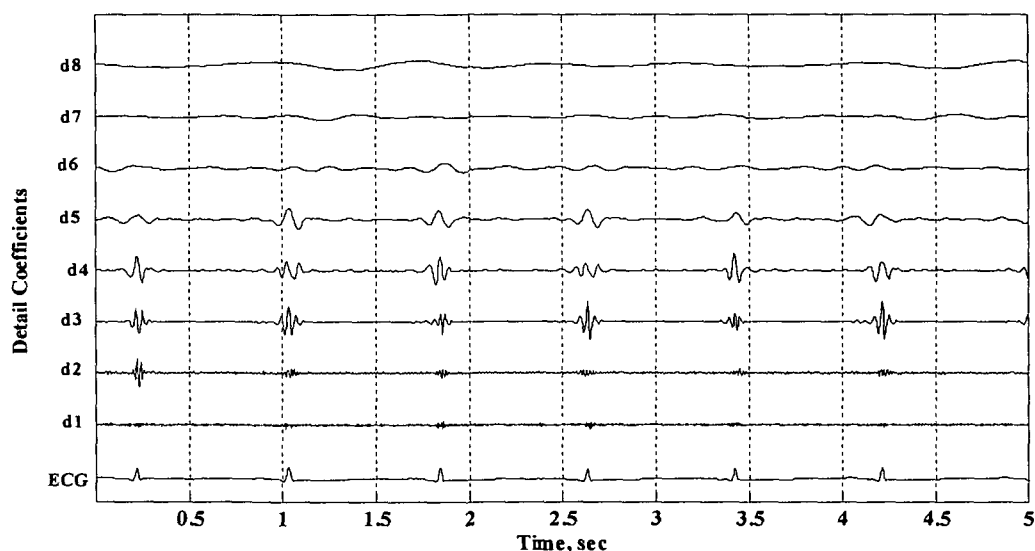
Figure 2.3(a) Decomposition of ECG signal (record 100, lead-II) using db6 wavelet

Most of the energy of a normal ECG signal is concentrated within QRS complex interval of about 80 ms and having a frequency range of 3-40 Hz [25,144], Normally isoelectric segments – PQ, ST and TP contain insignificant energy and the signal amplitudes are almost zero over these corresponding intervals. Definition and determination of signal energy is explained in detail in section 1.9.2.

The signals have been decomposed up to a certain approximate level using the selected wavelet and the energy content of each level is determined for the 8 decomposed levels of

ECG lead II signals from record 100 and record 123 and shown in Table 2.2. The plots for the energy distributions of the signals are shown in Figure 2.4(a-b).

The energy plots Figure 2.4(a-b) show that the energy is highest for detail signal at level 4 among the detail coefficients d1 to d8. Therefore, we consider that d4 carries the dominant details of QRS complexes. It is clear from Table 2.2 that the sum of energy of all detail coefficients and the remaining one approximation is equal to the energy of a single ECG record under test. Therefore, this energy level diagram of decomposition structure proves the energy conservation principle of wavelet transform. It means that original signal can be faithfully reproduced from the decomposed components and the information in the original signal is preserved maximally in the d4 component. It is also observed from Table 2.2 that energy is observed highest for a8 but as seen from Figure 2.3(c) that approximation coefficient at level 8 has no relevance with ECG features so it is not selected.
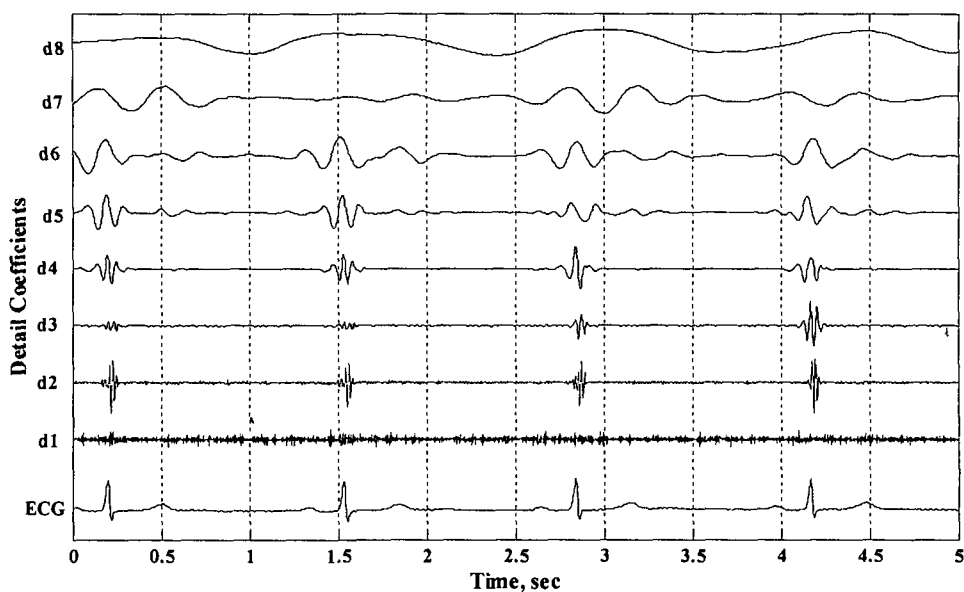
Figure 2.3 (b) Decomposition of ECG signal (record 123, lead-II) using db6 wavelet
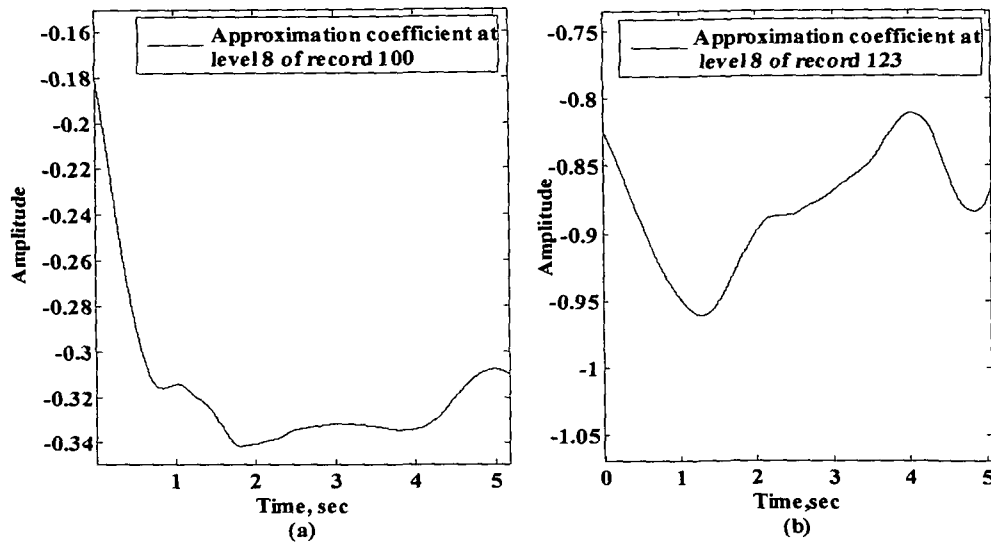
Figure 2.3 (c) Approximation coefficients of ECG signals at level 8 from
(a) record 100 and (b) record 123

Table 2.2 Energy content of detail and approximation coefficients

Signal	Energy content (%)	
	Record 100	Record 123
d1	0.0171	0.0031
d2	0.3574	0.0385
d3	4.1817	1.2069
d4	*9.7255*	*3.8376*
d5	6.5342	2.7949
d6	2.8434	1.8000
d7	1.4699	0.5596
d8	1.1600	0.3458
a8	73.7108	89.4136
Total energy content (%)	100	100

2.1.1.4. Selection of detail coefficient by frequency analysis

Time-amplitude representation of the signal is not always the best representation of the signal for most signal processing related applications. In many cases, the most distinguished information is hidden in the frequency content of the signal. In order to ensure that the frequency components available in the original ECG signal are preserved by the detail component, the only way is to analyze the signals in frequency domain. We propose here the second method of selecting the detail coefficient by frequency analysis as
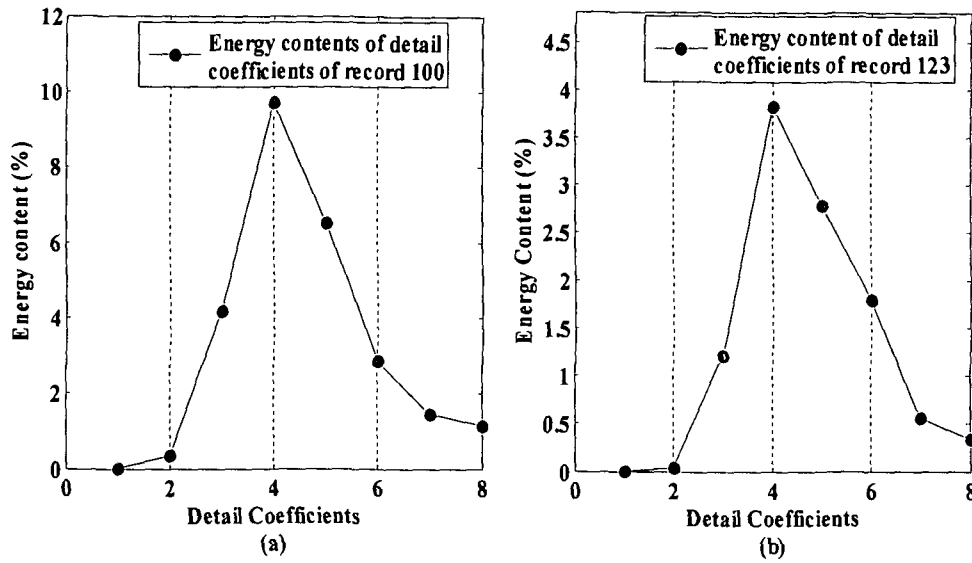
Figure 2.4 Energy plot of detail coefficients (d1-d8) of records (a) 100 and (b) 123

discussed below –Fourier transform of a signal shows how much of each frequency exists in the time domain signal. A fast Fourier transform (FFT) is an efficient algorithm to compute the discrete Fourier transform (DFT) and its inverse. FFTs are of great importance to a wide variety of applications in digital signal processing. Therefore, FFT of the detail coefficients and ECG signal is performed for records 100 and 123 and FFT plots for ECG signals from records 100 and 123 along with detail coefficients (d1-d8) are shown in Figure 2.5 (a-b) respectively. More clear FFT plots of ECG signals for records 100 and 123 along with their detail coefficients d4 are shown in Figure 2.6(a-b) and Figure 2.7(a-b) respectively.

Since in an ECG signal, except QRS complex, other complexes are mostly dc in nature and therefore, the frequency of the ECG is dominated by QRS frequency. A normal clinical cardiac signal has a maximum frequency of 40 Hz [145, 24]. The frequency ranges of all the detail coefficients are shown in Table 2.3. The frequency range of d4 signal for record 100 and 123 is found to be 1.994 Hz – 39.35 Hz and 1.056 Hz – 40.23 Hz respectively which is almost same as that of a QRS complex. Therefore we have chosen d4 signal of ECG of record 100 and record 123 for further analysis. It is also clear from Figure 2.6(a) and Table 2.3 that the original ECG signal (Figure 2.6(a)) is affected with power-line interference whereas d4 signal (Figure 2.5(a), Figure 2.6(b) and Table 2.3) is free from power-line interference.
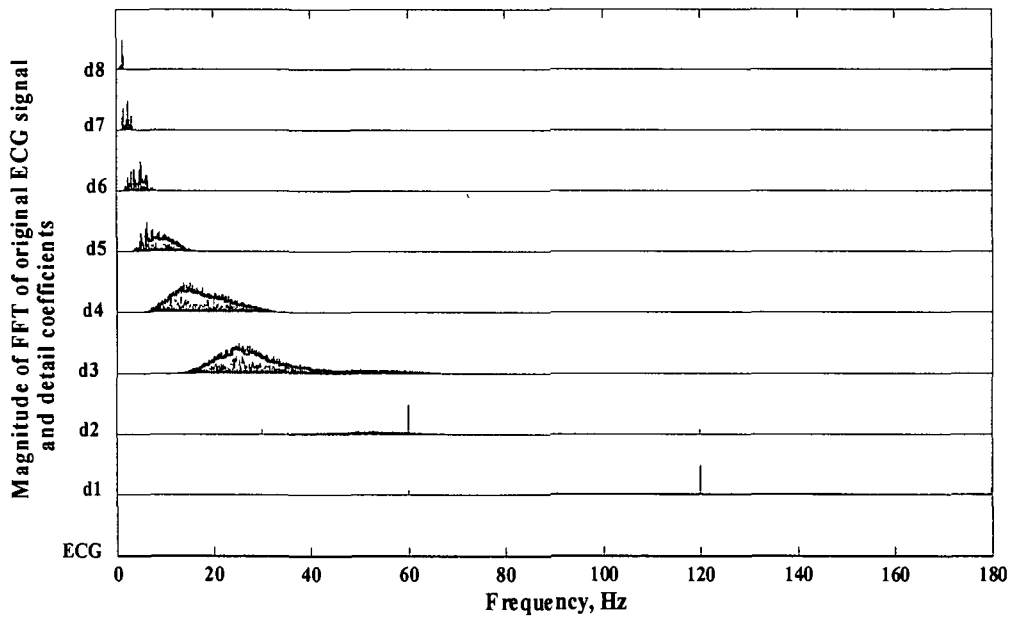
Figure 2.5(a) Frequency distribution of details and ECG signal (record-100, lead-II)
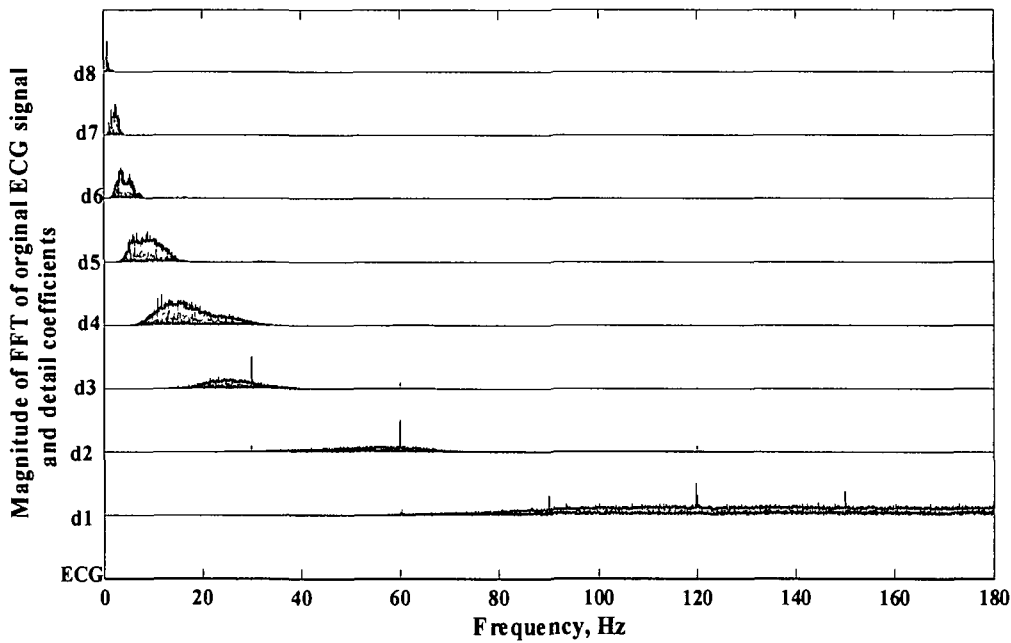
Figure 2.5(b) Frequency distribution of details and ECG signal (record-123, lead-II)
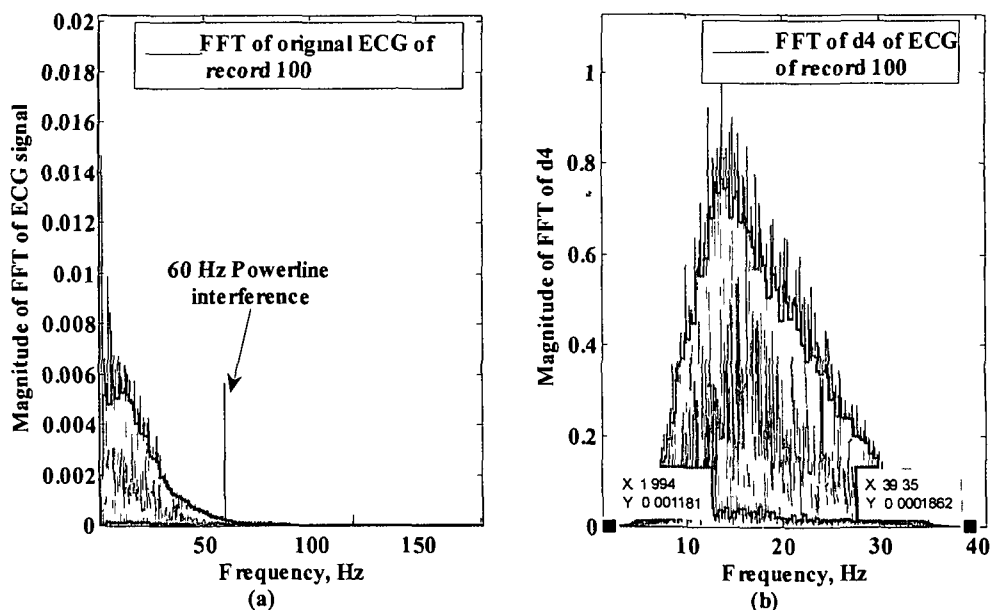
Figure 2.6 Frequency distribution of (a) ECG signal and (b) detail coefficient (d4) of
record-100, lead-II)
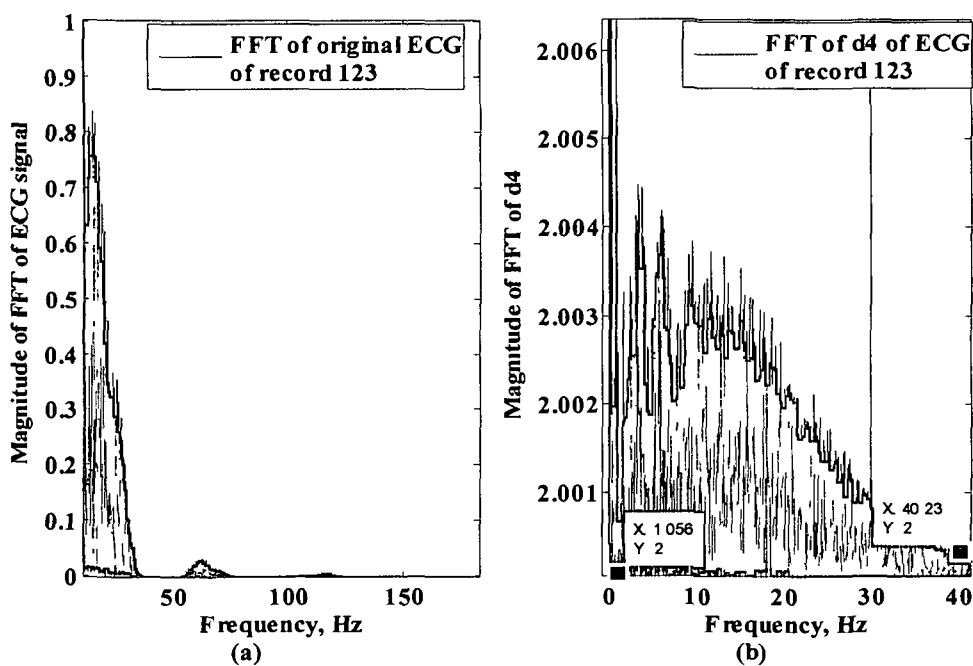
Figure 2.7 Frequency distribution of (a) ECG signal and (b) detail coefficient (d4) of
record-123, lead-II)

Table 2.3 Frequency range of detail coefficients

Detail signal	Frequency content (Hz)	
	Record 100	Record 123
d1	26.24-180	23.73-180
d2	12.56-131.4	14.06-133.3
d3	3.967-66.61	3.32-67.78
d4	1.994-39.35	1.056-40.23
d5	1.508-17.56	1.588-18.08
d6	0.8278-8.765	0.8712-8.737
d7	0.5931-4.253	0.5228-4.322
d8	0.3111-2.123	0.2221-2.248

2.1.1.5. Cross-correlation analysis

In addition to the above two analysis, cross-correlation analysis between all the decomposed signals individually with the original ECG signal was performed. This provides us the time domain relationship between the original and the decomposed signals. The cross-correlation between two different signals may be defined as the measure of match or similarity between one signal and time-delayed version of another signal. The cross-correlation coefficient takes on values ranging between +1 and -1. The mathematical equations of cross-correlation coefficients are discussed in section 1.9.4. The values of cross- correlation coefficients are shown in Table 2.4. From Table 2.4, it is clear that value of cross-correlation coefficient is highest for d4 for both records 100 and 123. Therefore, it is evident that d4 is highly correlated with the original ECG signal in time domain.

Table 2.4 Cross-correlation coefficients

Detail coefficients	Cross-correlation coefficients		Inference
	Record 100	Record 123	
d1	0.0246	0.0162	Weak
d2	0.1122	0.0574	Weak
d3	0.3838	0.3214	Moderate (low)
d4	*0.5851*	*0.5732*	*Moderate (high)*
d5	0.4785	0.4891	Moderate (low)
d6	0.3165	0.3925	Moderate (low)
d7	0.2275	0.2187	Weak
d8	0.2020	0.1719	Weak

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

2.1.1.6. Window based thresholding

The selected detail coefficient d4 is used to perform of R-wave detection in ECG. In thresholding, small coefficients are suppressed by introduction of a threshold. Such a procedure is called wavelet thresholding. There exist various thresholding procedures such as soft thresholding and hard thresholding which are discussed in details in section 1.9.5. In [65], author used hard thresholding method to remove high amplitude coefficients generated by ectopic beats from RR interval signals. In this case, we apply hard thresholding in which the samples below a predetermined threshold are set to zero. The threshold is selected as a percentage of maximum value of signal amplitude over certain duration as discussed below –

Let the length of signal be L that has to undergo window based thresholding and w is the window size. If the length of the signal is not divisible by the window size, zero padding can be done to make the signal length divisible by window size. The number of samples in the signal is divided by the window size and remainder value is obtained. If the signal length is defined as –

$$y = a : b \qquad (2.2)$$

Where, $'a'$ and $'b'$ the start sample and end sample number of the signal respectively. Let R denotes the number of samples obtained after dividing length of signal by window size, the new signal after adding zeros on the right side can be represented as

$$y_N = a:(L + R) \qquad (2.3)$$

The total number of windows in the signal be N then

$$N = \frac{L + R}{w} \qquad (2.4)$$

Where, w denotes the window size. The new signal is represented as the segments of window size. Let the total number of windows in the sample be $'N'$, then we can write

$$y_N = (a:w), \; [(w+1):2w], [(2w+1):3w] \cdots\cdots\cdots [((N-1)w+1):Nw] \qquad (2.5)$$

Let $th_1, th_2, th_3 \cdots\cdots th_N$ be the threshold for each window, then

$$th_N = A \times [\max(y_N ((N-1)w+1):Nw)] \qquad (2.6)$$

Where, A is a constant that that is fixed for overall process. A is defined as the percentage of maximum amplitude of d4 signal over the duration of the predefined window. In this case, we have used $A = 10$ and $w = 720$ which is equivalent to 2 second duration of signal. The signal d4 before and after thresholding is shown in Figure 2.8. It is clear from Figure 2.8 that the threshold varies for each window depending upon the amplitudes of samples falling in the window. The samples with amplitude below predefined threshold are set to zero and we get signal in the positive side of X-axis only.
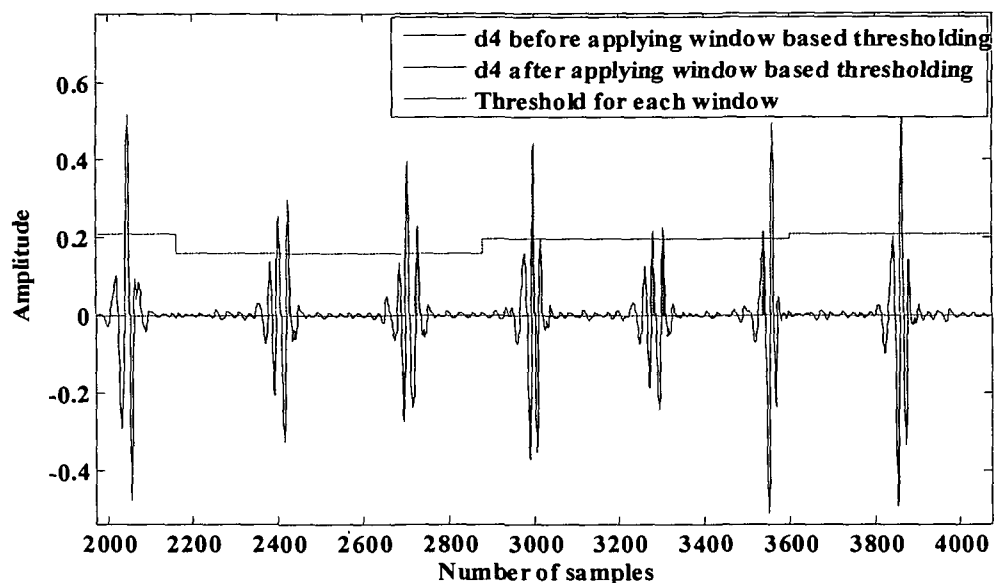
Figure 2.8 Signal d4 before and after window based thresholding

This method of thresholding leads to adaptive thresholding. In this method of thresholding, the signal is segmented in to equal segments by defining a window of particular duration in terms of samples and a distinct value of threshold is selected from each segment of the signal. This type of thresholding strategy limits any large variation in the signal amplitude at a certain instant to a particular segment of the signal due to which the true peaks with lower amplitudes may be ignored in other segments if threshold is defined for the entire signal.

2.1.1.7. Maxima detection and R-wave positions

The numbers of remaining coefficients of d4 after thresholding as shown in Figure 2.8 undergo maxima detection. The signal d4 after thresholding and maxima detection is shown

in Figure 2.9. A maxima is detected if the signal amplitude at any instant is higher than the amplitudes of previous and succeeding samples. Mathematically, this condition may be represented as –

$$if \quad d4(k) > d4(k+1) \quad and \quad d4(k) < d4(k-1)$$

$$d4(k) = \text{maxima}$$
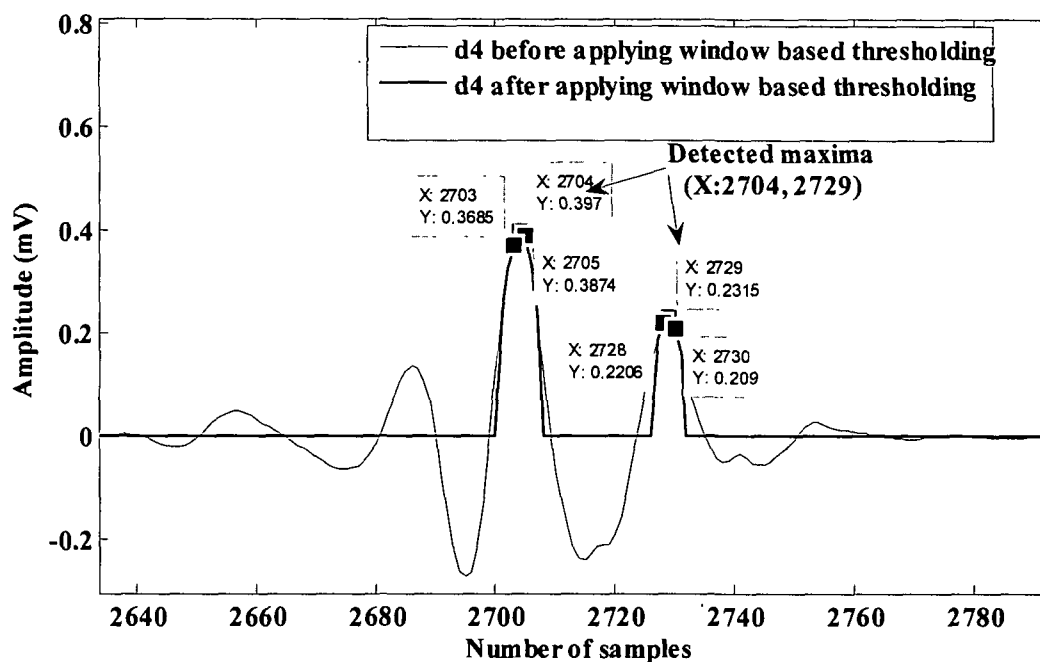
Figure 2.9 Maxima detection in signal d4 after window based thresholding

The remaining number of d4 coefficients after maxima detection are taken as the number of R waves and their positions are taken as the R peaks locations. Since R wave has the highest peak in QRS complex and hence no two QRS complexes can be found during less than 200 ms, a refractory period of 200 ms is applied after detection of first peak which gives rise to actual number of R-waves [21]. Detected peaks positions in ECG lead II signal from record 105 after application of refractory period of 200 ms is shown in Figure 2.10. The positions of the R peaks has been detected and marked on the original signal. The waveforms with the positioned R peaks for ECG signals of records 100, 101, 102 and 123 are shown in Figure 2.11 (a),(b) & (c) respectively.
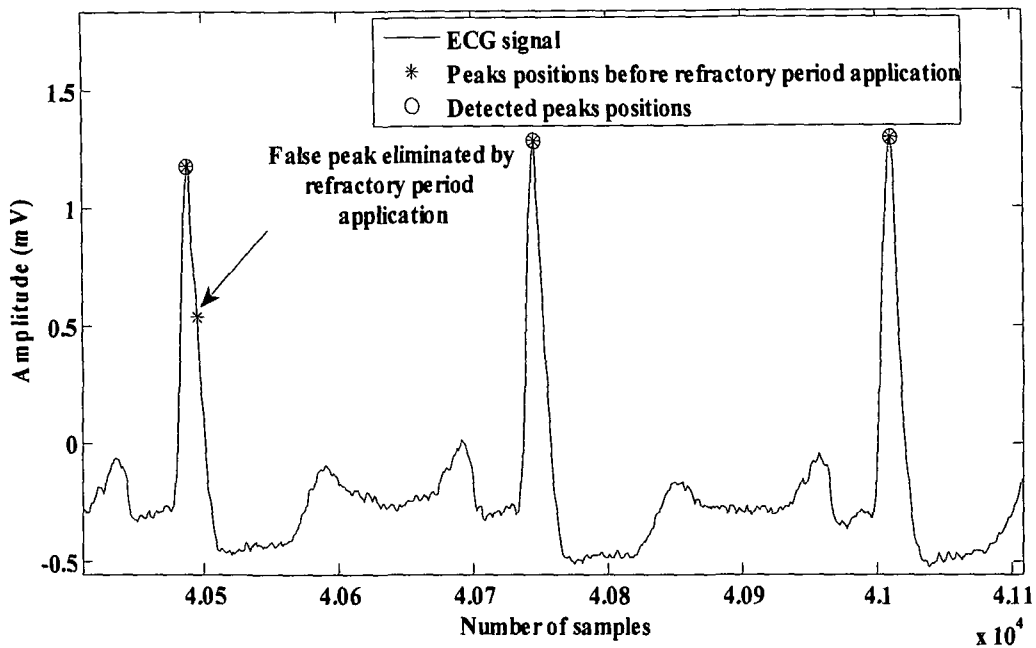
Figure 2.10 Signal d4 of ECG lead II signal from record 105 after thresholding
and application of 200 ms refractory period
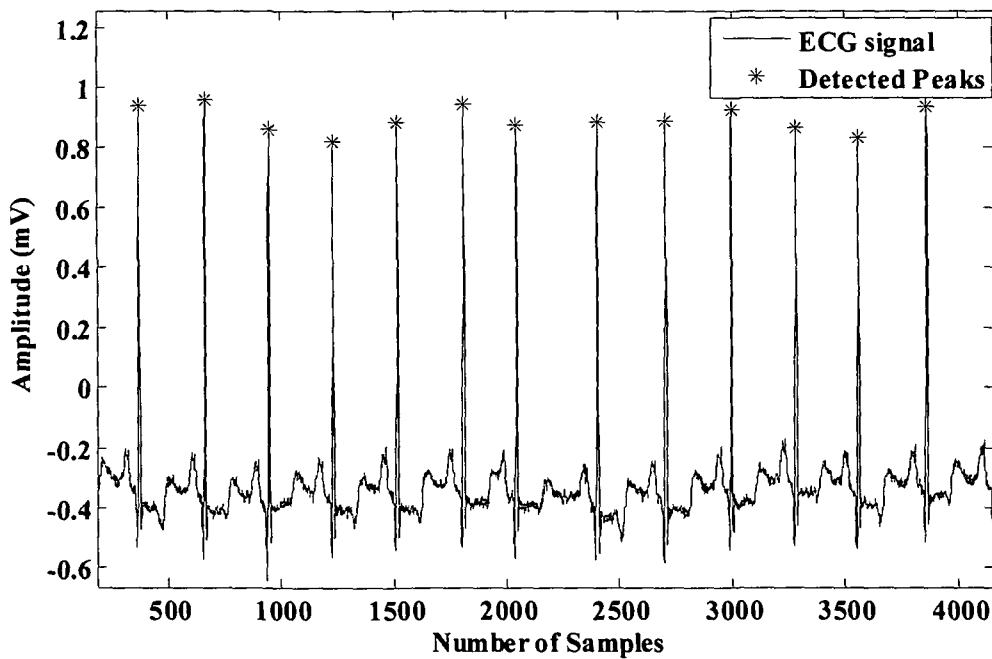
Figure 2.11(a) Positioned R-peaks in ECG (record-100, lead-II) signal
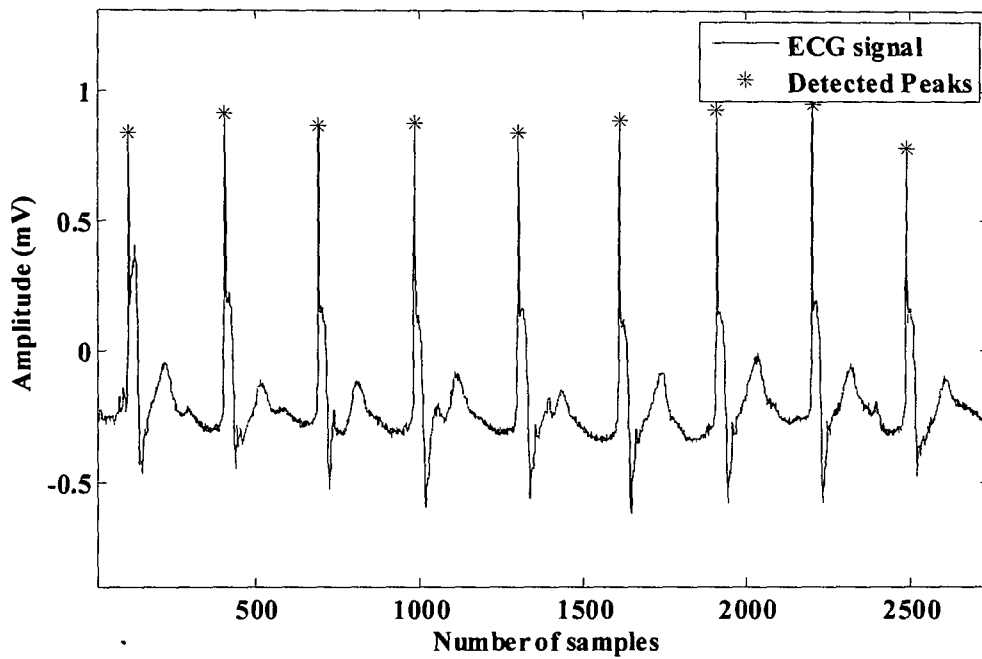
Figure 2.11(b) Positioned R-peaks in ECG (record-102, lead-V5) signal
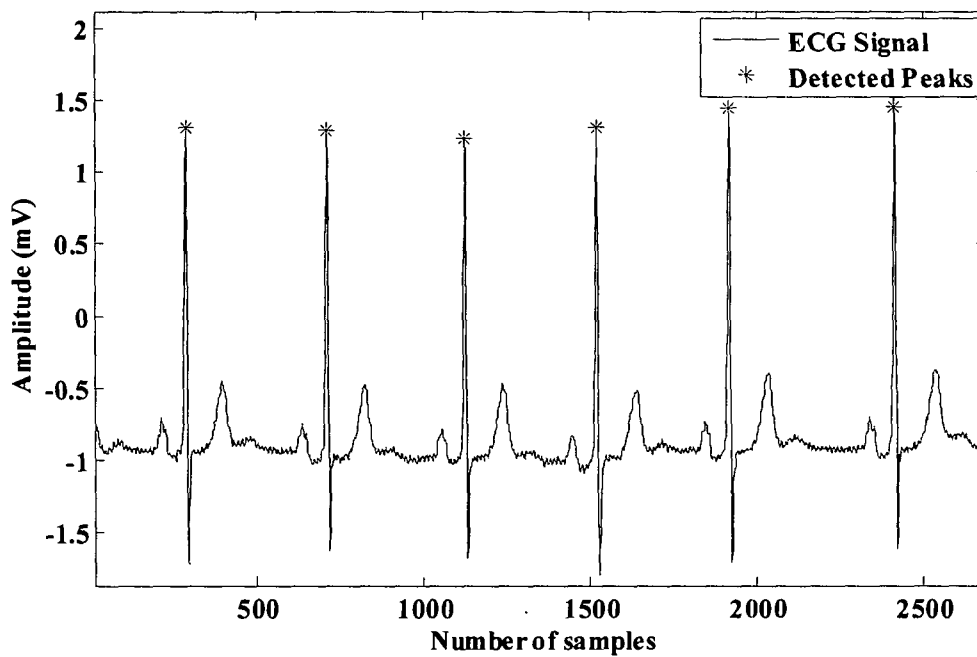
Figure 2.11(c) Positioned R-peaks in ECG (record-123, lead-II) signal

2.1.2. ECG peak detection using energy analysis technique

As discussed in section 2.1.1.3 that the detail signal d4 generated from ECG signal after wavelet analysis comprise of maximum information of QRS complexes and consequently used for the detection of R peaks. It is further established in section 2.1.1.7 that detail coefficient d4 after supplementary analysis resulted in the detection of R peaks in ECG signal. This concept motivates us to detect ECG peaks if the window based energy analysis is performed and the resultant energy signal may be further analyzed for detection of ECG peaks. R wave is the highest amplitude wave in ECG; therefore it is presumed that window based energy analysis of ECG signal will result in the higher energy values wherever ECG peaks exist. Converting the ECG signal into energy signal facilitates suppression of lower peaks (P and T waves) and resulting a comparatively distinct and smooth signal facilitating the ease of peak detection. In this section, we represent the detection of ECG peaks by energy analysis.

2.1.2.1. Method of detection

The ECG signals from upper lead of all the forty four downloadable records (5 minute duration) are taken for detection of peaks by energy analysis. The algorithm of ECG peak detection using energy analysis technique is shown in Figure 2.12.

The detection process is performed by two methods –

    a) Window shifted by window size

    b) Window shifted by one sample

2.1.2.2. Data preprocessing

Most of the ECG signals used in this work are affected at source itself by low as well as high frequency artifacts such as baseline drift, power-line interference etc, therefore it is necessary to remove these artifacts before starting the detection process. Base-line drift may be caused in chest lead ECG signals by coughing or breathing with large movement of chest, or when an arm or leg is moved in case of limb lead ECG acquisition. Poor contact of electrodes may also cause low frequency artifacts. Base-line drift may sometimes be caused by variations in temperature and bias in the instrumentation amplifiers as well.
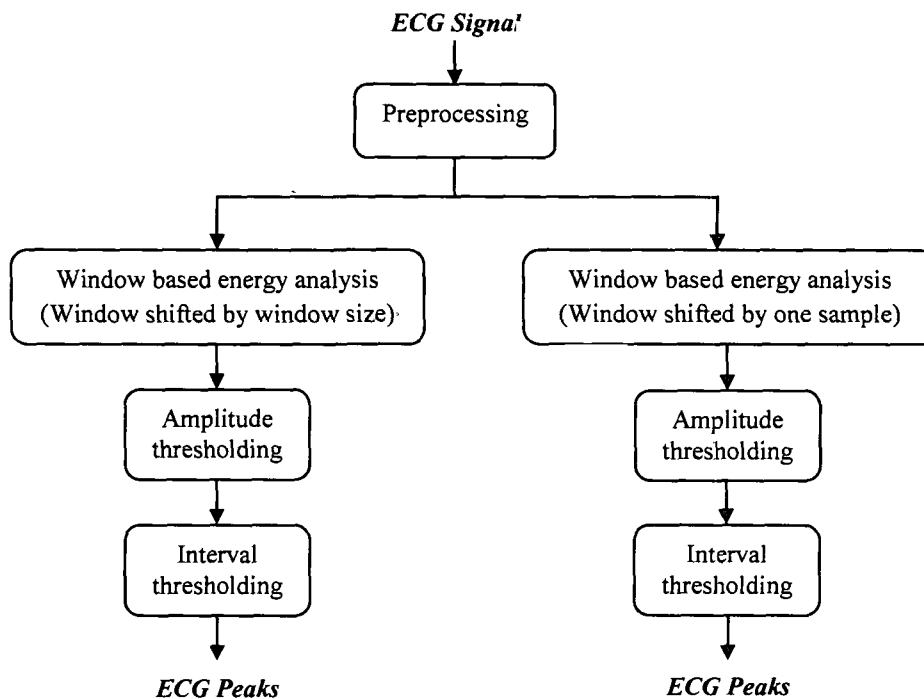
Figure 2.12 Block diagram of ECG peak detection algorithm by energy analysis

A large base-line drift may cause the positive or negative peaks in ECG to be clipped by the amplifiers or the ADC [146]. Since energy calculation of a signal is directly related to absolute amplitude of the signal being analyzed and base line drift causes a large variation in signal amplitude energy as stated in equation (1.13-1.14), hence removal of base line drift is the key step before starting the detection process. The baseline drift frequency is 0.5 Hz as mentioned in [148] and FFT of ECG signal of record 109 in Figure 2.13 shows a peak at 0.0006466 Hz.

As stated in 2.1.1.3 that most of the energy of a normal ECG signal is concentrated within the interval of about 80 ms spanned by QRS complex and having a frequency range of 3-40 Hz [25,144], the information below 2 Hz is redundant for R-peak detection. Therefore, the ECG signal under test is filtered by fourth order Butterworth high-pass filter with cut-off frequency 2 Hz to remove baseline drift. The ECG signal from record 109 affected by base line drift and filtered ECG signal is shown in Figure 2.14. From Figure 2.14, it is clear that the filtered output retains the characteristics of QRS complex but distort the P and T waves

to some extent and base-line drift is removed which ensures the correct calculation of energy.

However, if the signal under study is affected with high frequency noise such as power-line interference, energy distribution of the signal does not result in high peaks at lower amplitudes. Energy distribution of the signal (record 100) affected with power-line interference is shown in Figure 2.15(a-b) and FFT of the signal is shown in Figure 2.6(a).
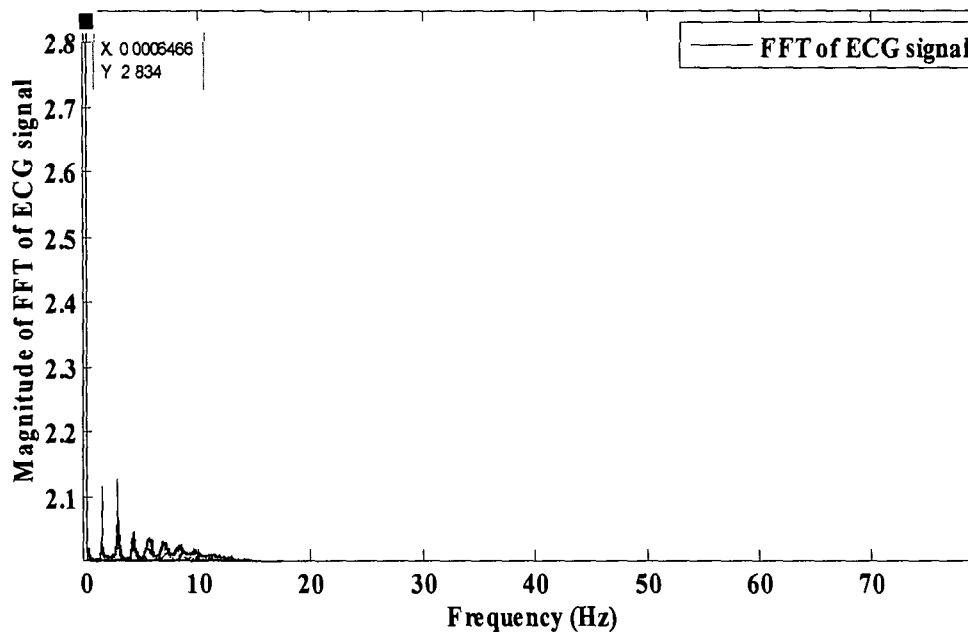
Figure 2.13 FFT of ECG signal of record 109

2.1.2.3. Signal energy analysis by moving window

The filtered ECG signals are then processed for energy calculation for a window equivalent to the duration of the QRS complex (approximately 100 ms for 36 samples). Further, window is shifted by the following two methods –

a) Window shifted by window size

If the length of the signal is not divisible by the window size, zero padding can be done to make the signal length divisible by window size. But in this case, zero padding is not required as length of five minute segment of each record of the database is divisible by window size. Further, the window is shifted by the window size and energy of samples falling in the succeeding window is calculated. This process is repeated for the whole signal

length. Thus first calculated energy value corresponds to first 36 samples (100 ms) and the next calculated energy value corresponds to next 36 samples i.e. 37-72 samples of ECG signal and so on. The calculated energy values are then expanded to the whole length of the
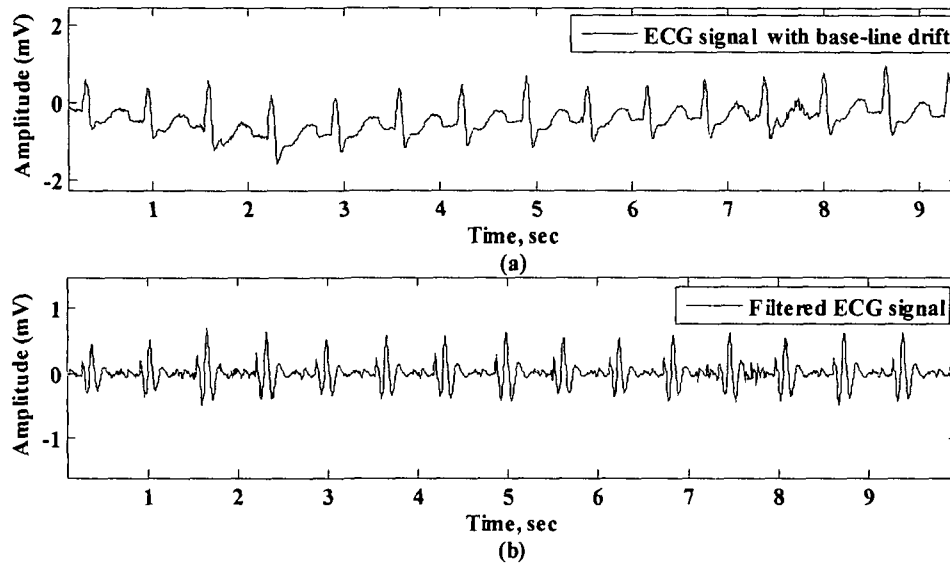
Figure 2.14 ECG signal from record 109 with (a) base line drift (Top) and (b) filtered ECG (bottom) after filtering with fourth order Butterworth highpass filter

signal by repeating each calculated window energy value by the number equivalent to window size. Mathematically, this method of window shift can be represented as mentioned below –

The total number of windows (m) in the entire signal length is given by –

$$m = \frac{n + n'}{w} \qquad (2.7)$$

Where,

n = Number of samples in ECG signal

n' = Number of zeros required to be added to make the signal length divisible by window size

w = Number of samples in one window (i.e. window size)

if $m = \frac{n}{w}$ is an integer

then $n' = 0$

The number of samples (n), can be distributed in each window as per window size to the entire signal length as follows-

$$n = N_1, N_2, N_3 \ldots\ldots\ldots\ldots\ldots N_m \tag{2.8}$$

Where,

$$N_1 = a : w$$

$$N_2 = (a + w) : 2w$$

$$N_3 = (a + 2w) : 3w \ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$N_m = [\{(a + (m - 1)w)\} : mw]$$

Where 'a' refers to start of window and is usually taken as 1 since we start from the beginning of the signal and $N_1, N_2, N_3, \ldots\ldots\ldots\ldots N_m$ denote the start sample number and end sample number of corresponding window. Total calculated window energy values E for the entire signal length can be written as –

$$E = E_1, E_2, E_3 \ldots\ldots\ldots\ldots\ldots E_m \tag{2.9}$$

Therefore, calculated energy value of first window value (E_1) corresponds to energy of N_1 samples in the signal falling in the first window; calculated energy value of second window (E_2) corresponds to energy of N_2 samples in the signal falling in the second window and so on.

This calculated energy of individual windows are expanded to the entire length of ECG signal. Thus, a new energy signal is formed whose length is equal to the length of the ECG signal under test. This energy distribution for record 100 is shown in Figure 2.15(a). It is clear from Figure 2.15(a) that energy of the certain window is higher where QRS complexes are prominent than the other windows.

b) Window shifted by one sample

In this case of energy calculation, window is shifted by one sample and energy of samples falling in the next window is calculated. Thus, energy of first window is the energy of first 1-36 samples and energy of succeeding window is the energy of 2-37 samples and so on. But in this method of energy calculation, zero padding of number of samples equivalent to window size *(w)* is essential factor to calculate the energy of samples falling in the last window. For this purpose, half of the number of zeros required to be added, are pad to both

sides of the signal. This process of energy calculation with ECG record 100 under test with its window energy distribution is shown in Figure 2.15(b). Here also it is clear that higher energy values correspond to QRS complex.

Thus a new signal is formed whose length (L) is given by-

$$L = \frac{w}{2} + n + \frac{w}{2} \qquad (2.10)$$

Where, 'n' is the number of samples in ECG signal under test and 'w' is the window size.

In this case, equation (2.8) can be rewritten as-

$$L = N_1, N_2, N_3 \ldots\ldots\ldots\ldots\ldots N_n \qquad (2.11)$$

Where,

$$N_1 = a : w$$

$$N_2 = (a+1) : (w+1)$$

$$N_3 = (a+2) : (w+2)\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$N_n = [\{(a+(n-1))\} : \{(w+(n-1))\}]$$
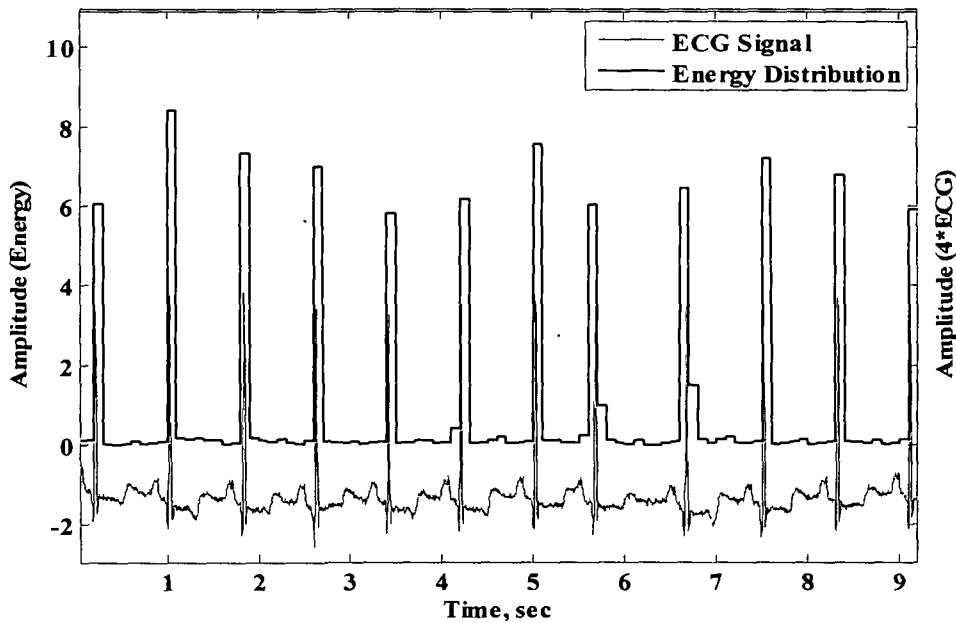
Figure 2.15(a) ECG signal from record 100 and its energy distribution using window shifted by window size
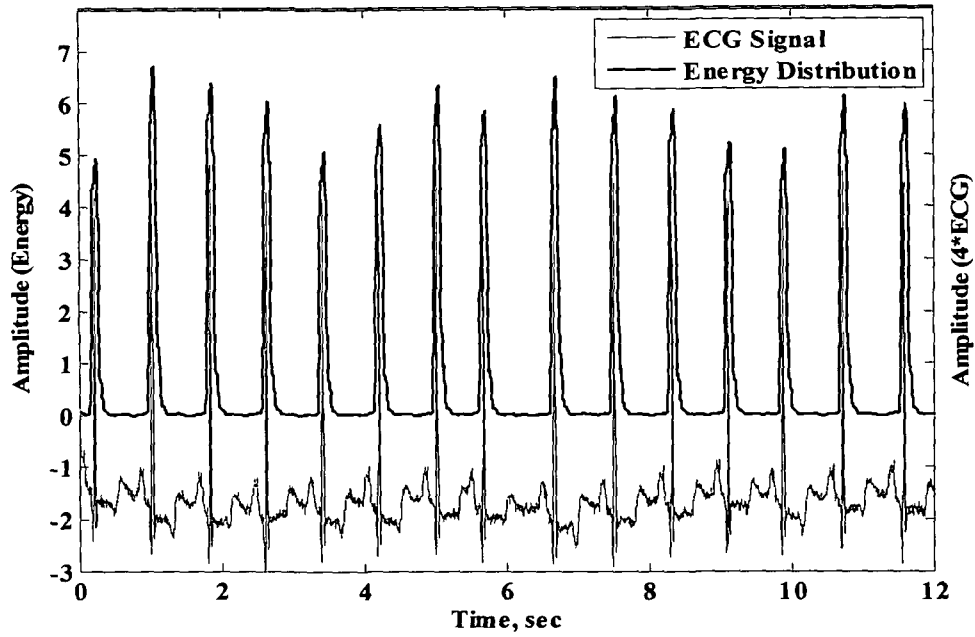
Figure 2.15(b) ECG signal from record 100 and its energy distribution using
window shifted by one sample
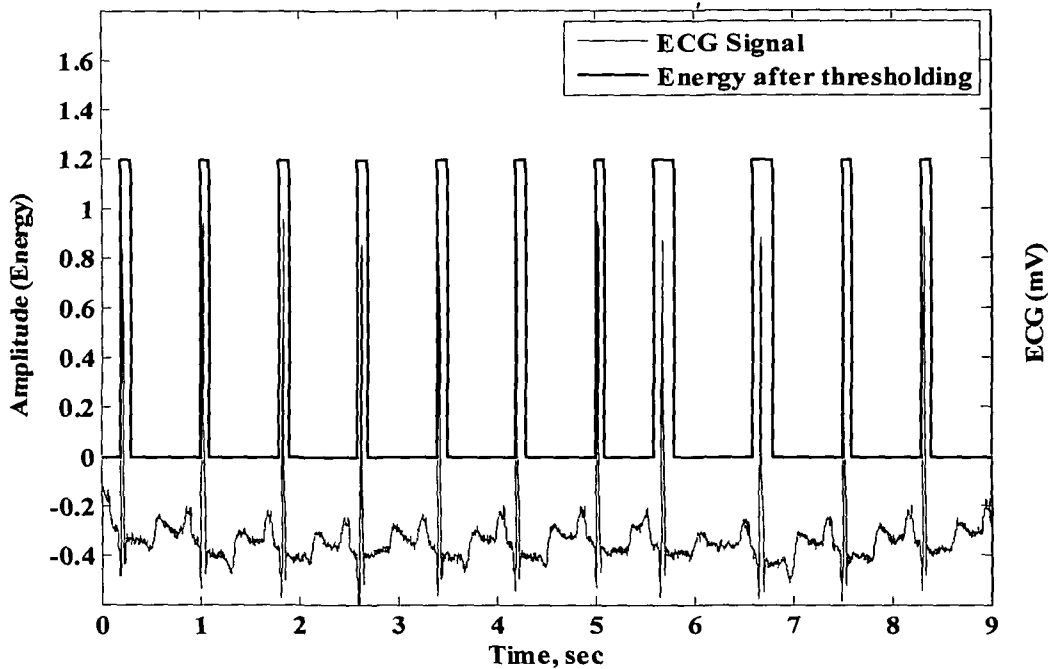
Figure 2.16(a) ECG signal from record 100 and its energy after thresholding
(window shifted by window size)
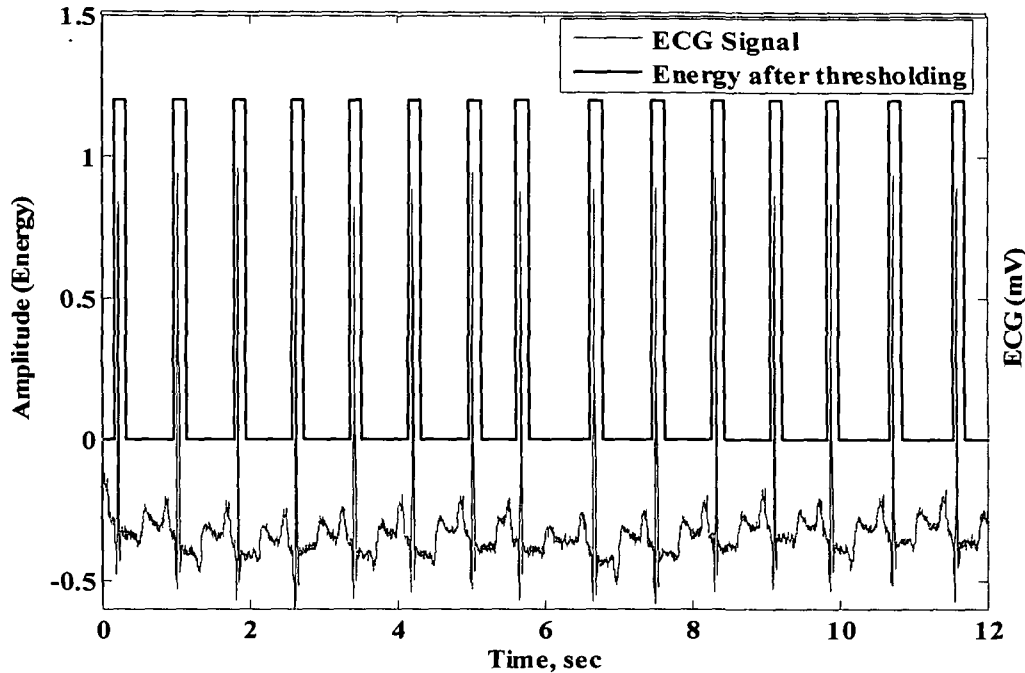
Figure 2.16 (b) ECG signal from record 100 and its energy after thresholding
(window shifted by one sample)

Here, also 'a' refers to beginning of window and is usually taken as 1 since we start from the start of ECG signal and N_1, N_2, N_3.......................N_m comprise of the first and last sample numbers of corresponding window.

The resulting energy signal comprises of length equal to the length of ECG signal. Figure 2.15(b) shows the energy distribution of record 100 by using this method of window shift. It is clear from Figure 2.15(b) that the energy is observed higher in the areas where the QRS complexes exist.

2.1.2.4. Thresholding

The generated energy signals obtained after energy analysis of ECG signal described in section 2.1.2.3 (a-b) undergo thresholding to remove lower energy values. For this purpose, a small value of threshold is applied to the generated ECG signals. The energy samples below this threshold are set to zero and the samples with higher amplitudes are assigned the value equal to 1.2. A threshold value of 5% of maximum amplitude of the signal has been found to be suitable. ECG signal from record 100 and energy signals obtained after thresholding are shown in Figures 2.16(a-b) respectively.

2.1.2.5. R wave positions

The onset samples of remaining thresholded energy windows provide an idea of approximate number of peaks available in the signal and their positions are assumed as R peaks. As no two QRS complexes can be found during less than 200 ms, a window equivalent to 200 ms (72 samples) is skipped in the signal after detection of first peak which gives rise to actual number of R-waves [21]. ECG signal with its energy distribution after thresholding and application of refractory period is shown in Figure 2.17. From Figure 2.17, it is clear that application of refractory period eliminates the centre peak which may occur in thresholded energy signal within the interval of 200 ms. The ECG waveform with the positioned R peaks by windowing method mentioned in section 2.1.2.3(b) for record 100, lead-II is shown in Figure 2.19.
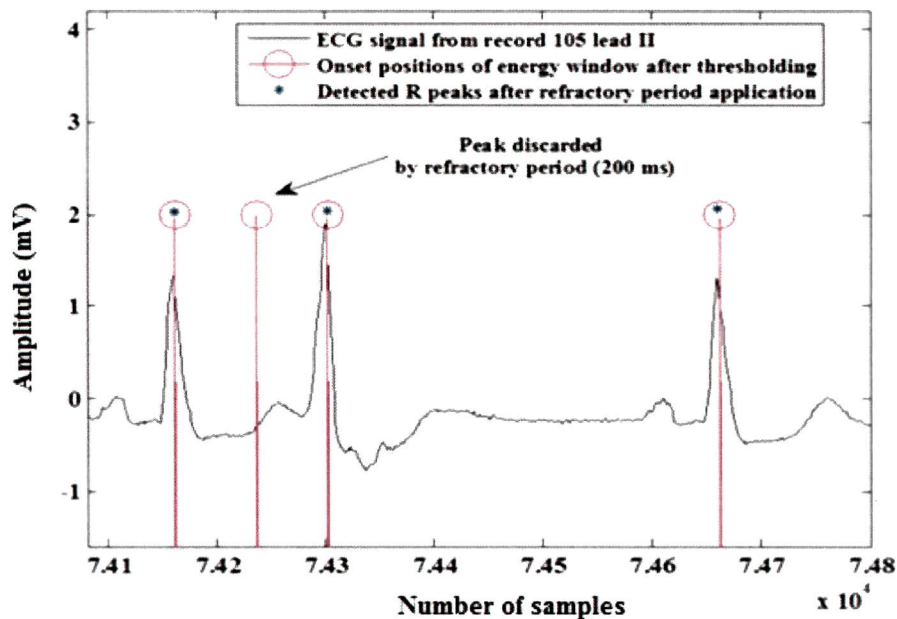
Figure 2.17 ECG signal with its energy distribution after thresholding and application of refractory period
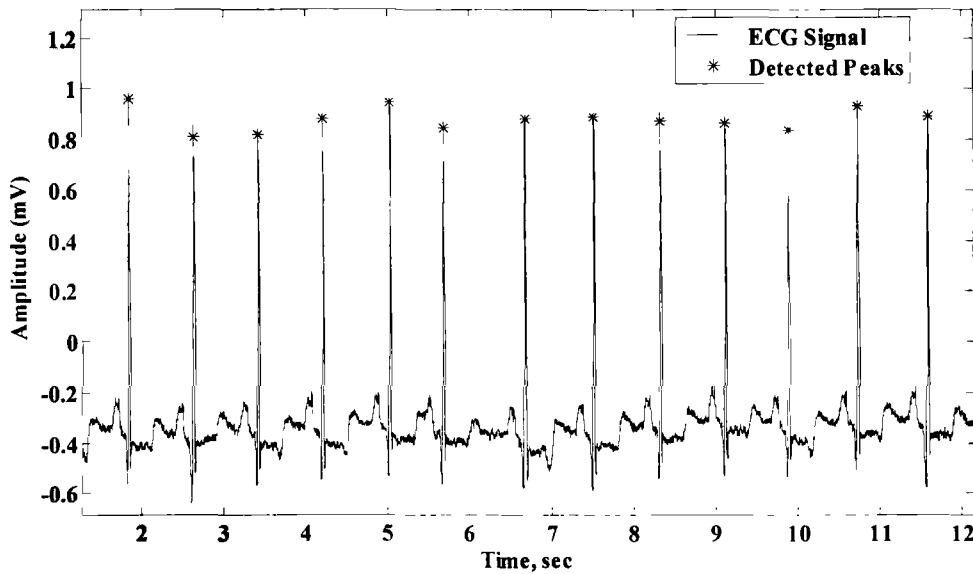
Figure 2.18 Positioned R-peaks in ECG (record-100, lead-II) signal

2.1.3. Q and S wave detection

The selected detail coefficient d4 is further analyzed to detect Q and S wave positions considering R peak positions as reference. Generally, Q and S waves occur in ECG signal as minima before and after R peaks positions. Detection of Q and S wave in d4 signal involves the following steps as shown in the flow chart in Figure 2.19.

i.   Detail coefficient d4 is obtained.

ii.   All minima positions are found if the previous and subsequent samples in the d4 signal have higher amplitude than the sample amplitude between them. Mathematically, this condition can be expressed as -

$$if \ d4(i) < d4(i-1) \ \& \ \ d4(i) < d4(i+1)$$
(2.12)

iii.   Zero crossing points of d4 are detected and their positions are marked along with ECG peaks positions.

iv.   The minima that occur between two zero crossing points before the detected ECG peak position is registered and marked as Q wave.
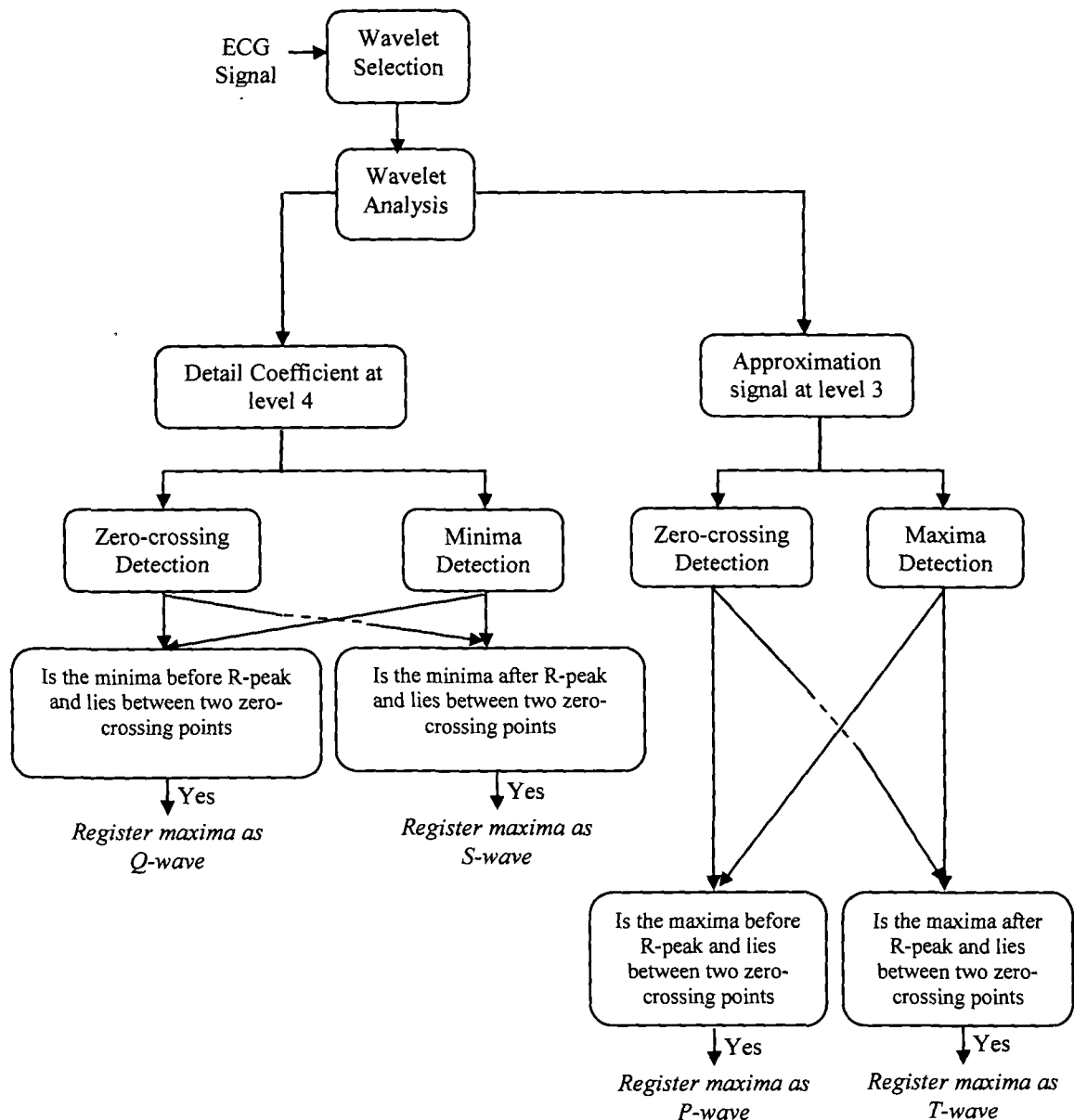
Figure 2.19 Flow chart of P, Q, S and T wave detection algorithm

v. The minima that occur between two zero crossing points after the detected ECG peak position is registered and marked as S wave.

vi. The detected Q and S wave positions are marked on the d4 signal. Detected Q and S wave positions in d4 and original ECG signal along with all minima, ECG peaks and zero crossing points for ECG lead II signal of record 123 are shown in Figure 2.20.
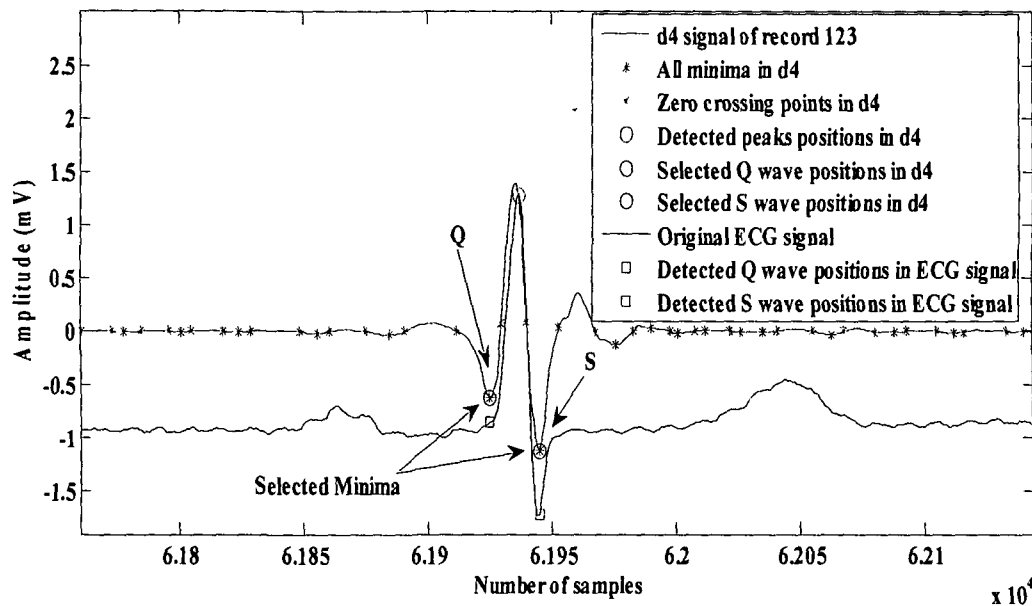
Figure 2.20 Q and S wave positions in the original ECG signal and d4 signal along with all minima, ECG peaks and zero crossing points

2.1.4. P and T wave detection

Detection of P and T waves is performed using approximation coefficient at level 3 (a3). The signal *a3* preserves the shape of P and T wave; however, higher frequency components are removed. Also, it is evident that decomposition of a3 results in d4, therefore, a3 also has entire information of d4 along with P and T waves. Detection of P and T waves in a3 signal involves the following steps as shown in the flow chart in Figure 2.19.

i. The signal a3 is analyzed to detect all maxima positions if the previous and subsequent samples in the resulting signal have lower amplitude than the amplitude of sample between them. Mathematically, this condition can be expressed as -

$$if \ a_3(k) > a3(k-1) \ \& \ a3(k) > a3(k+1)$$
(2.13)

ii. Zero crossing points of a3 are detected and their positions are marked in the a3 signal along with detected ECG peaks positions.

iii. The maxima between two zero crossing points (excluding the maxima detected as ECG peak) that occur before the detected ECG peak position is registered and marked as P wave.

iv.   The maxima between two zero crossing points (excluding the maxima detected as ECG peak) that occur after the detected ECG peak position is registered and marked as T wave.

v.   Detected P and T wave positions in a3 signal and original ECG signal are marked along with detected maxima as shown in Figure 2.21.
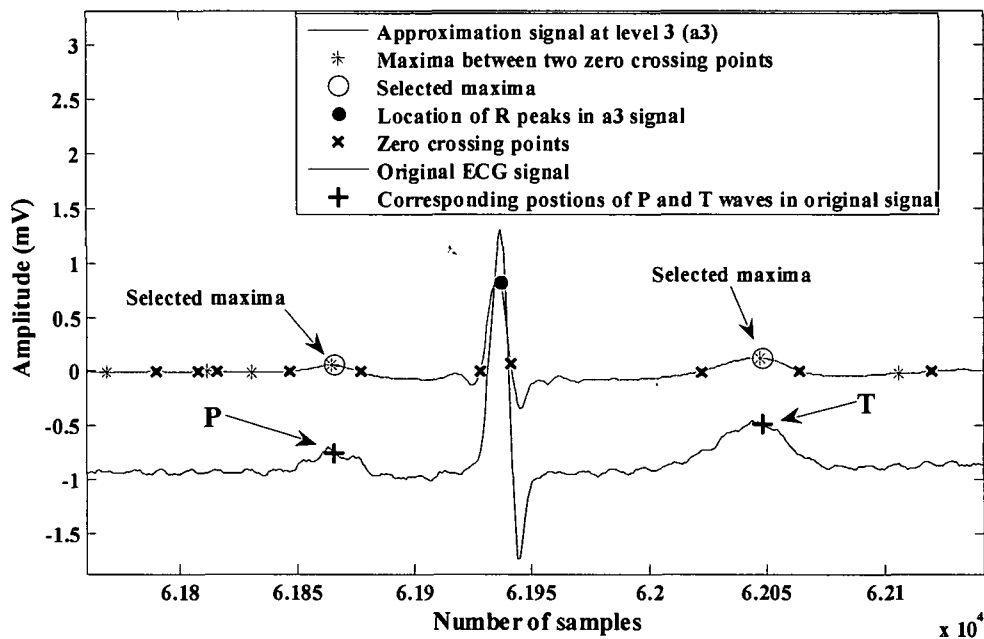
Figure 2.21 Plot of approximation coefficient at level 3 (a3), maxima between two zero crossing points, selected maxima between two zero crossing points, ECG peaks, original ECG signal and detected P and T wave positions in original ECG signal

Although the technique for detection of P,Q,R,S and T wave location on record 100 of the database individually, the techniques are jointly applied to records 115, 123 of MIT-BIH database and one record mgh007 of MGH/MF waveform database to validate the technique. The detected ECG peak positions by wavelet method along with P, Q, S and T wave positions in ECG signal from record 115 and 123 are marked on the original signal and are shown in Figures 2.22(a-b) respectively. The performance of the algorithm is also tested on ECG signal from MGH/MF waveform database for record mgh007 of female patient of 60 years of age. The patient has normal sinus rhythm with first degree heart block @ 85 beats per minute and hemodynamic arterial pressure of 120/80 as specified in the annotation file by expert for ECG signal in mgh007 record. The positions of P, Q, R, S and

T waves are marked on the original ECG lead II signal from record mgh007 and are shown in Figure 2.22(c).
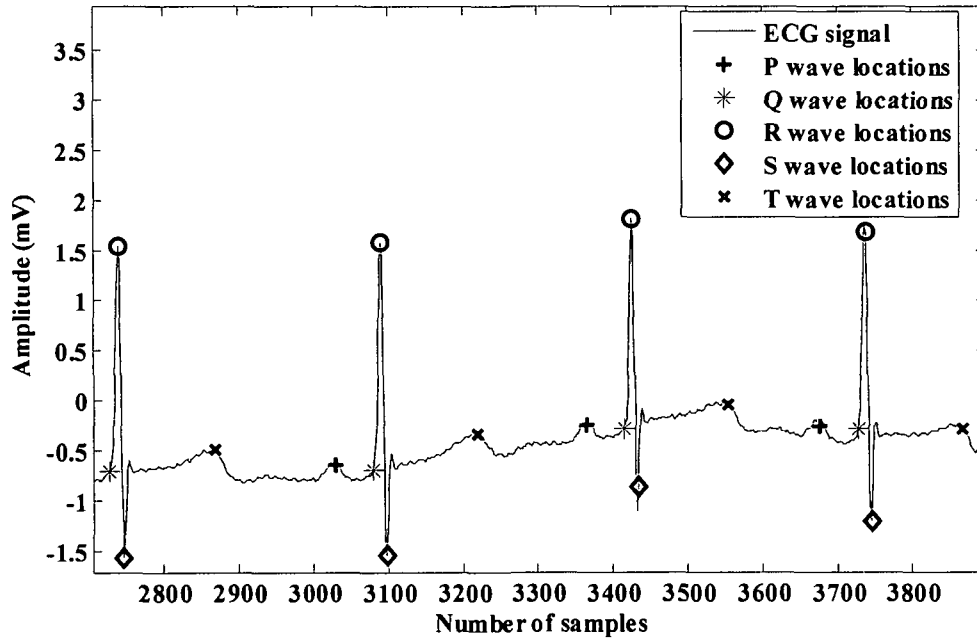
Figure 2.22(a) ECG signal from record 115 along with marked P, Q, R, S and T wave positions
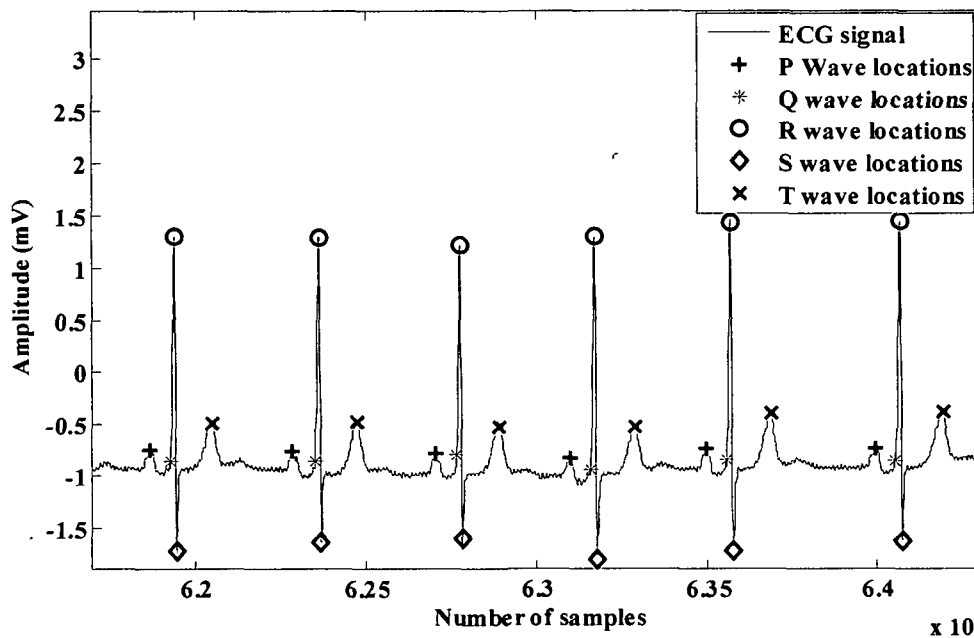
Figure 2.22(b) ECG signal from record 123 along with marked P, Q, R, S and T wave positions
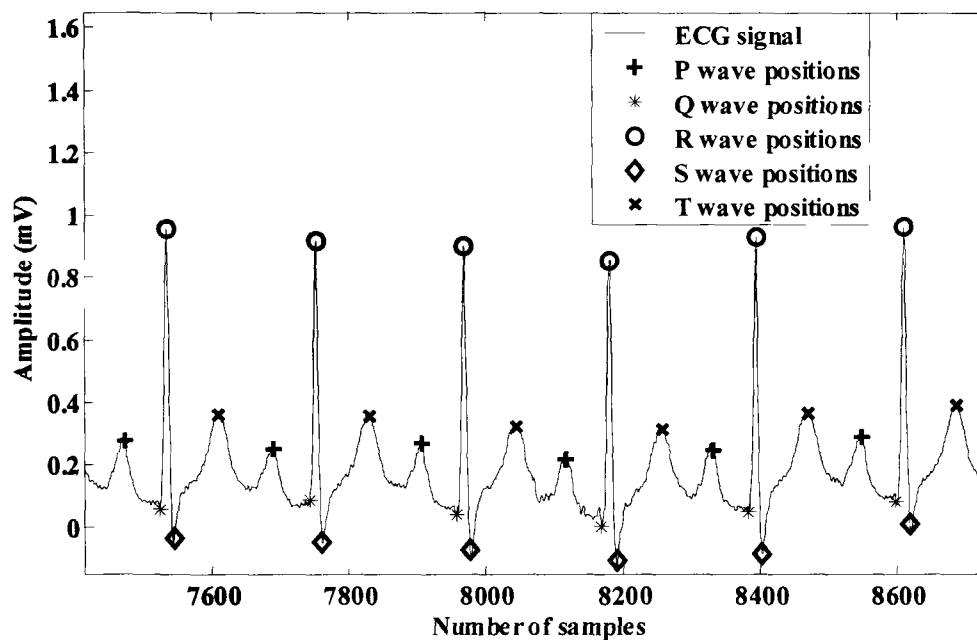
Figure 2.22(c) ECG signal from mgh007 record of MGH/MF waveform database along with marked P, Q, R, S and T wave positions

2.2. PVC detection

The normal heart conduction system is accomplished by the specialized pacemaker cells situated at the sino-atrial node. Sometimes, abnormal conduction of heart may originate in atrial or ventricular region which is independent of the pace set by sino-atrial node. Premature beats may be atrial, atrioventricular (A-V) junctional or ventricular in origin, and may be found in apparently healthy individuals but these beats may become clinically significant if they occur too frequently, originate from multiple foci or found in individuals with proven heart disease. The most common cause of cardiac arrhythmias is the coronary artery disease.

The occurrence of an arrhythmia is unpredictable. Automated systems provide clinicians the tools to be altered in real time if the life threatening conditions surface in their patients. As a result, automatic detection and classification of cardiac electrophysiology using biomedical signal processing techniques have become a critical aspect of clinical monitoring. The purpose of this study is to develop a method to distinguish healthy and

abnormal subjects with premature beats using the combination of RR interval and energy analysis of ECG signal.

2.2.1. Premature ventricular contraction

Premature beats that occur in the ventricular region are called as premature ventricular contractions (PVC) or ventricular premature beats (VPB). PVC is evident in ECG waveform as an abnormal wave shape of QRS complex. Spontaneous occurrence of PVCs in the ECG waveform is the hallmark of drug toxicity and cardiac abnormal conduction. PVCs are diagnosed by prematurity, wide QRS complexes in ECG waveform. The occurrence of these beats is occasional and is accessible in ECG signal as wider or prolonged and premature beats. The sporadic appearance of these beats is not a cause of worry but these beats become of utmost importance when they occur too frequently and require immediate treatment to avoid the problems of consequent life threatening arrhythmias [35]. Figure 2.23 shows two ECG signals from lead II and lead V1 of record 105 of MIT-BIH database with normal ('dot') and PVC ('V') beats annotated by the cardiologist.
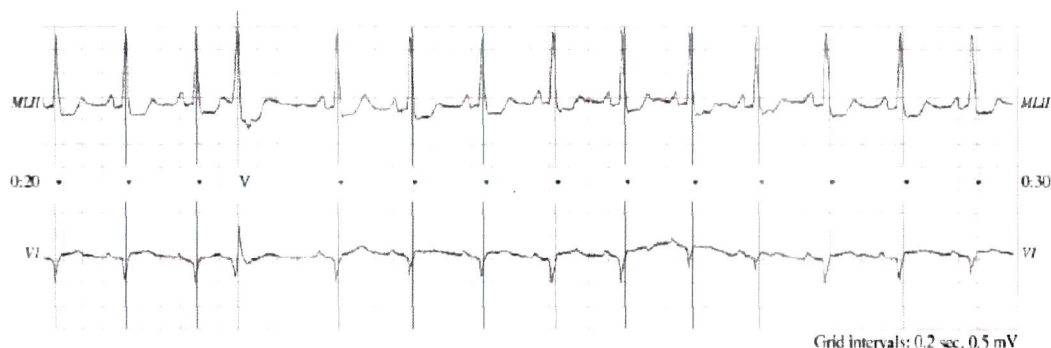
Figure 2.23 ECG signals from record 105 of MIT-BIH database with normal ('dot') and PVC('V') annotated by the cardiologist

Many PVC detection and classification algorithms have been developed so far. The PVC detection algorithms include Discrete Cosine Transform (DCT) and autoregressive modeling [148], symbolic dynamics [149], correlation coefficient in ECG signal [15], Bayesin network framework [150], while the PVC classification algorithms are developed mainly using fuzzy neural network [151], Artificial Neural Networks considering timing information between the detected peaks as a feature set for classification [16].

Other methods of PVC detection include QRS template matching [152, 153], Kth nearest-neighbours rule [154], artificial neural networks, discriminant analysis, Fuzzy logic [153, 155], Karhunen- Loeve transform [156], pattern recognition and optimal parameter selection [157] and fast approximation of vectorcardiogram loop [158].

The results from previous works suggest that the combination of waveform shape and timing interval features is critical for robust detection. We have already proved that window based energy analysis of ECG signal results in a energy signal in which low amplitude peaks are suppressed. Following this fact, it is presumed that energy analysis of ECG signal with PVC beats will result in higher energy values compared to QRS complex. Energy based technique for PVC detection is not reported in literature. Finally, a proficient PVC detection algorithm has been developed which combines RR interval features of ECG by R wave detection in ECG using either wavelet analysis or window based energy analysis of ECG signal using window of 100 ms followed by energy analysis of ECG signal using a window of 600 ms duration. The combination of RR interval feature using either wavelet or energy analysis method using a window of 600 ms contribute to the development of an expert PVC detection algorithm.

2.2.2. Method of PVC detection

Both RR interval and energy analysis techniques are used to detect PVC beats in ECG signal under test. The block diagram of PVC detection algorithm is shown in Figure 2.24.

2.2.2.1. PVC detection based on RR interval

Detection process is performed on the first five minute segment of record 102 and 105, the details of which are given in Table 2.5 and completed in the following steps –

As already stated in section 2.2.1 that PVC appears in the form of abnormal shape of QRS complex and the resulting QRS complex is wider and RR interval between PVC beat and next beat is high. Therefore, RR interval is considered as the prominent feature for PVC detection in ECG signal. An ECG signal with increased RR interval at the location of PVC beats is shown in Figure 2.25.

RR Interval window by wavelet method

RR interval window by energy method
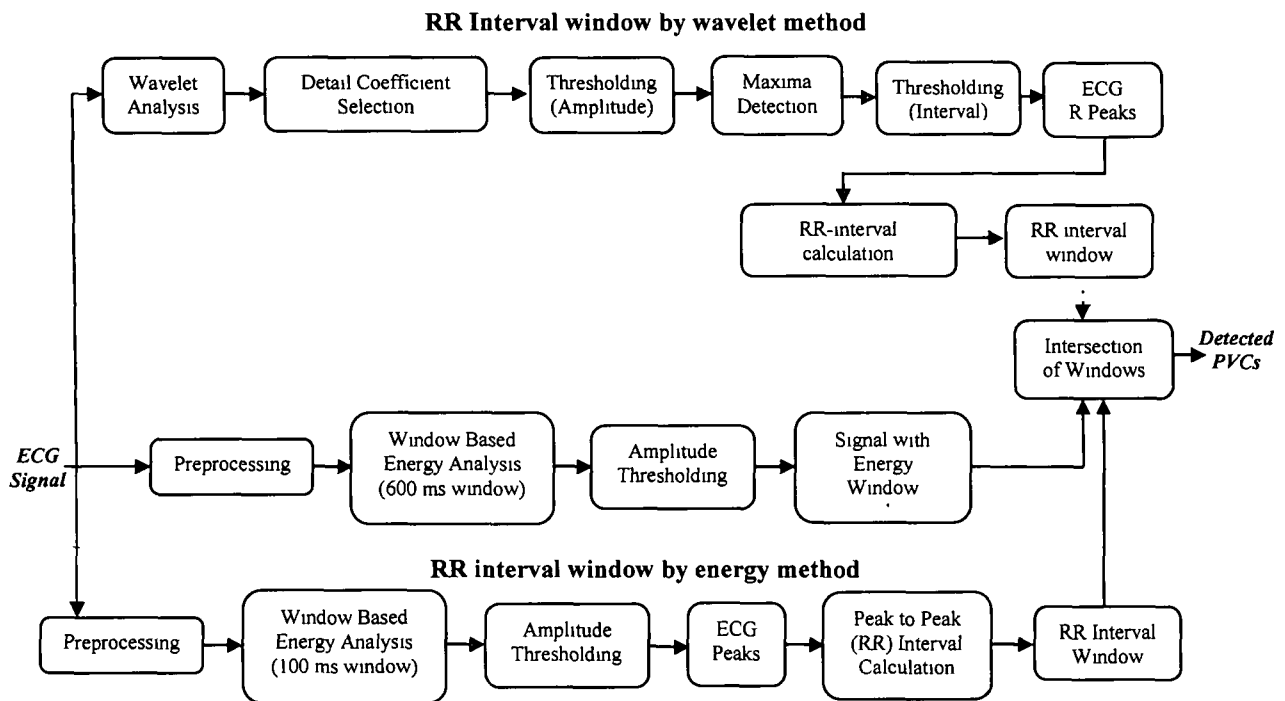
Figure 2.24 Block diagram of PVC detection algorithm

The detected R peaks by any of the methods mentioned in section 2.1.1 and section 2.1.2 (a-b) can be used to calculate RR interval. In this case, we consider the R peaks detected by energy method which comprise of energy analysis of ECG signal using a window of 100 ms and window shifted by one sample as mentioned in section 2.1.2(b). At first, RR interval is calculated between all the consecutive R peaks and a threshold of 75% of the maximum value of RR interval is found suitable. Two consecutive R peaks with peak to peak interval greater than threshold are selected and remaining peak to peak interval values are discarded. A window is placed between these two selected peaks as shown in Figure 2.26.

Table 2.5 Details of record 102 and record 105

Record No & lead information	Age	Sex	Number of PVC beats (annotated)			
			Type of beats	Before 5:00 min	After 5:00 min	Total
102, lead V5	84 Years	Female	Normal	98	01	99
			PVC	01	03	04
			Paced	243	1785	2028
			Pacemaker Fusion	24	32	56
105, lead II	73 years	Female	Normal	405	2121	2556
			PVC	12	29	41
			Unclassifiable	-	05	05

Page 108
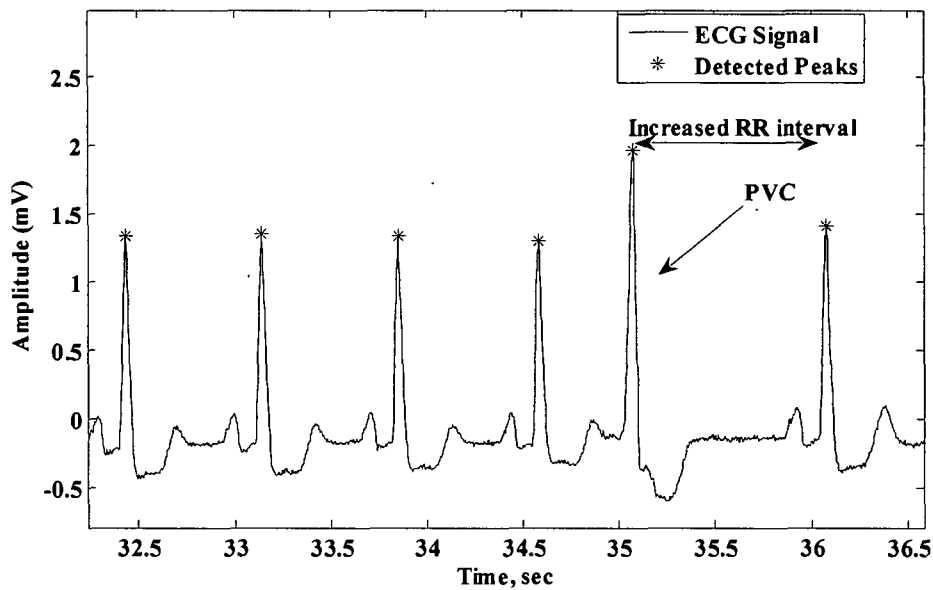
**Figure 2.25 ECG signal from record 105 with detected peaks showing increased
RR interval at the location of PVC beats**
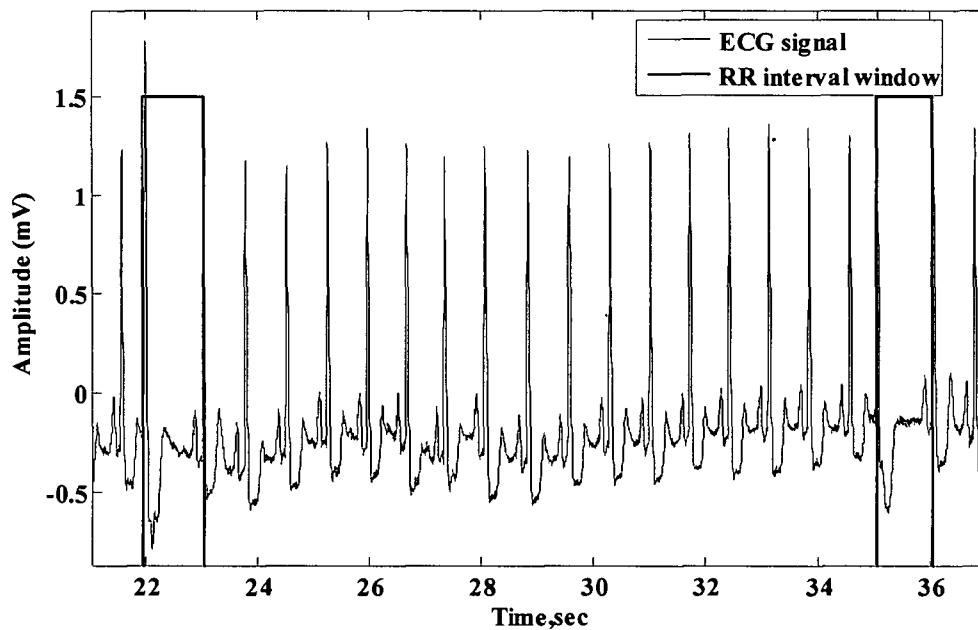
Figure 2.26 ECG signal from record 105 with RR interval windows at the location of PVC beats

The onsets of these windows are selected as possible number of PVCs and their positions
are marked. ECG signals from records 102 and 105 with energy distribution using a
window of 100 ms duration, energy thresholded signal along with selected RR interval

*Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous
Pressure Signals by Signal Processing Techniques*

windows are shown in Figures 2.27 and Figure 2.28 respectively. In this case, two PVC beats are found in record 102 lead V5 and twelve PVCs are found in record 105 lead-II in five minute duration of signal.

2.2.2.2. PVC detection based on energy analysis

Generally the duration of normal QRS-T wave is 350-400 ms. In case of PVCs, this duration is prolonged and QRS-T complexes become wider than normal. PVCs have their energy distribution over longer time interval than normal QRS-T duration. The concept RR interval cannot only be applied to find the premature beats in ECG since there may be other type of beats in the signal such as paced beats. It is not possible to select the two consecutive R peaks having maximum RR interval if the signal possesses more than one PVC beat and also ECG signal characteristics differ from patient to patient. RR interval concept does not work in case of record 102 where RR interval concept detects false PVC as shown in Figure 2.27.
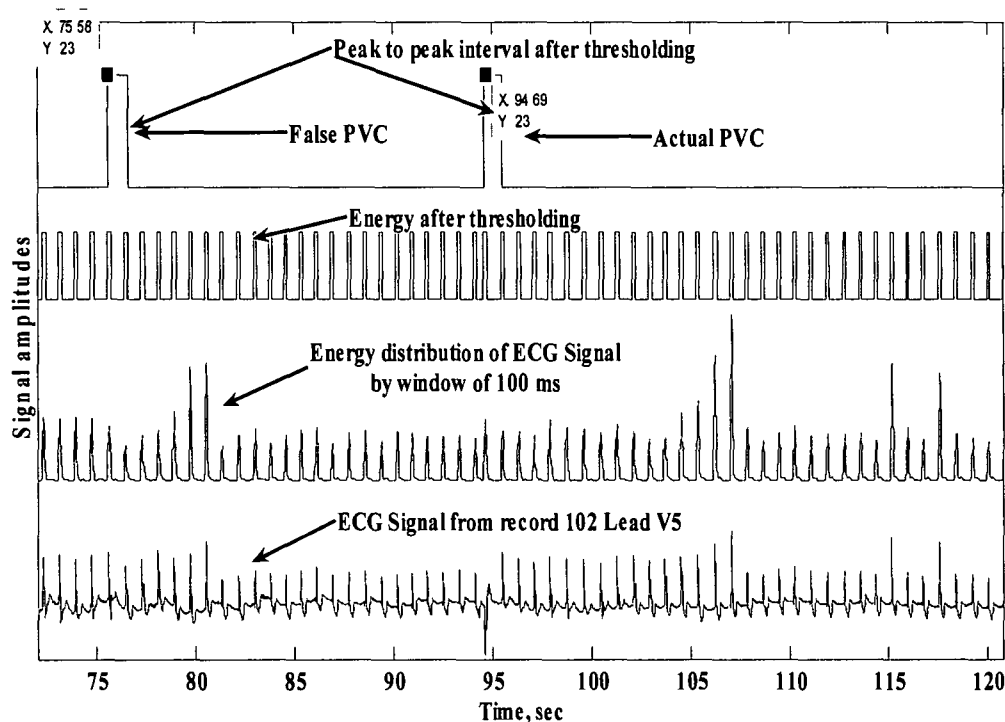
Figure 2.27 ECG signal from record 102 with energy distribution using a window of 100 ms duration, energy signal after thresholding and selected RR interval windows at the possible location of beats
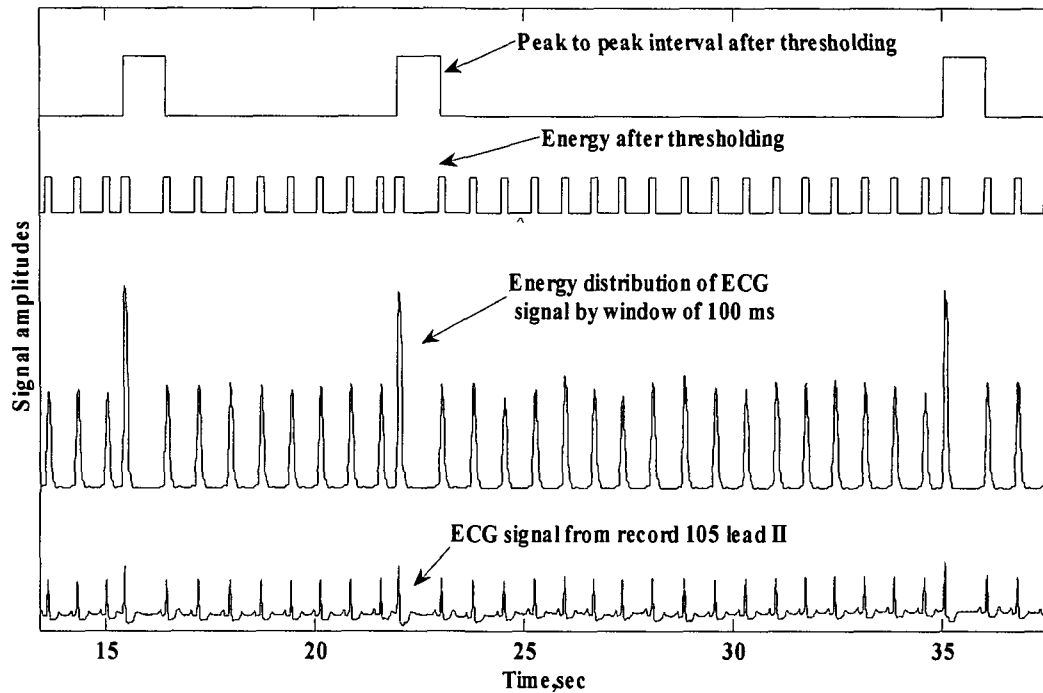
Figure 2.28 ECG signal from record 105 with energy distribution using a window of 100 ms duration, energy signal after thresholding and selected RR interval windows at the possible location of PVC beats
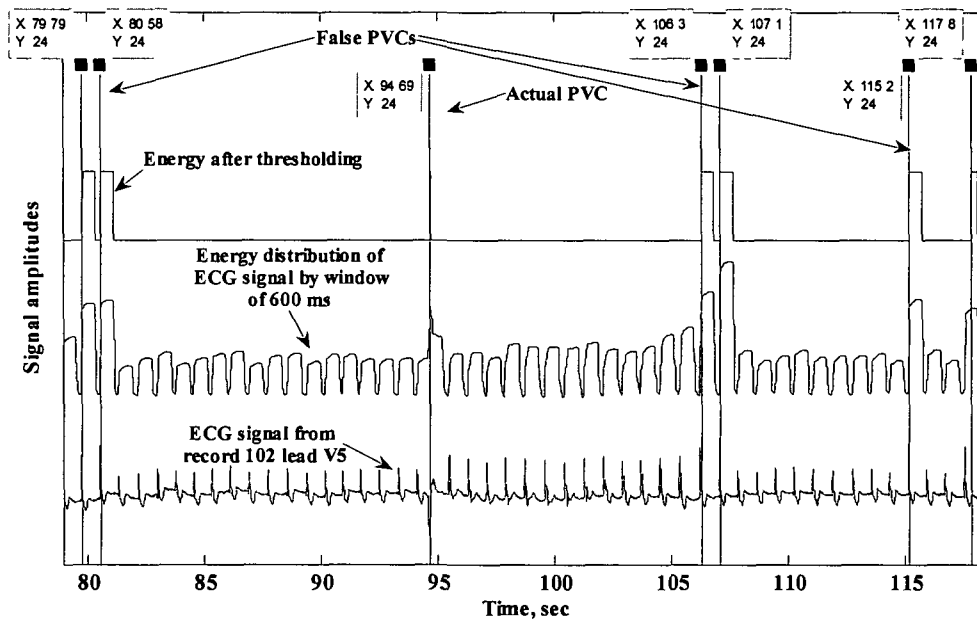
Figure 2.29 ECG signal from record 102 with energy distribution using a window of 600 ms duration, energy signal after thresholding and selected energy windows at the possible location of PVC beats
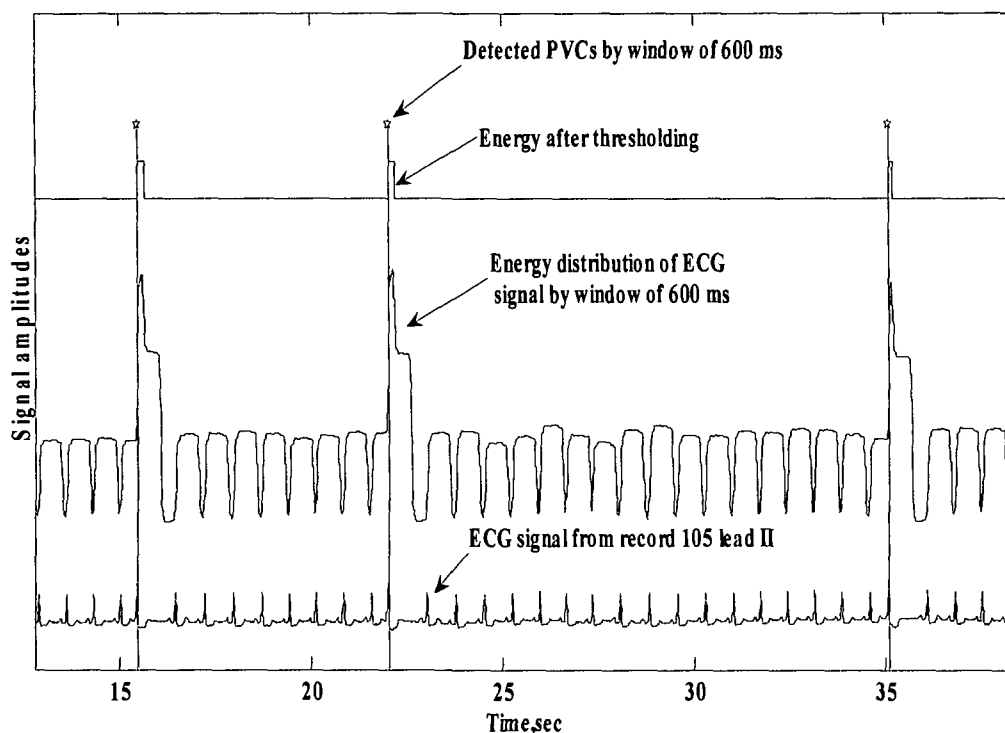
`Figure 2.30 ECG signal from record 105 with energy distribution using a window of 600 ms duration, energy signal after thresholding and selected energy windows at the possible location of PVC beats

Therefore, RR interval concept has to be correlated with some other concept to take a decision of actual number of PVCs. It is clear from section 1.9.2 that energy is defined as the area under the curve and also since PVC is wider in shape, the region associated with PVC in the calculated energy of ECG signal is apparent as higher energy compared to other QRS complexes in the signal under test. Therefore, we propose a new method of energy analysis of ECG signal with extended window size for the detection of PVC beats. The filtered ECG signal is processed for window based energy analysis using a window of 600 ms duration. Energy is calculated for this signal duration and an energy signal is generated as stated in section 2.1.2.3 (a). The length of new energy signal formed is the same as the length of ECG signal under test. The lower energy values are eliminated by applying a threshold value of 65% of maximum amplitude of energy signal.

The resulting signals with possible PVCs for records 102 and 105 are shown in Figure 2.29 and Figure 2.30 respectively. In the energy distribution of record 102, lead V5 in Figure 2.29, energy is observed higher at 6 more locations whereas PVCs do not exist at those locations. It is evident from Figure 2.30 that the energy is higher in the region where PVCs exist. The onsets of energy windows are considered as PVCs in the signals. The

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

onsets of energy windows give fourteen PVCs in record 105 lead-II and seven PVCs in the record 102, lead V5.

2.2.2.3. PVC detection based on intersection of RR interval and energy analysis

As already stated in section 2.2.2.1 and Figure 2.28 that RR-interval concept for PVC detection hold good for one signal (record 105) but does not work as in case of record 102 (Figure 2.27). Similarly, window based energy analysis using 600 ms window results in 06 and 02 false PVCs for records 102 and 105. Therefore, intersection of both RR-interval and energy windows is considered suitable for the detection of actual number and position of PVCs.

Hence in this approach, RR interval window obtained in section 2.2.2.1 is superimposed over energy window as discussed in section 2.2.2.2 and an intersection window is found which gives the actual number and positions of PVCs in the signal as shown in Figure 2.31.
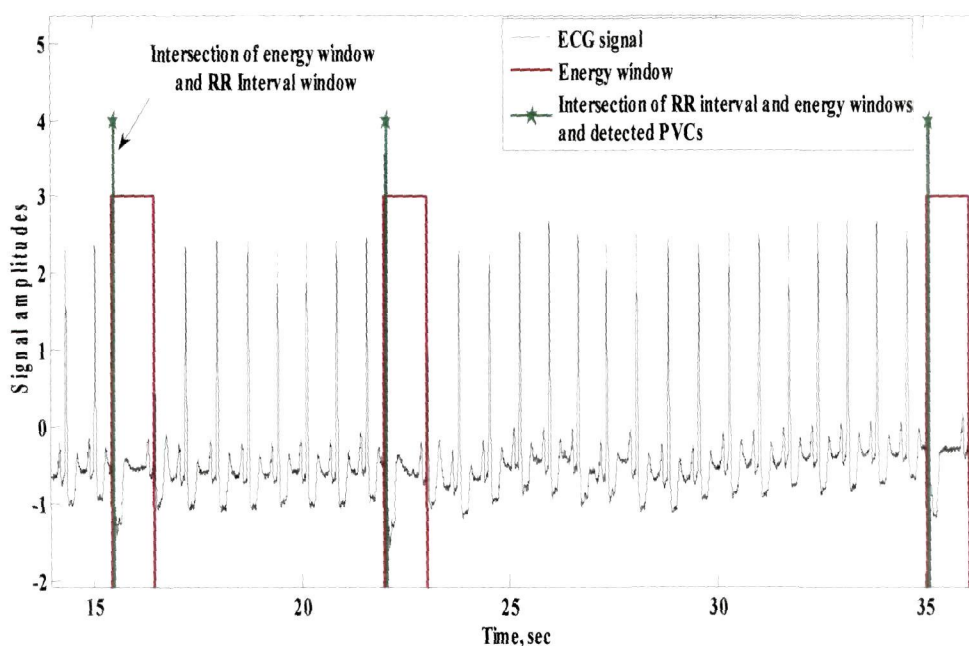
**Figure 2.31 Plot of ECG signal and all detected PVC beats by RR interval
and energy analysis method**
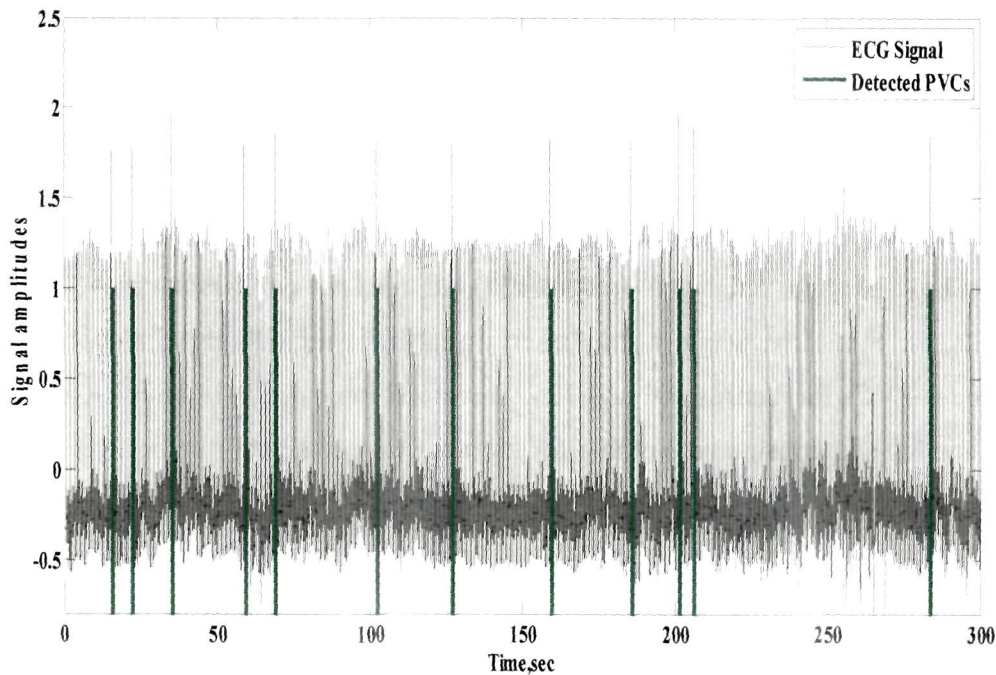
Figure 2.32 Plot of ECG signal from record 105 and all detected PVC beats (12) by RR interval and energy analysis method

The onsets of all the intersection windows are registered as positions of PVCs and their numbers gives rise the number of PVCs in the signal. In Figure 2.32, the location of all twelve PVCs are marked in the original ECG lead II signal from record 105. By using this intersection method one PVC is detected in record 102 and remaining six false PVC beats detected by energy method are discarded.

2.3. Results and conclusion

The results of developed algorithms for feature extraction and PVC detection are presented in the following sections –

2.3.1. ECG peak detection using wavelet technique

The algorithm has been tested on MIT-BIH arrhythmia database in which each recording is of 30 minutes duration. Forty four records from first lead signal (lead II signal) including records 102 and 104 (which have lead V5 signal as first lead) are tested for R peaks detection. All beats positions in MIT-BIH database are annotated by experts. The

Page 114

performance of the developed algorithm is tested with expert annotations available on the database. The beat positions that are correctly detected are named as *'true peaks'* whereas the beats positions that are missed by the algorithm are named as *'missed peaks'*. 'False peaks' are the peaks which are not actually available but reported as the algorithm output. The overall accuracy obtained using db6 wavelet, out of total 102869 peaks, 102773 true peaks are detected and 396 peaks are missed by the algorithm whereas 138 false peaks are detected. The algorithm performance is judged based on three measures viz accuracy (A), sensitivity (S_e) and positive predictivity (P_P) [159].

The overall accuracy for R-wave detection is given by the relation [16]

$$A = \left(1 - \frac{N_e}{N_b}\right) \times 100 \qquad (2.14)$$

Where A denotes the average accuracy for all the forty four records under test and the variables N_e and N_b represent the total number of detection errors and peaks available in the records. The detection error N_e is the sum of missed peaks (M_P) and false peaks (F_P). The sensitivity and positive predictivity of detector are given by the following relation [45]-

$$S_e = \frac{T_P}{T_P + M_P} \times 100 \qquad (2.15)$$

$$P_P = \frac{T_P}{T_P + F_P} \times 100 \qquad (2.16)$$

Where 'T_P' denotes the number of correctly detected peaks where as 'M_P' denotes the peaks missed by the detector and 'F_P' represents the number of detected false peaks. Sensitivity measures how successfully the detector recognizes the peaks and positive predictivity measures how exclusively the detector detects the peaks in the signal. The algorithm achieved an accuracy of 99.48 %, sensitivity of 99.62 % and positive predictivity of 99.87 % respectively over all the forty four records of MIT-BIH database. The results for R-wave detection using db6 wavelet are given in Table 2.6.

2.3.2. ECG peak detection using energy analysis technique

The algorithm has been tested on first 5 minute segments of forty four downloadable records out of total forty eight records of MIT-BIH database. The records 102 & 104 do not

have modified lead II signals. The accuracies obtained for these records individually using energy analysis when window energy analysis is done as per section 2.1.2.1(a) is 99.73 % and 98.93 % respectively. The accuracies obtained for these records are 100 % & 98.15 % respectively when the energy analysis of ECG signal is followed as per section 2.1.2.1(b).

The overall accuracy, sensitivity and positive predictivity using all forty four records of database for section 2.1.2.1(a) are found to be 98.17 %, 98.82 %, 99.36 % and 98.63 %, 99.36 % and 99.28 % respectively for energy calculation as per section 2.1.2.2(b). The summary of results by window shifted by window size and window shifted by one sample is given in Table 2.7.

The consolidated results of wavelet based and energy based peak detection are shown in Table 2.8. From Table 2.8, it is clear that wavelet based technique outperforms to energy based peak detection method.

It is mentioned in literature survey in section 2.0 that ECG feature extraction has been carried out by several researchers using different techniques [20-29]. Further, a comparison of the results of proposed wavelet based technique with the techniques mentioned in [21,24,30] has been carried out in terms of accuracy, sensitivity and positive predictivity. The details of comparative study are mentioned in Table 2.9. It is clear from Table 2.9 that proposed wavelet based technique performs better in terms of accuracy compared to results of Sahambi et al.[21] for 8 records analyzed in their work. The results of the proposed technique are better in terms of sensitivity and positive predictivity (98.37 % and 98.15 %) for 22 records as compared to that of Mahmoodabadi et al. [24]. Lee R.G. et al. [30] reported an accuracy of 95.26 % for 42 records whereas our algorithm achieves an accuracy of 99.53 % on 42 records of database excluding the records 108, 233 which are not tested by Lee R.G. et al.

2.3.3. PVC detection

The performance of the algorithm is evaluated for thirty seven records on MIT-BIH database in which PVC beats exist. Eleven records 101, 103, 112, 113, 115, 117, 122, 212, 220, 222 and 232 do not have PVC beats. The overall accuracy of detection is determined by using the equation 2.14.

Where, N_e states for the detection errors which is defined as the sum of the missed and false beats and N_b stands for the total number of PVC beats in five minute segment of ECG signal from MIT-BIH database. Out of total 1183 PVCs, 1164 PVCs are correctly detected whereas 20 PVCs are missed and 18 false PVCs are detected. Therefore, achieved overall accuracy of detection is 96.79 %. The sensitivity and positive predictivity are calculated from equations 2.15 and 2.16 respectively. The achieved values of sensitivity and positive predictivity are 98.31 % and 98.48 % respectively. The results for PVC detection using the combination of energy analysis of ECG signal by 100 ms and 600 ms windows are displayed in Table 2.10.

As stated in section 2.2.1, PVC detection has been performed by many authors in [152-158]. It is to be noted that PVC detection has been carried out by the authors on 37 records in [152 -155, 157-158], however, Herrereo et al. [156] has performed PVC detection on 33 records. Therefore, to justify the results of our proposed technique of PVC detection, a comparison of results of proposed method is carried out with other techniques [152-158]. The results of comparative study are shown in Table 2.11. From Table 2.11, it is clear that the proposed method performs better than the results shown in [153-154, 157-158] in terms of sensitivity for all 37 records under study. However, the obtained sensitivity value is close to the results shown in [152]. In terms of sensitivity and positive predictivity values shown in [155], the proposed method has shown better sensitivity value and the positive predictive values are nearly matching.

Table 2.6 Results of R-wave detection using Daubechies (db6) wavelet

Record No.	Beat Annotations from database	True peaks detected	Missed peaks	False peaks
100	2273	2270	03	-
101	1865	1864	01	-
102 (V5)	2187	2186	01	02
103	2084	2080	04	01
104 (V5)	2229	2227	02	05
105	2572	2521	51	04
106	2027	2004	23	01
107	2137	2137	-	08
108	1774	1741	33	10
109	2532	2532	-	13
111	2124	2124	-	-
112	2539	2520	19	-
114	1879	1878	01	10
115	1953	1953	-	-
116	2412	2412	-	03
118	2288	2283	05	01
119	1987	1987	-	14
121	1863	1850	13	-
122	2476	2472	04	-
123	1518	1518	-	-
124	1619	1608	11	02
200	2601	2601	-	06
202	2136	2135	01	02
203	2980	2830	150	10
205	2656	2650	06	09
207	2332	2328	04	-
208	2955	2951	04	01
209	3005	3005	-	16
210	2650	2949	01	01
212	2748	2748	-	03
213	3251	3249	02	-
214	2262	2259	03	01
215	3363	3350	13	-
217	2208	2198	10	-
219	2287	2286	01	02
220	2048	2048	-	-
221	2427	2427	-	03
222	2483	2481	02	-

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

223	2605	2600	05	-
228	2053	2053	-	05
230	2256	2256	-	03
231	1573	1573	-	02
233	3079	3069	10	-
234	2573	2560	13	-
Total number of records = 44	102869	102773	396	138

Table 2.6 concluded

Table 2.7 Results of R-wave detection using energy analysis technique

Performance	Window shifted by window size	Window shifted by one sample
Accuracy	98.17%	98.63%
Sensitivity	98.82%	99.36%
Positive Predictivity	99.36%	99.28%

Table 2.8 Comparison of results of wavelet and energy based R peaks detection

Performance	Wavelet method	Window shifted by window size	Window shifted by one sample
Accuracy	99.48 %	98.17%	98.63%
Sensitivity	99.62 %	98.82%	99.36%
Positive Predictivity	99.87 %	99.36%	99.28%

Table 2.9 Comparison of results of proposed wavelet based R wave detection method with other techniques

Performance	For 8 records (100, 101, 102, 103, 104, 105, 106, 107)		For 22 records (100, 101, 103, 105, 106, 107 118, 119, 200, 202, 203, 205, 207, 208, 209, 210, 212, 213, 214, 215, 217, 219)		For 42 records (All records mentioned in Table 2.5 except 108, 233)	
	Sahambi et al.[21]	Proposed method	Mahmoodabadi et al.[24]	Proposed method	Lee R.G. et al.[30]	Proposed method
Accuracy	98.94 %	99.38 %	-	-	95.26 %	99.53 %
Sensitivity	-	-	98.37 %	99.46 %	-	-
Positive Predictivity	-	-	98.15 %	99.81 %	-	-

Table 2.10 Results of PVC detection

Record No	Age/ Sex	Lead information	Number of PVCs In 5 min	Detected PVCs			True PVC	Missed	False PVC
				RR Interval	PVC by energy method	PVC by combining both methods			
100	69/M	MLII, V5	01	01	182	01	01	-	-
102	84/F	V5/V2	01	01	07	01	01	-	-
104	66/F	V5/V2	01	02	03	02	01	-	01
105	73/F	MLII/V1	12	12	14	12	12	-	-
106	24/F	MLII/V1	60	58	56	55	55	05	-
107	63/M	MLII/V1	02	03	02	02	02	-	-
108	87/F	MLII/V1	03	05	04	04	03	-	01
109	64/M	MLII/V1	06	12	10	09	06	-	03
111	47/F	MLII/V1	01	01	333	01	01	-	-
114	72/F	V5/ML II	13	12	14	11	11	02	-
116	68/M	MLII/V1	11	10	10	10	10	01	-
118	69/M	MLII/V1	03	05	04	03	03	-	-
119	51/F	MLII/V1	80	110	94	81	80	-	01
121	83/F	MLII/V1	01	01	289	01	01	-	-
123	63/F	MLII/V5	01	01	02	01	01	-	-
124	77/M	MLII/V4	20	19	21	19	18	02	01
200	64/M	MLII/V1	126	135	205	127	127	-	01
201	68/M	MLII/V1	20	23	204	21	16	04	05
202	68/M	MLII/V1	04	08	12	04	04	-	-
203	43/M	MLII/V1	71	79	95	72	71	-	01
205	59/M	MLII/V1	06	15	18	06	06	-	-
207	89/F	MLII/V1	101	135	120	102	101	-	01
208	23/F	MLII/V1	168	202	292	168	168	-	-
209	62/M	MLII/V1	01	01	87	01	01	-	-
210	89/M	MLII/V1	30	32	58	30	29	01	-
213	61/M	MLII/V1	25	24	32	23	23	02	-
214	53/M	MLII/V1	44	48	54	45	43	01	01
215	81/M	MLII/V1	33	39	94	33	33	-	-
217	65/M	MLII/V1	22	25	35	23	22	-	01
219	NA/M	MLII/V1	13	15	15	13	13	-	-
221	83/M	MLII/V1	80	105	134	80	80	-	01
223	73/M	MLII/V1	18	22	24	18	18	-	-
228	80/F	MLII/V1	60	72	62	60	60	-	-
230	32/M	MLII/V1	01	03	14	01	01	-	-
231	72/F	MLII/V1	04	04	05	04	04	-	-
233	57/M	MLII/V1	139	205	165	138	137	02	-
234	56/F	MLII/V1	01	02	22	01	01	-	-
Total number of PVCs			1183				1164	20	18

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

Table 2.11 Comparison of results of proposed PVC detection method with other authors work

Proposed Method (37 records)		Krasteva et al. [152] (37 records)		Christov et al. [154] (37 records)		Bortolan et al. [153] (37 records)		Herrero et al. [156] (33 records)		Jekova et al.[157] (37 records)		Jekova et al.[155] (37 records)		Jekova et al. [158] (37 records)	
Se %	PP %	Se %	PP %	Se %	PP %	Se %	PP %	Se %	PP %	Se %	PP %	Se %	PP %	Se %	PP %
98.31	98.48	98.40	-	97.3	-	97.7	-	99.62	-	90.5	-	97.99	98.96	93.3	-

Publication on this chapter

[1] Pachauri, A., & Bhuyan, M. Robust detection of R-wave using wavelet technique, in International Conference on Signal Processing, Communication and Networking (ICSPCN' 2009), Singapore, 335-339, World Academy of Science, Engineering and Technology 56, 2009, 901-905.

[2] Pachauri, A., & Bhuyan, M. Wavelet and energy based approach for PVC detection, in IEEE International Conference on Emerging Trends in Electronic and Photonic Devices & Systems (ELECTRO' 2009), Varanasi, India, 258-261.

[3] Pachauri, A., & Bhuyan, M. PVC detection by energy analysis", in 2nd International Conference on RF and Signal Processing Systems (RSPS' 2010), Guntur, India, 380-384.

[4] Pachauri, A., & Bhuyan, M. A new approach to ECG peak detection, in International Conference on Biomedical Engineering and Assistive Technologies (BEATS' 2010), Jalandhar, Punjab, India.

CHAPTER

Feature Extraction of Arterial Blood Pressure and Central Venous Pressure Signals

3

3.0. Introduction

Present trend of cardiac diagnosis is conceded manually by the cardiologists using mainly ECG waveform. So far ECG has been used to diagnose the heart condition by the cardiologists. Like ECG, ABP and CVP waveform are also rich in pathological information about cardiovascular function. As stated in section 2.2, that life threatening disturbances such as PVC are diagnosed by studying the change in the intervals and amplitudes of ECG signal components. Analysis of arterial blood pressure and central venous waveforms can help to supplement the information required by the cardiologists. It is already stated in section 1.6.4 and 1.7.4 (Table 1.6 and 1.7) that ECG as well as hemodynamic waveforms (ABP and CVP) depicts cardiac physiology. In certain cases, abnormal conduction of heart displayed in one signal waveform i.e. ECG signal, is also present in other signal waveform such as ABP, CVP (Figure 1.8). Analysis of ectopic beats is generally carried out using ECG signal but changes in ABP waveform are also present in case of these ectopic beats [34,35]. Therefore, analysis of ABP waveform in addition can be used to illustrate the cardiac signatures efficiently in the diagnosis of cardiac health. Parallel analysis of ABP waveform can help to develop robust method for the diagnosis of certain heart diseases such as ectopic beats those result in sudden decrease in systolic blood pressure [34,35].

As seen from the literature [40-46], that most of the ABP feature extraction developed so far include the detection of either peaks [44] or onsets [43,46] or dicrotic notch [41] only and there is only one algorithm in which detection of all three features is proposed [45]. There is no algorithm reported in literature for detection of all four features of ABP signal.

Wavelet and energy based feature extraction technique applied on ECG signal have been found suitable as discussed in chapter 2. The merits of these techniques have motivated us to apply for ABP and CVP signals. Systolic peak is the most prominent feature in ABP waveform that is apparent as the highest amplitude, thereby it is considered as the basis for the detection of all other features such as onset, dicrotic notch and dicrotic peak of ABP signal.

In this chapter, we suggest a wavelet transform based technique for full characterization of ABP waveform that is robust to physiological interferences and varying signal amplitude, does not require any preprocessing of signals, developed and validated on open access MGH/MF waveform database [139], Fantasia database [137], MIT-BIH polysmographic database [138] and CSL database [12]. Moreover selection of detail coefficients after wavelet decomposition has been justified by energy, frequency and cross-correlation analysis of detail coefficients. Further application of window based threshold overcomes the setback of missing peaks due to large variations in the signal amplitude at any particular instant. The developed algorithm is applicable to any signal length. We have implemented the algorithm in Fantasia database, MIT BIH Polysmographic database and selected segments of MGH/MF and CSL databases for manual validation of algorithm.

In this chapter, another algorithm for the detection of systolic peaks in the Arterial Blood Pressure (ABP) signal using energy analysis of ABP signal is also proposed. The proposed algorithm deals with the implementation of energy based approach for the detection of systolic peaks in ABP signal. Energy based peak detection algorithm has been developed on five minute segment of ABP signals for the duration for which ECG signals are also available. The algorithm has been validated with the available expert annotations for ECG lead II signal for five minute segment.

Like ECG and ABP signals, a great deal of cardiac physiological information is available in CVP signal. CVP monitoring is used for assessment of blood volume and right heart function. Normal mechanical events of the cardiac cycle are responsible for the sequence of waves seen in a typical CVP trace. CVP signals are used to know normal and abnormal heart behaviour more precisely than ABP waveforms by the analysis of constituent five phasic events, three peaks ('a', 'c', 'v') and two descents ('x', 'y'). To address a few abnormal heart condition, atrioventricular dissociation, tricuspid regurgitation, tricuspid

stenosis, right ventricular ischemia, pericardial constriction, cardiac tamponade etc. can be characterized with the variations in its constituent signal components which are discussed in Table 1.5. However, present trend of cardiac diagnosis is dependent more on the analysis of ECG signal whereas there is physical correlation with ABP signal to some extent but there is no attempt reported in literature pertaining to feature extraction of CVP signal. The reason why CVP analysis is helpful is - In certain conditions ECG signals may be excessively noisy or ECG acquisition may not be possible due to surgical dressing of patients. The noise in CVP which is a pressure signal, is mechanical in nature whereas electrical noise interferes more in ECG [42]. The cardiac signatures can be well understood to a great extent if an algorithm for CVP features is also available which can assist the cardiologists to know cardiac state precisely. Parallel analysis of ECG signal with ABP signal or CVP signal or a combination of all these three signals can formulate robust algorithms for reducing the false alarms for arrhythmia or other diseases in intensive care units. These algorithms can not only reduce the time taken by the cardiologists to a great extent but also can assist them to update the knowledge of cardiac internal signatures to a large extent. In normal central venous pressure waveform, x-descent is the most prominent feature which has least amplitude compared to other features in the signal, therefore, it sets up the basis of CVP feature extraction.

The use of wavelet based technique by changing the wavelet for the analysis and selection of appropriate detail/ approximation coefficient is again found to be a promising method for CVP feature extraction in this research. We present a wavelet based technique for CVP waveform delineation using daubechies db4 wavelet. The selection of relevant detail coefficient is again supported by energy, frequency and cross-correlation analysis of detail coefficients with the original CVP signal. Here, window based negative amplitude thresholding is used to detect 'x'-descent precisely which is further used to mark remaining peaks ('a', 'c', 'v') and descent ('y') waves of CVP signal.

3.1. ABP feature extraction

Feature extraction of ABP signal is carried out using wavelet and energy analysis technique as mentioned below –

3.1.1. Wavelet based ABP peak detection

We have used mgh007 record of MGH/MF waveform database with 8 signals, the plot of which is shown in Figure 3.1 and description of record is given in Table 3.1. The raw signals are converted to physical units by using the following conversion rule.

$$Physical\ Unit\ = \frac{Raw\ Signal\ Amplitude - Base}{Gain} \qquad (3.1)$$

It is clear from Figure 3.2 that raw ABP signal has amplitude -400 to 600 on Y-axis whereas signal after conversion into physical units has amplitude 80-140 mm Hg. In the database, ECG signals are available for complete duration of the signal whereas ABP signals are missing in certain duration. Therefore, the samples of ABP and ECG II for the duration when both signals are available are used for analysis for matching the number of peaks in ABP signals using ECG annotations by expert. ABP signal is processed for wavelet analysis and selected relevant detail coefficients are used for window based amplitude thresholding. Maxima are detected in the signal obtained after common point detection in two selected detail coefficients and an threshold is applied to detect the ABP peaks. The block diagram of ABP peak detection algorithm is shown in Figure 3.3 and schematic representation of steps involved in the detection method is given in Figure 3.4. The detection process is performed on ABP signal of mgh007 record of MGH/MF database and 'abp1' signal of CSL database and completed in the following steps –

Table 3.1. Description of mgh007 record of MGH/MF waveform Database

Record	Signal	Gain	Base	Units
Mgh007	ECG lead I	1341	242	mV
	ECG lead II	1382	-618	mV
	ECG lead V	1295	-452	mV
	ABP	12.17	-1218	mmHg
	PAP	19.26	-1016	mmHg
	CVP	19.04	-1007	mmHg
	Resp. Imp.	1000	0	mV
	CO_2	1000	0	mV
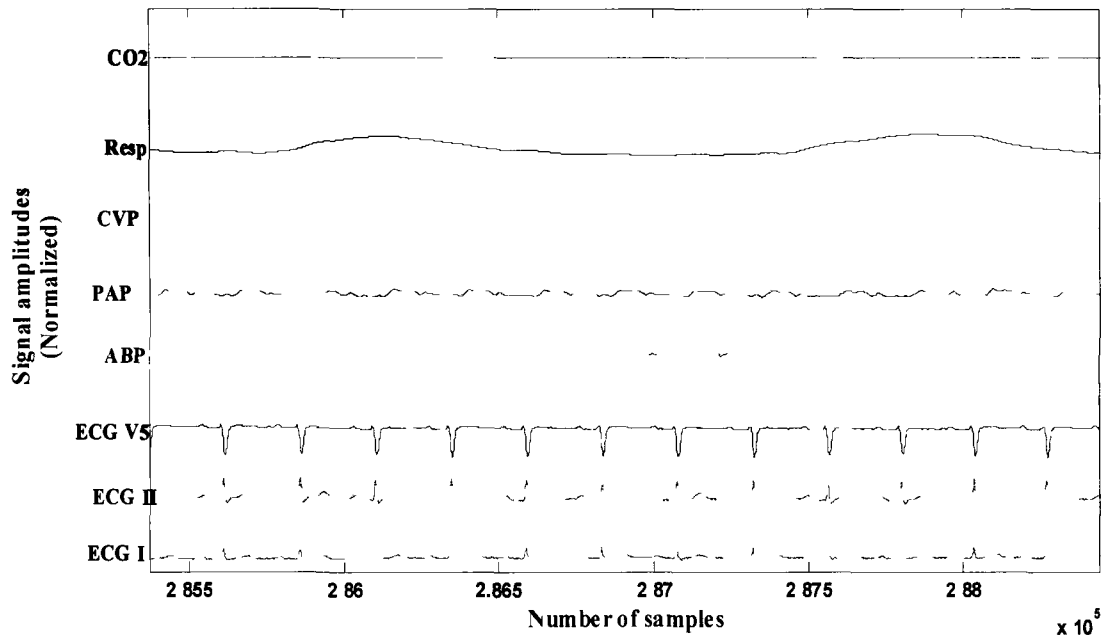
Figure 3.1 Plot of all signals of mgh007 record
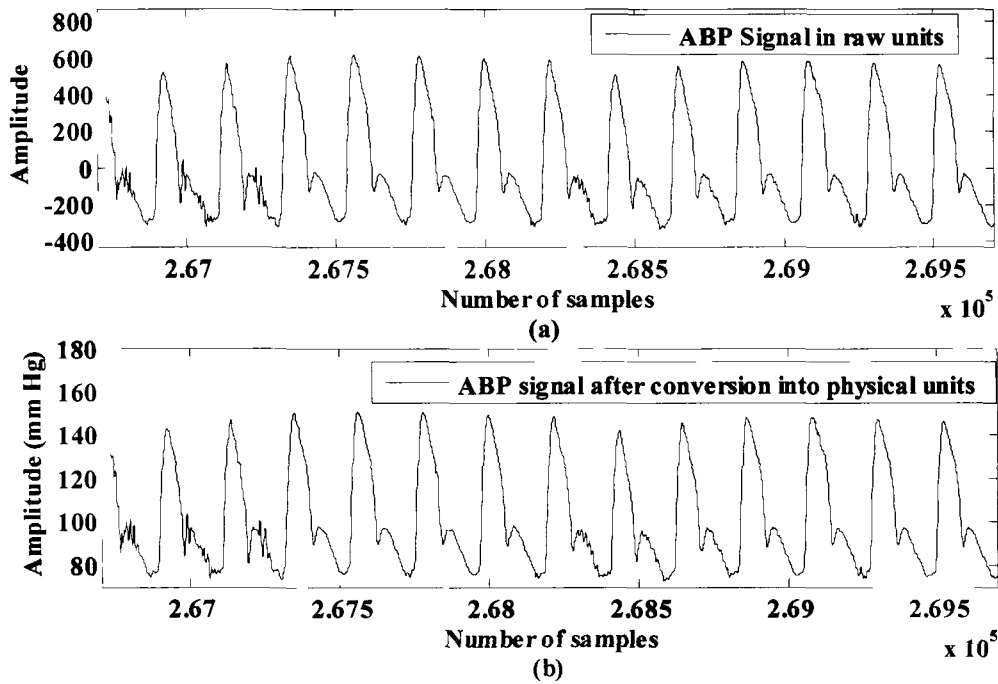
Figure 3.2 (a) ABP signal in raw units and (b) ABP signal after conversion into physical units

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques
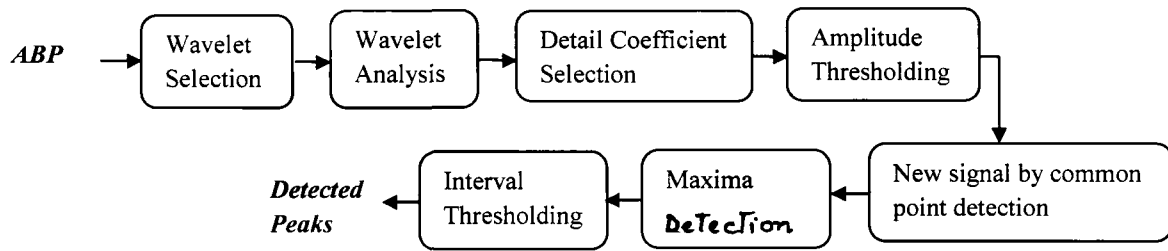
Figure 3.3 Block diagram of ABP peak detection algorithm using wavelet technique

3.1.1.1. Wavelet analysis

For wavelet analysis, the choice of wavelet depends upon type of application. Generally, a wavelet similar in shape to the signal being analyzed is considered suitable for the analysis [24,22]. Symlets (sym4) and Daubechies (db4) family of wavelets have been found to give details more accurately than others for ABP feature extraction in this research. Wavelet function (ψ) of sym4 and db4 wavelets are shown in Figure 3.5(a-b) respectively.

The ABP signals are decomposed to the required level using db4 and sym4 wavelets depending upon the dominant frequency components present in the signal. The detailed wavelet components (d1-d9) and approximation coefficient at first level is shown decomposition structure of ABP signal (mgh007 record) and approximation coefficient at first level is shown in Figure 3.6. From this decomposition structure, the relevant detail coefficient is selected based on the energy, frequency and cross-correlation analyses as mentioned below –

3.1.1.2. Selection of detail coefficients by energy analysis

Applying the same analysis as was done for peak detection of ECG signal, it is decided that maximum energy of an ABP signal is available in its higher amplitude and wider systolic complex. Other segments of the signal such as onset and dicrotic notch possess lower energy. It is observed that d7 possess highest energy for mgh007 while d6 has maximum energy for abp1 signal of CSL database as shown in Table 3.2. It is observed from Table 3.2 that the sum of energy of all detail coefficients and remaining one approximation coefficient is equal to the energy of ABP signal under test, which proves the energy

conservation principle of wavelet transform. It means that original signal can be faithfully reproduced from the decomposed components and the information in the original signal is distributed at different scales but remain preserved during decomposition.

Figure 3.6
Figure 3.7 & Table 3.3
Figure 3.8

Detect maxima in new signal
Figure 3.8

a. The detected maxima has higher amplitude than previous and succeeding sample
Figure 3.8

b. Two corresponding samples possess the same amplitudes but Ist & IVth have lower amplitudes

c. Three corresponding samples possess the same amplitude values but Ist & Vth samples have lower amplitudes
Figure 3.9

d. Four corresponding samples possess the same amplitudes but Ist & VIth samples have lower amplitudes

e. Go to next sample and repeat the process till maxima is detected

Yes No Yes No Yes No Yes No Yes

Valid Maxima Detection

Interval thresholding
Figure 3.10

Detected Peaks
Figure 3.11

Figure 3.4 Schematic representation of steps involved in ABP peak detection algorithm

(a)

(b)
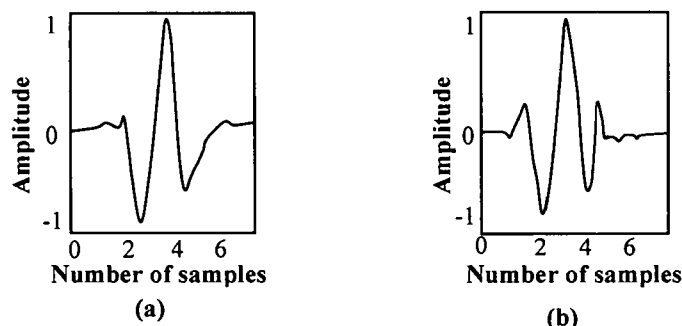
Figure 3.5 Wavelet function (ψ) of (a) symmetric (sym4) wavelet and (b) daubechies (db4) wavelet
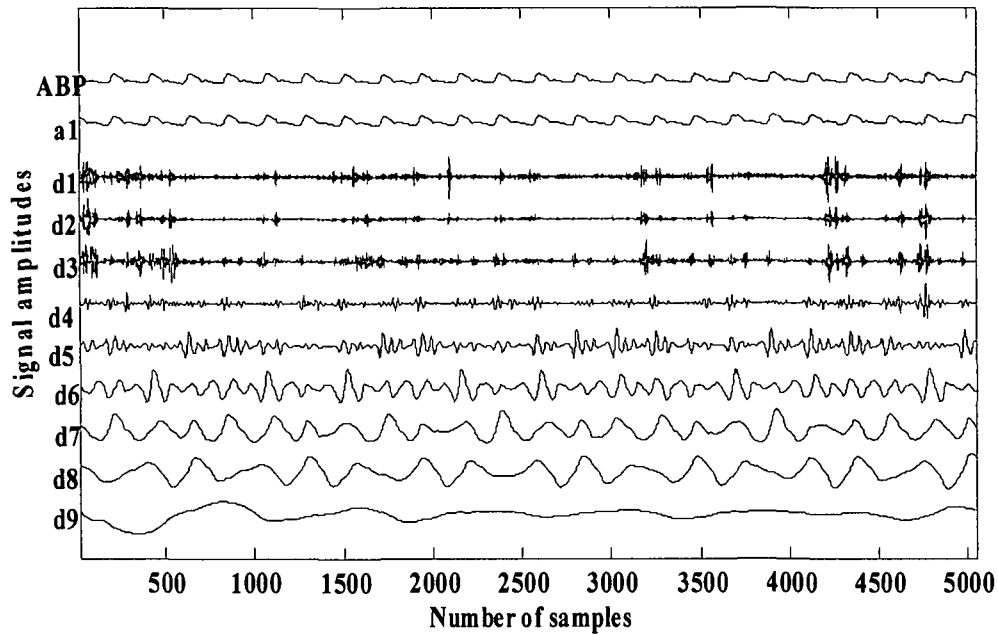
Figure 3.6 Decomposition of ABP signal of mgh007 record using db4 wavelet

Table 3.2 Energy content of detail and approximation coefficients

Signal	Detail Coefficients	Energy contents Using sym4	Energy contents Using db4
mgh007	d1	0.0000	0.0000
	d2	0.0001	0.0001
	d3	0.0025	0.0023
	d4	0.0405	0.0377
	d5	0.2521	0.2505
	d6	*1.0197*	*1.0019*
	d7	*2.5670*	*2.5178*
	d8	0.9724	0.9523
	d9	0.0212	0.0184
	a9	95.1246	95.2190
abp1	d1	0.0001	0.0001
	d2	0.0018	0.0018
	d3	0.0901	0.0870
	d4	0.9122	0.9260
	d5	*2.1670*	*2.1364*
	d6	*2.3076*	*2.3362*
	d7	0.1158	0.1026
	d8	0.1398	0.1604
	d9	0.1127	0.0984
	a9	94.1528	94.1512

3.1.1.3. Selection of detail coefficients by frequency analysis

Another justification of selecting d7 for mgh007 and d6 for abp1 is its available frequency components correlated with that of ABP signal. In pressure signals, most of the signal power is in the frequency range of 0.7-3.5 Hz in humans [44]. Therefore, the Fourier analysis of ABP signal, a1 signal and all its decomposed detail signals is computed as shown in Figure 3.7. From Figure 3.7, it is observed that d7 for mgh007 record and d6 for abp1 signal matches in frequency content of the original ABP signal. Frequency range of all the detail signals (d1-d9) obtained from FFT plot is shown in Table 3.3.

Table 3.3 Frequency content of detail coefficients

Signal	Detail coefficients	Frequency content (Hz) using sym4	Frequency content (Hz) using db4
mgh007	d1	40.99 - 180	39 - 180
	d2	27.5 - 126.9	25.41 - 126.7
	d3	10.85 - 69.68	10.62 - 68.05
	d4	5.573 - 34.92	5.583 - 35.63
	d5	2.123 - 17.86	2.17 - 18.02
	d6	*0.8001 - 9.127*	*0.8533 - 8.997*
	d7	*0.3533 - 4.567*	*0.45 - 4.423*
	d8	0.16 - 1.923	0.6033 - 1.833
	d9	0.09667 - 1.123	0.2267 - 1.107
abp1	d1	10.94 - 62.5	12.38 - 62.5
	d2	5.399 - 48.68	4.305 - 46.68
	d3	2.832 - 25.71	3.005 - 24.38
	d4	1.873 - 11.98	1.504 - 12.92
	d5	*0.6275 - 5.939*	*0.7567 - 5.861*
	d6	*0.3331 - 3.432*	*0.2411 - 3.403*
	d7	0.1514 - 1.581	0.1453 - 1.599
	d8	0.04889 - 0.8556	0.05306 - 0.8056
	d9	0.0125 - 0.4042	0.04389 - 0.3922
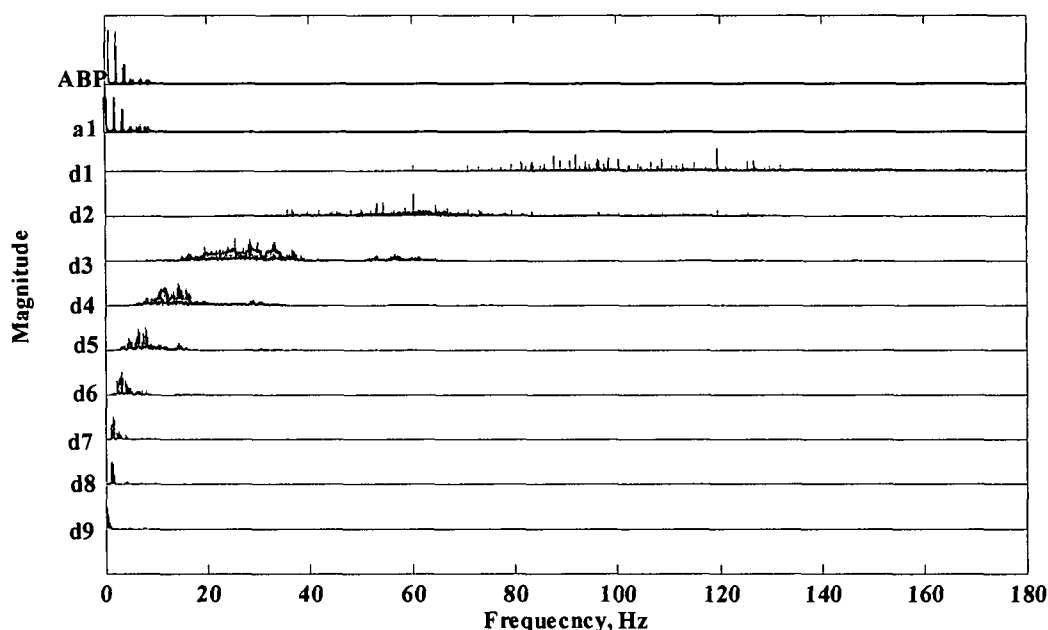
Figure 3.7 FFT of ABP signal from mgh007 record, approximation coefficient at first level (a1) and detail coefficients (d1-d9)

3.1.1.4. Selection of detail coefficient by cross-correlation analysis

In addition to above two analyses, cross-correlation analysis of detail coefficients individually with the original ABP signal is also performed. The results of correlation analysis of detail signals with ABP signals are given in Table 3.4. It is observed that the cross-correlation coefficient is highest at d7 for mgh007 record while it is in d6 for abp1 signal. This fact is observed while wavelet decomposition is followed using both sym4 and db4 wavelets in mgh007 and abp1 signals.

From Table 3.4, it is evident that d7 for mgh007 ABP signal and d6 of abp1 are highly correlated with the original ABP signals in time domain. The results from the above mentioned three analyses show that d7 for mgh007 and d6 for abp1 signal carries maximum information in regard to the ABP peaks signals in terms of energy content, frequency content and time domain similarity. It is clear from Table 3.3 that both d6 and d7 for mgh007 comprise of required frequency component of ABP signal. Similarly, d5 and d6 for abp1 record also encompass required frequency component. Hence, in this case, we propose to use two most nearest detail coefficient [d6, d7] and [d5, d6] pairs for mgh007

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

and abp1 records respectively which are also close to original signals in terms of energy content and cross-correlation coefficients.

Table 3.4 Cross-correlation coefficients

Signal	Detail coefficients	Cross-correlation coefficients using sym4	Cross-correlation coefficients using db4	Inference
mgh007	d1	0.0012	0.0011	Weak
	d2	0.0042	0.0043	Weak
	d3	0.0215	0.0211	Weak
	d4	0.0871	0.0851	Weak
	d5	0.2171	0.2185	Weak
	d6	0.4371	0.4370	Moderate low
	d7	*0.6930*	*0.6927*	*Moderate high*
	d8	0.4253	0.4256	Moderate low
	d9	0.0465	0.0466	Weak
abp1	d1	0.0039	0.0039	Weak
	d2	0.0157	0.0157	Weak
	d3	0.1106	0.1088	Weak
	d4	0.3522	0.3548	Moderate low
	d5	0.5427	0.5388	*Moderate high*
	d6	*0.5599*	*0.5633*	*Moderate high*
	d7	0.1249	0.1174	Weak
	d8	0.1361	0.1469	Weak
	d9	0.1198	0.1123	Weak

3.1.1.5. Window based thresholding and maxima detection

The selected detail coefficients undergo window based thresholding as mentioned in section 2.1.1.6. A threshold value 18% of maximum signal amplitude within the window is found suitable. Then, a technique of feature intersection is adopted for maxima detection. A new signal is generated by selecting the intersecting samples with positive amplitude values in both d6 and d7 after thresholding. Detection of common samples in d6 and d7 signals is shown in Figure 3.8. It is clear from Figure 3.8 that the amplitude of intersecting samples are equivalent of sample amplitude in original ABP signal.

In this newly generated signal, a maxima $d(k)$ is detected when the amplitude of previous and the succeeding sample are lower than the middle sample following condition (3.2) for maxima detection. It is possible that the newly formed signal may comprise of two or more maxima of same amplitude because of the fact that pressure signals are slow varying. So detection of maxima is difficult among three consecutive samples. In such cases, the

algorithm looks for the third sample or next sample if it possesses lower amplitude. This process is repeated till a valid maxima is detected. An example for valid maxima selection is shown in Figure 3.9. The rules for valid maxima $d(k)$ detection are given below –

a) If $d(k) > d(k-1)$ and $d(k) > d(k+1)$ (3.2)

b) If two corresponding samples have same amplitude i.e. $d(k) = d(k+1)$

Then valid maxima condition

$d(k) > d(k-1)$ and $d(k) > d(k+2)$ (3.3)

c) If three corresponding samples have same amplitude i.e. $d(k) = d(k+1)$ and $d(k) = d(k+2)$

Then valid maxima condition

$d(k) > d(k-1)$ and $d(k) > d(k+3)$ (3.4)

3.1.1.6. Interval thresholding and peak positions

The numbers and position of detected maxima denote the approximate number of systolic peaks and approximate systolic peaks positions. As already stated that peak to peak interval in case of ECG and ABP is same (Figure 1.10). Therefore, a physiological refractory period of 200 ms should exist before next ABP peak appears. Assuming this fact, an interval threshold of 200 ms is applied after detection of first peak in the signal that gives rise to actual number and positions of peaks [21, 160]. Application of interval thresholding is shown in Figure 3.10. It is clear from Figure 3.10 that the second maxima detected as peak is eliminated by the application of interval threshold. ABP signal with positioned systolic peaks for mgh007 record and 'abp1' signal are shown in Figure 3.11 (a-b) respectively.
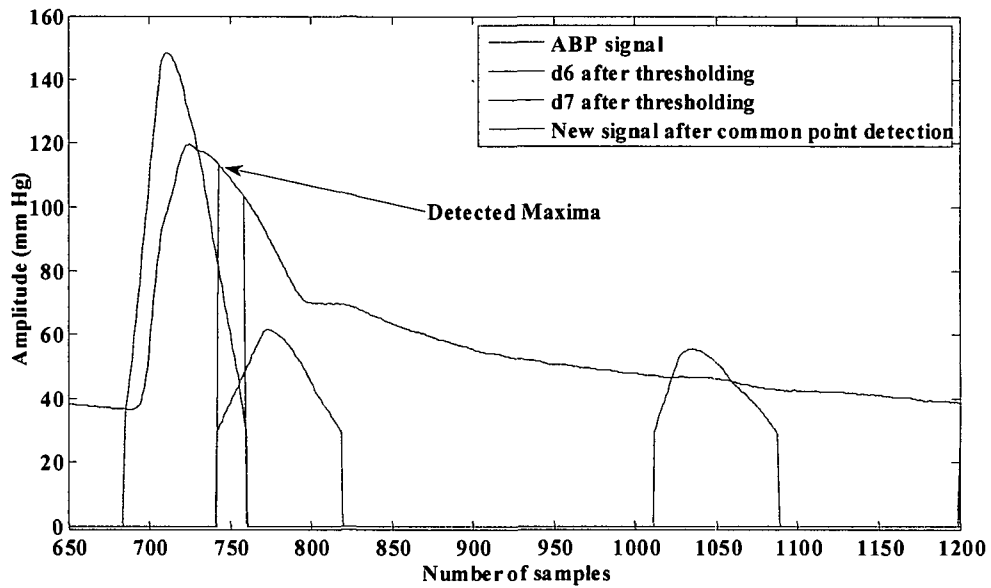
Figure 3.8. Detected maxima from the new signal after common points detection
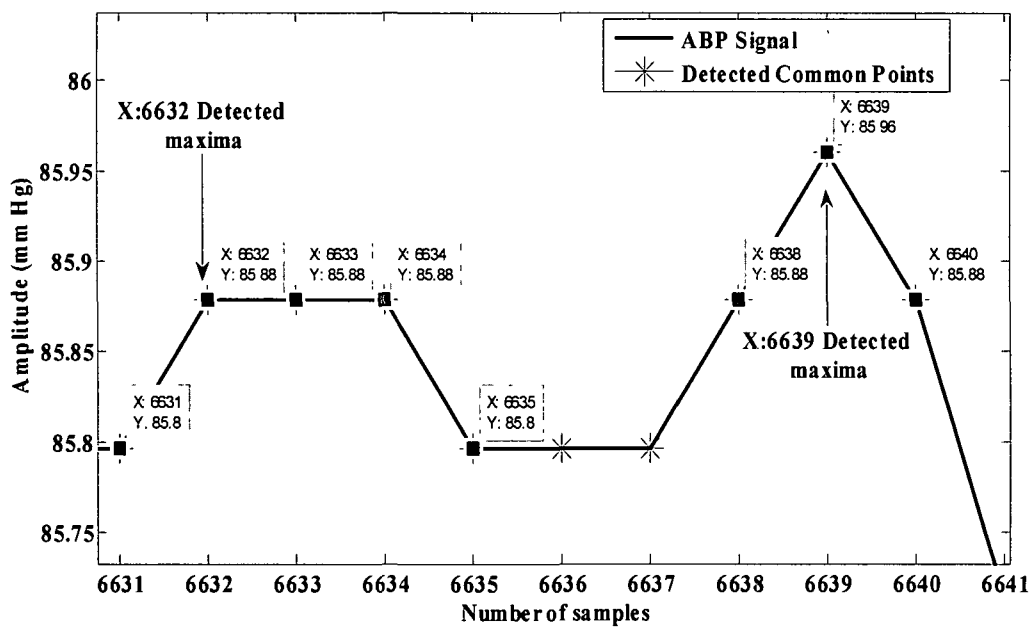
Figure 3.9 Detected maxima when the three corresponding samples possess the same amplitudes but first and fifth sample possess lower amplitudes
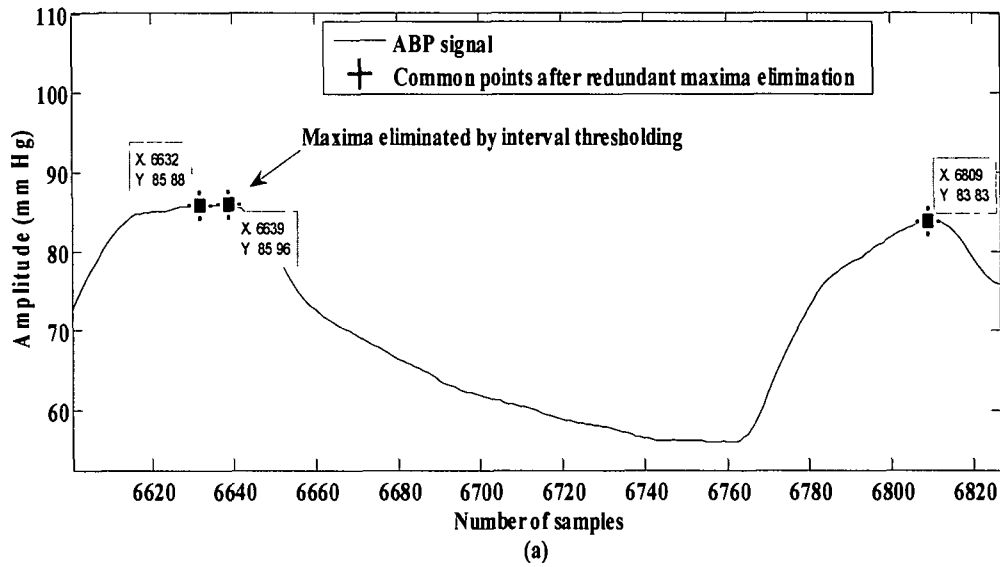
Figure 3.10 ABP signal along with common points after redundant maxima elimination
and application of interval thresholding (Refractory period)
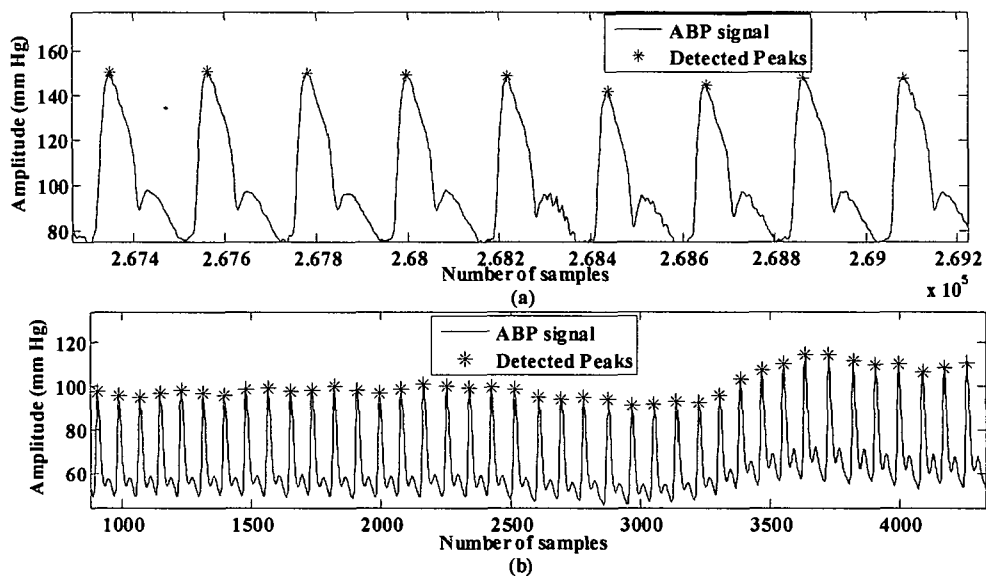
Figure 3.11 Detected peaks positions in ABP signal of (a) mgh007 record of MGH/MF
waveform database and (b) abp1 signal of CSL database

3.1.2. ABP peak detection using energy analysis

As already stated in section 2.1.2 that window based energy analysis of ECG signal has been efficiently used for the detection of R peaks in ECG signal. Like R wave in ECG signal, systolic peak in ECG signal is also the highest amplitude peak. Following this fact, window based energy analysis of ABP signal is computed using a window of 100 ms duration. This energy domain offers an easy interpretation of the ABP signal for the detection of systolic peaks in ABP signal. The block diagram of ABP peak detection using energy analysis is shown in Figure 3.12. The detection process is performed on ABP signal of mgh001 record as discussed below –

3.1.2.1. Preprocessing

Variations in signal amplitude may result in the erroneous calculation of energy as energy calculation includes the absolute value of signal amplitude. The biomedical signals are corrupted by artifacts such as base line drift, power-line interference etc. Base line drift causes sharp variation in signal amplitude. That is why removal of base line drift is the key

ABP Signal

```
┌──────────────────┐
│  Preprocessing   │
└──────────────────┘
         │
┌──────────────────┐
│     Energy       │
│    Analysis      │
└──────────────────┘
         │
┌──────────────────────┐
│   Window based       │
│ amplitude thresholding│
└──────────────────────┘
         │
┌──────────────────┐
│    Interval      │
│  thresholding    │
└──────────────────┘
         │
```
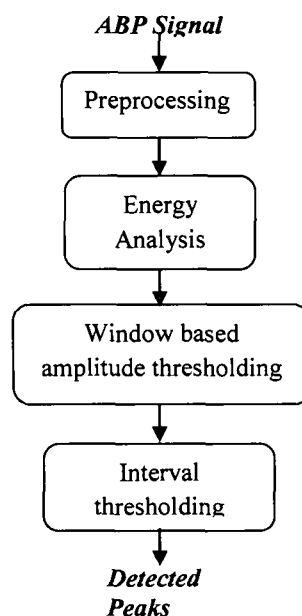
Detected
Peaks

Figure 3.12 Block diagram of ABP peak detection algorithm using energy analysis

step before starting the detection process by this method. Blood pressure signals have frequency range from 0.7-3.5 Hz [44]. It is further noticed from FFT of ABP signal shown

in Figure 3.13 that the frequency of baseline drift is 0.0018 Hz. The signal under test is filtered by fourth order Butterworth highpass filter with cutoff frequency of 0.5 Hz to eliminate base line drift [147].
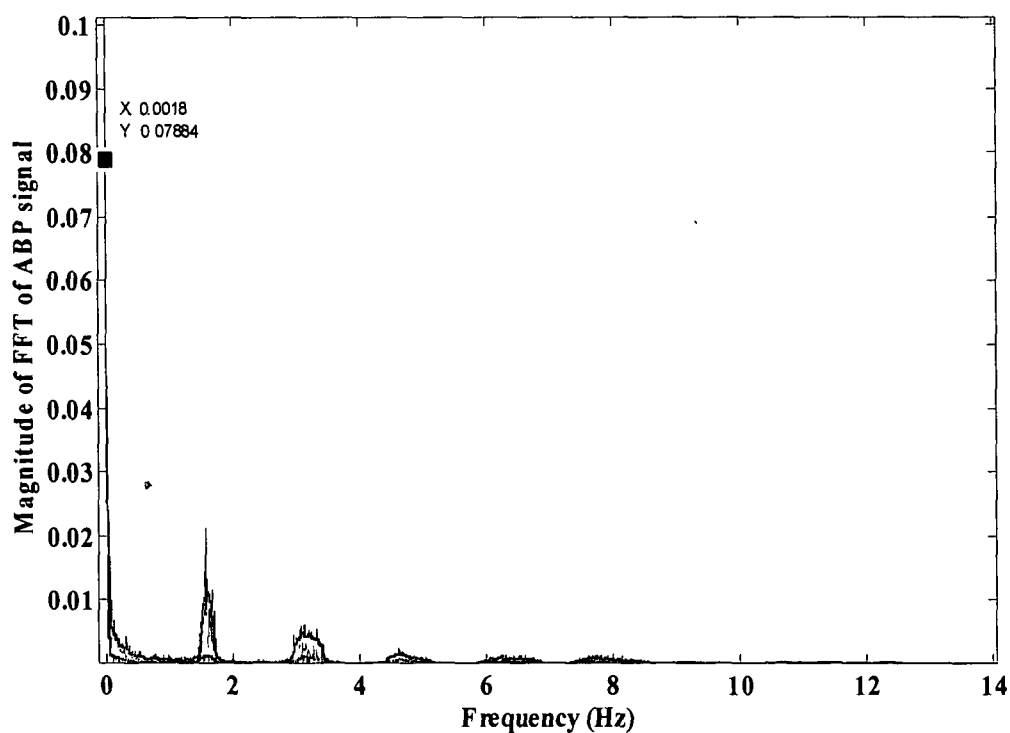
Figure 3.13 FFT of ABP signal

3.1.2.2. Energy analysis by moving window

The energy of filtered ABP signal is calculated considering a window of 100 ms. then undergoes energy analysis by windowing approach considering a window of 100 ms. To match the window to the signal length, zero padding is done in both start and end of the signal equally. The length (L) of the new signal is given by equation 2.10 and number of samples are distributed in each window according to equation 2.11. Energy distribution of the ABP signal using a window of 100 ms duration is shown in Figure 3.14(a). It is clear from Figure 3.14(a) that energy is higher at systolic peaks locations.

3.1.2.3. Thresholding

A threshold value is applied to the generated energy signal as discussed in The generated energy signal now undergoes window based thresholding as discussed in section 2.1.1.6. A threshold value of 5 % of the maximum amplitude of the signal within the predefined window of 2 sec duration is found suitable for the detection of peaks by this method. The ABP and thresholded energy signals are shown in Figure 3.14(b).

3.1.2.4. Interval thresholding and peak positions

The onset or offset samples of remaining thresholded energy windows provide an idea of approximate number of peaks available in the signal and their positions as systolic peaks. An interval threshold of 200 ms duration is applied after detection of first peak in ABP signal under test [21,160]. ABP waveform with the positioned systolic peaks is shown in Figure 3.15(a) and corresponding ECG lead II signal with approved beat annotations by expert is shown in Figure 3.15(b).
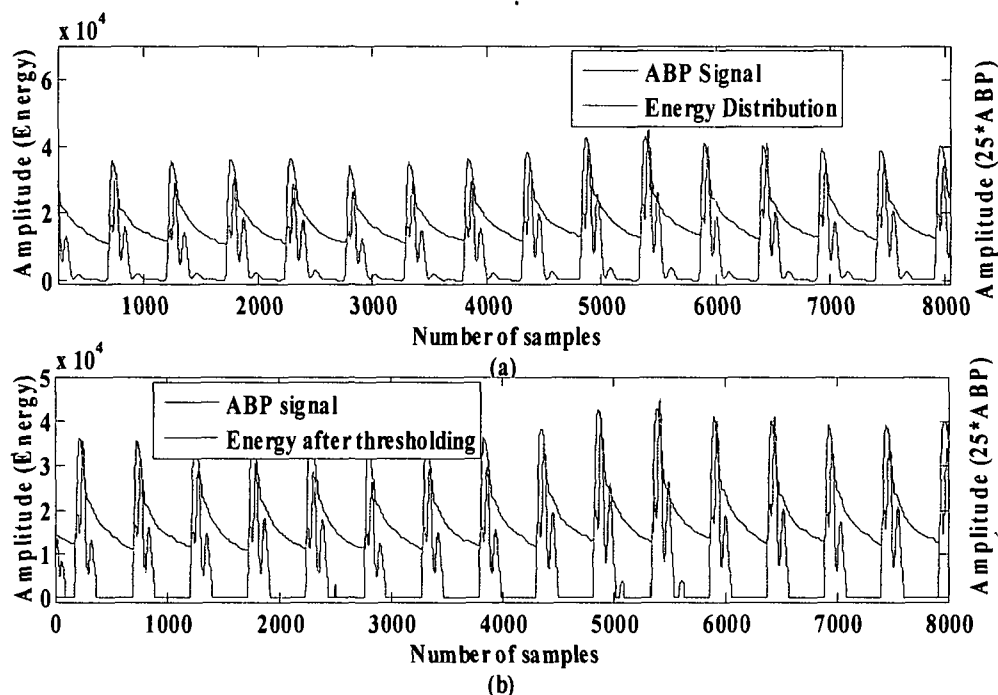
Figure 3.14 (a) Energy distribution of ABP signal of mgh001 record and
(b) Energy after thresholding
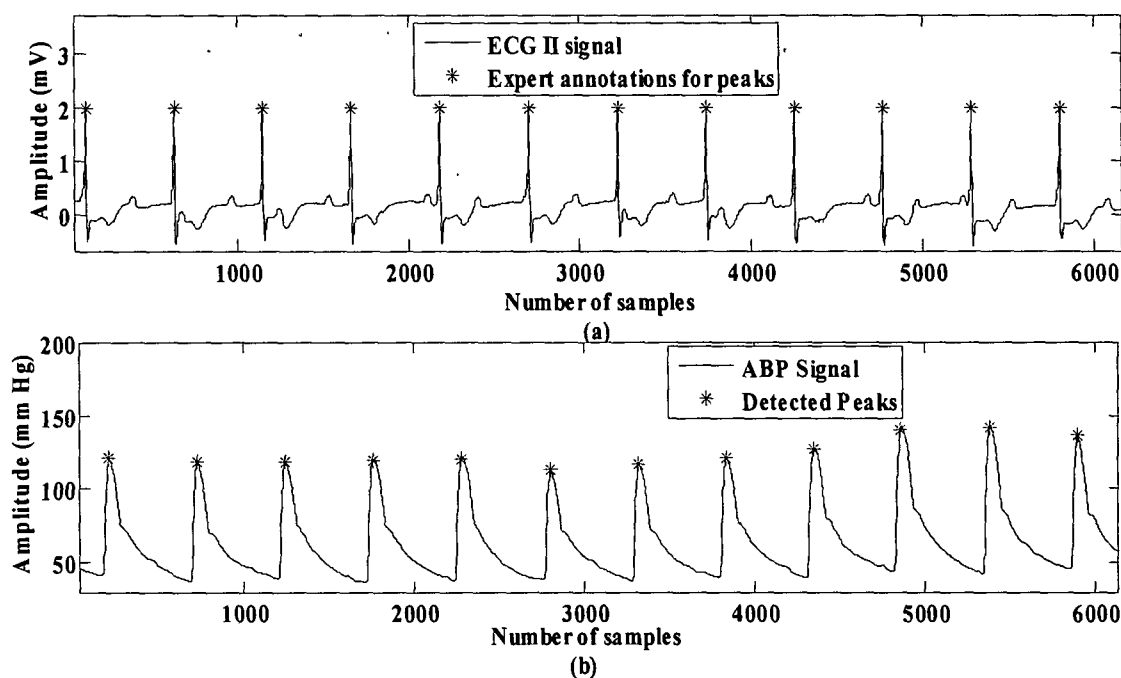
Figure 3.15 (a) ECG lead II signal with expert annotations for peaks and (b) ABP waveform with positioned systolic peaks for mgh001 record

3.1.3. ABP onset, dicrotic notch and dicrotic peak detection

After the detection of peaks of ABP waveform, the remaining features such as onset, dicrotic notch and dicrotic peak are detected using approximation signal at first level ($'a_l'$) The detected peaks positions serve as the reference for detection of remaining features of signal. As discussed earlier in section 2.1.1.2, wavelet decomposition of the signal using the selected wavelet at first level results in detail coefficient and average coefficient. The approximation signal (a_l) obtained after wavelet analysis is shown in Figure 3.6 along with detail coefficients. This approximation coefficient comprises of all features of ABP signal and the shape of ABP signal is also retained. The analysis of (a_l) has been found suitable for extracting the remaining features of ABP signal. Flow chart of ABP onset, dicrotic notch and dicrotic peak detection algorithm are shown in Figure 3.16. The onset positions of ABP waveform are determined after eliminating the redundant minima obtained from approximation signal by thresholding. For this purpose, detected peaks positions in the ABP signal are taken into account. Amplitude thresholding is applied to a_l signal. After thresholding, minima in the signal are detected if the previous and subsequent samples in

the signal have higher amplitude than the sample amplitude between them. The resulting signal with minima positions undergoes window based amplitude thresholding. The minima below the threshold are kept and remaining minima are discarded.

i)   Onset detection

Out of all these minima, only those minima are taken into account which appear between the detected peaks. If several minima are detected between two peaks, minima with the minimum amplitude is registered as ABP onset. The detected peaks and onset positions in the ABP signal are now used to detect the dicrotic notch in ABP signal. The method of onset detection is shown in Figure 3.16 and Figure 3.17.

ii)   Dicrotic notch detection

The dicrotic notch is found between the peak of first ABP cycle and onset of next ABP cycle. For this purpose, the algorithm looks for the minima between detected peaks and onset locations. If there are several minima found between a certain peak and onset, the algorithm takes into account the minima with maximum amplitude and register the minima as dicrotic notch.

iii)   Dicrotic peak detection

Dicrotic peak appears as a peak with weak amplitude between the dicrotic notch and onset of next ABP pulse. It is the result of reflected waves from the lower extremities and the aorta. For the detection of dicrotic peak, the algorithm takes into account the detected onset and dicrotic notch positions. All the maxima in the average signal at first level are determined. The maxima with maximum amplitude appearing between dicrotic notch and onset of next ABP pulse are registered as dicrotic peaks. The method of dicrotic peak detection is shown in Figure 3.16 and Figure 3.18.

The detected four signal components (peak, onset, dicrotic notch and dicrotic peak) for ABP signals of MGH/MF waveform database, Fantasia database, MIT-BIH polysmographic database are shown in Figures (3.19–3.23) respectively. ABP peak detection is proposed by author in [44] on abp1 signal of CSL database.

ABP Signal

Perform wavelet decomposition using selected wavelet up to 9 levels

Approximation coefficient at level 1

Maxima detection
if $a_1(k) > a_1(k-1)$ & $a_1(k) > a_1(k+1)$

Minima detection
if $a_1(k) < a_1(k-1)$ & $a_1(k) < a_1(k+1)$

Window based thresholding

Select the minima those appear between two detected peaks

Select the minima those appear between detected peaks and next onset positions

Are several minima detected — No — Is single minima detected

Select the minima that appears after the peak

Yes

Yes

Select the minima with minimum amplitude — *Register minima as onset*

Register minima as Dicrotic Notch

Select the maxima those appear after the dicrotic notch and before onset

Is single maxima detected — No — Are several maxima detected

Yes

Yes

Register minima as Dicrotic Peak — Select the maxima with maximum amplitude

Figure 3.16 Flow chart of ABP onset, dicrotic notch and dicrotic peak detection algorithm

Annotations for peaks from author [44] and expert are available for abp1 signal. Therefore, we have compared the detected with the annotations of author and expert1. Figure 3.24 shows author and expert1 annotations for peaks along with detector annotations for the four signal components. It is clear from Figure that annotations by our method are the same as that for author and expert1 annotations.
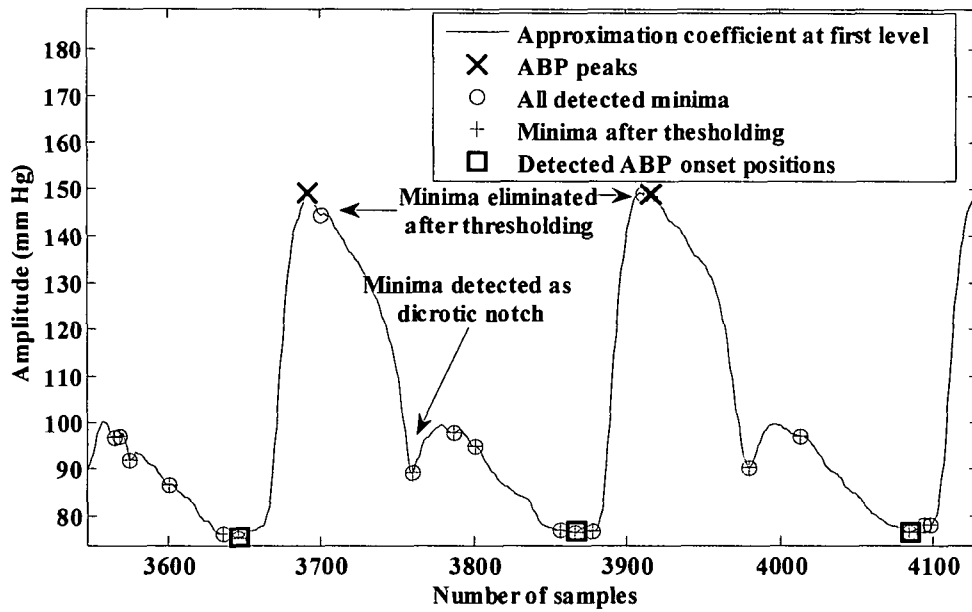
Figure 3.17 Detection of ABP onset and dicrotic notch in approximation coefficient at first level
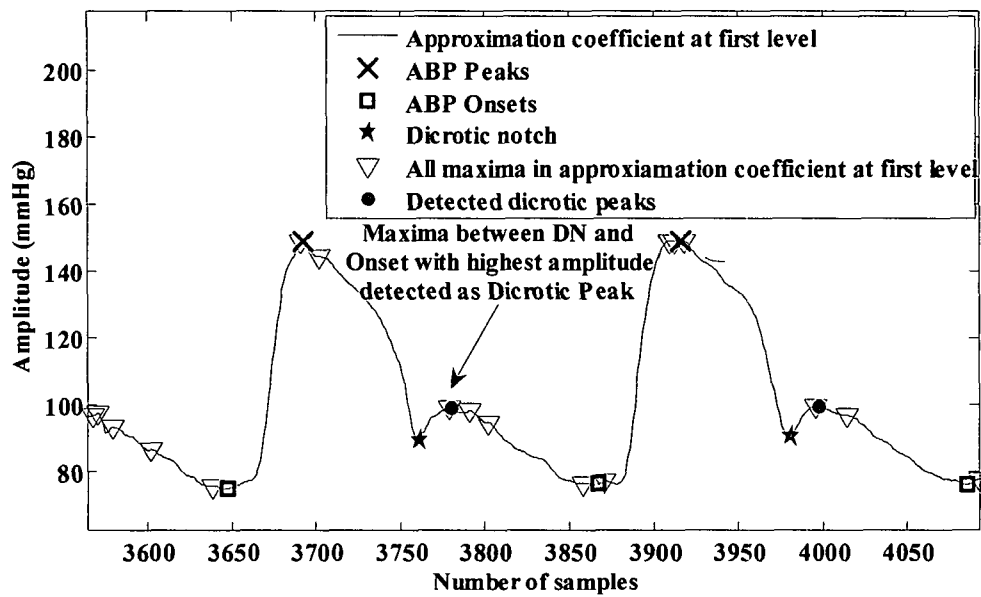
Figure 3.18 Detection of dicrotic peak in approximation coefficient at first level
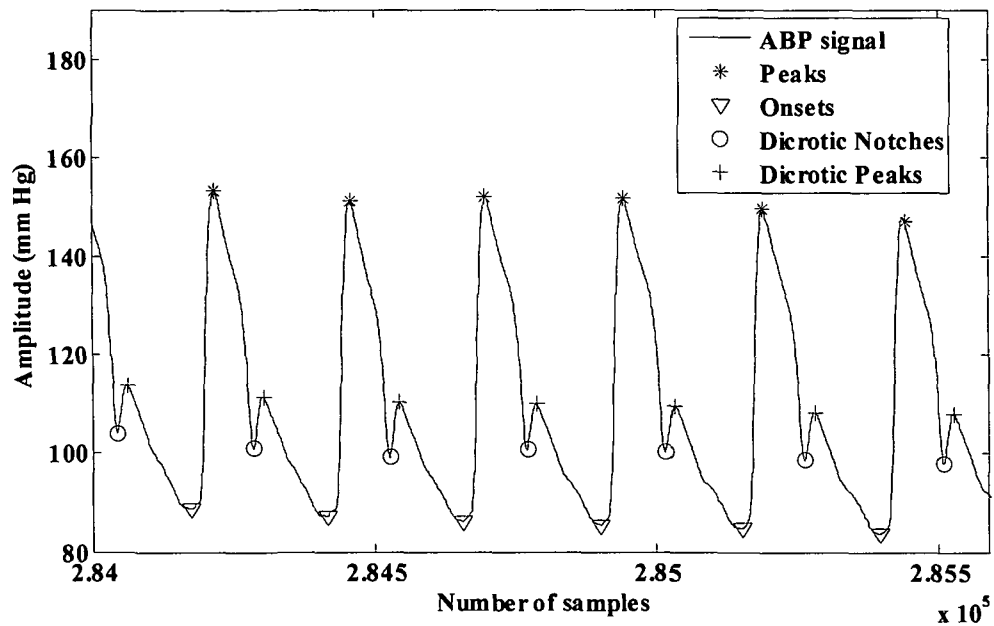
Figure 3.19 Detected peaks, onsets, dicrotic notch and dicrotic peaks in ABP signal of mgh007 record of MGH/MF waveform database using db4 wavelet
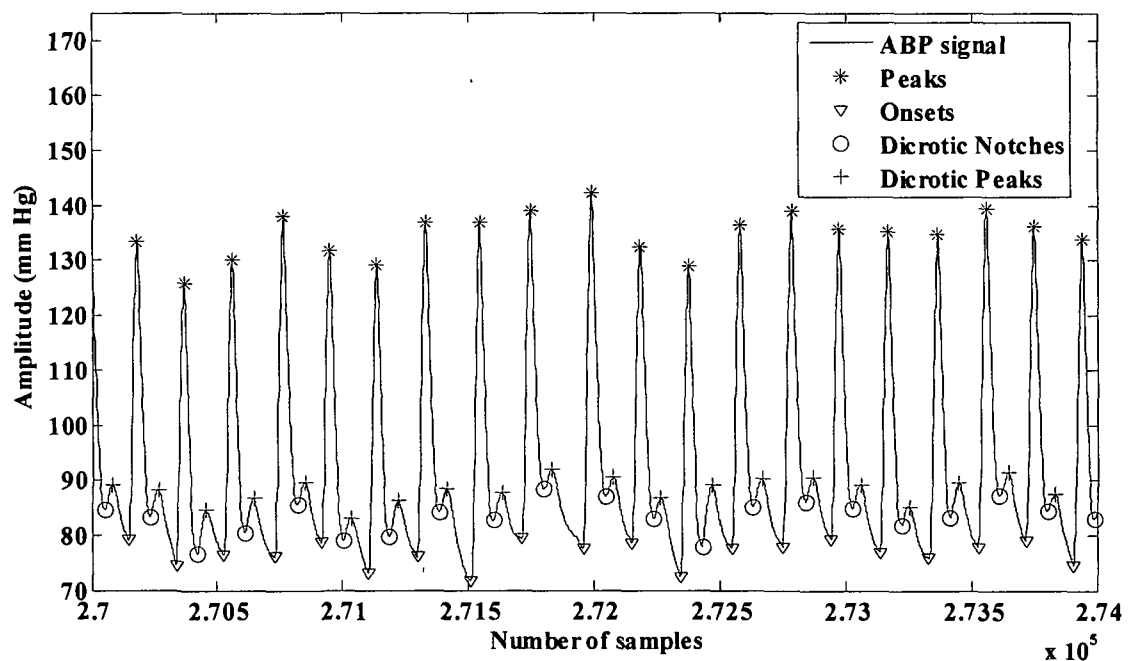
Figure 3.20 Detected peaks, onsets, dicrotic notch and dicrotic peaks in ABP signal of mgh010 record of MGH/MF waveform database using sym4 wavelet
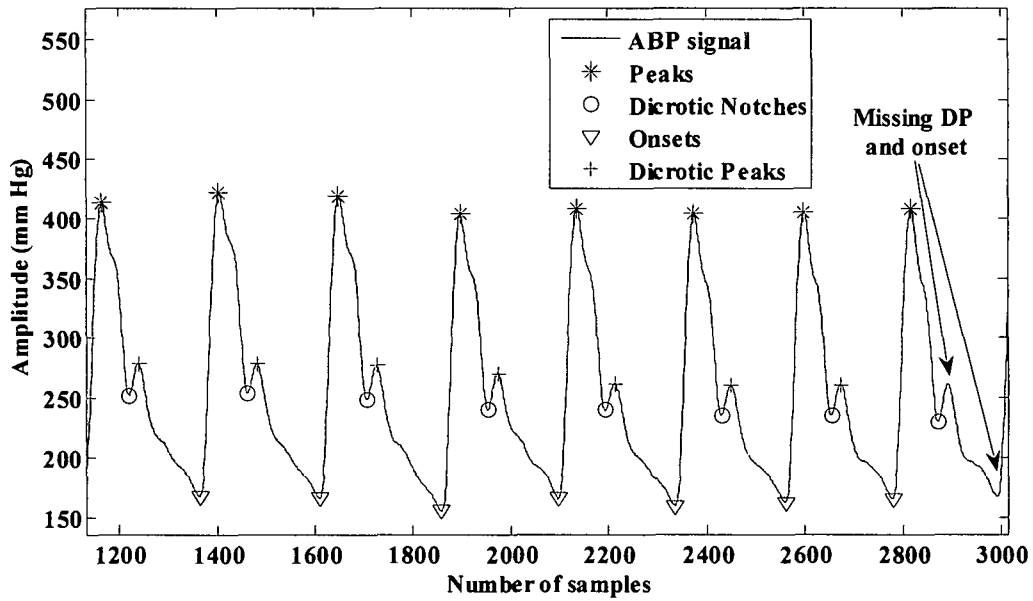
Figure 3.21 Detected peaks, onsets, dicrotic notch and dicrotic peaks in ABP signal of f2o04 record of Fantasia database using db4 wavelet. One DP and onset is missed by the algorithm
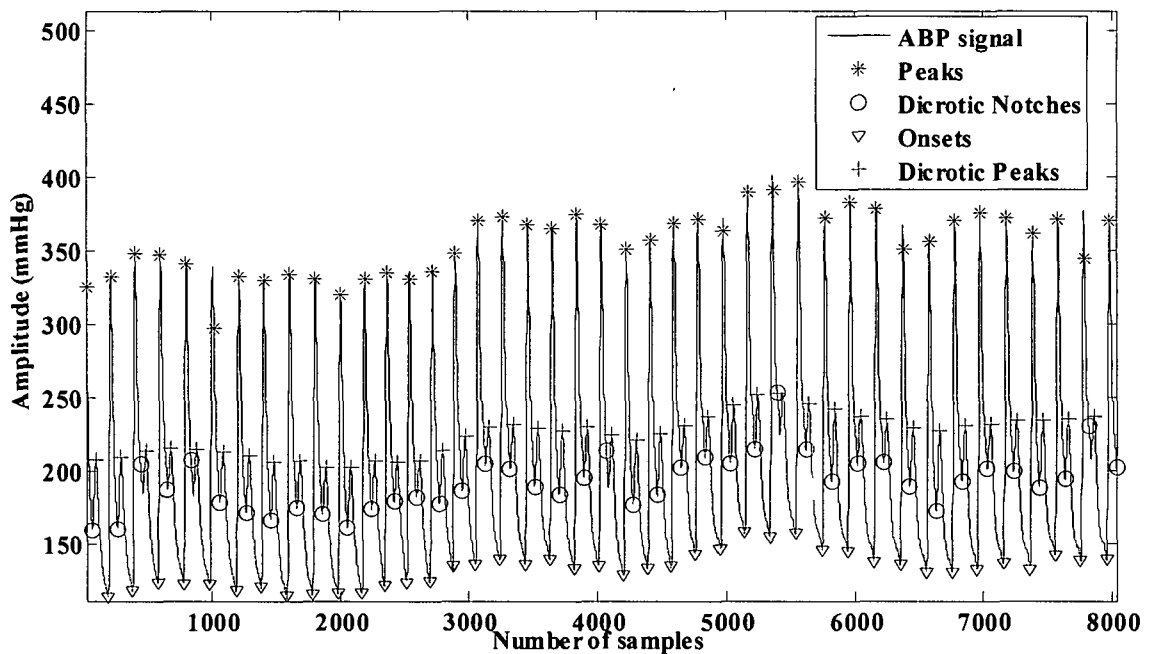
Figure 3.22 Detected peaks, onsets, dicrotic notch and dicrotic peaks in ABP signal of f2y05 record of Fantasia database
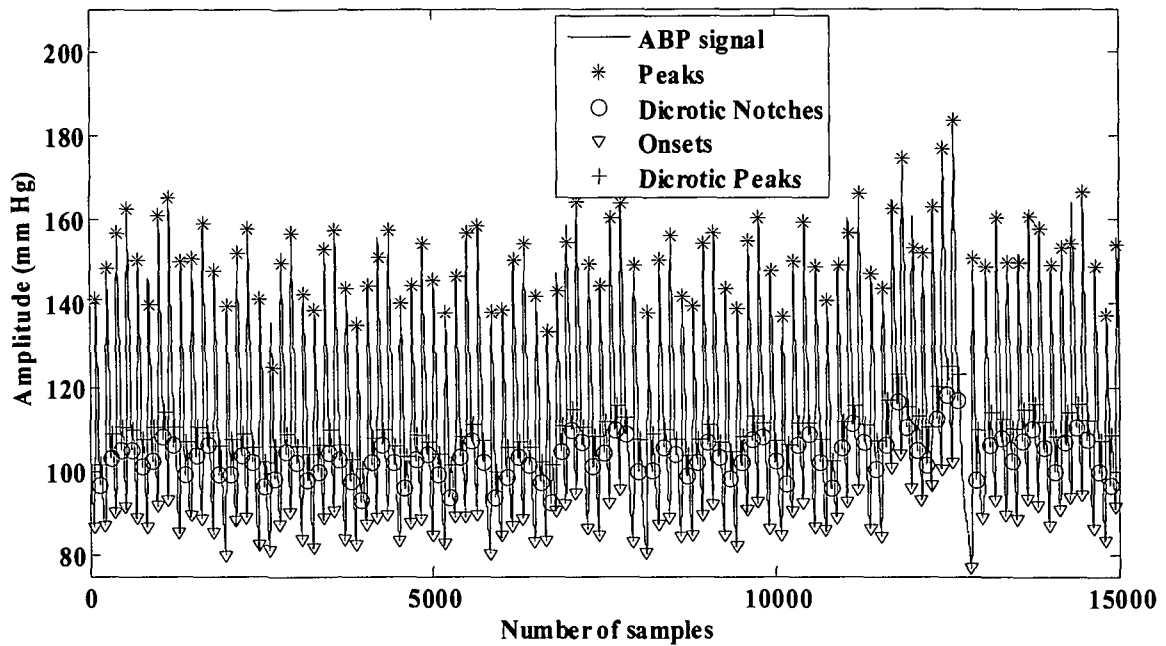
Figure 3.23 Detected peaks, onsets, dicrotic notch and dicrotic peaks in ABP signal (one minute segment) of slp02am record of MIT BIH polysmographic database using sym4 wavelet
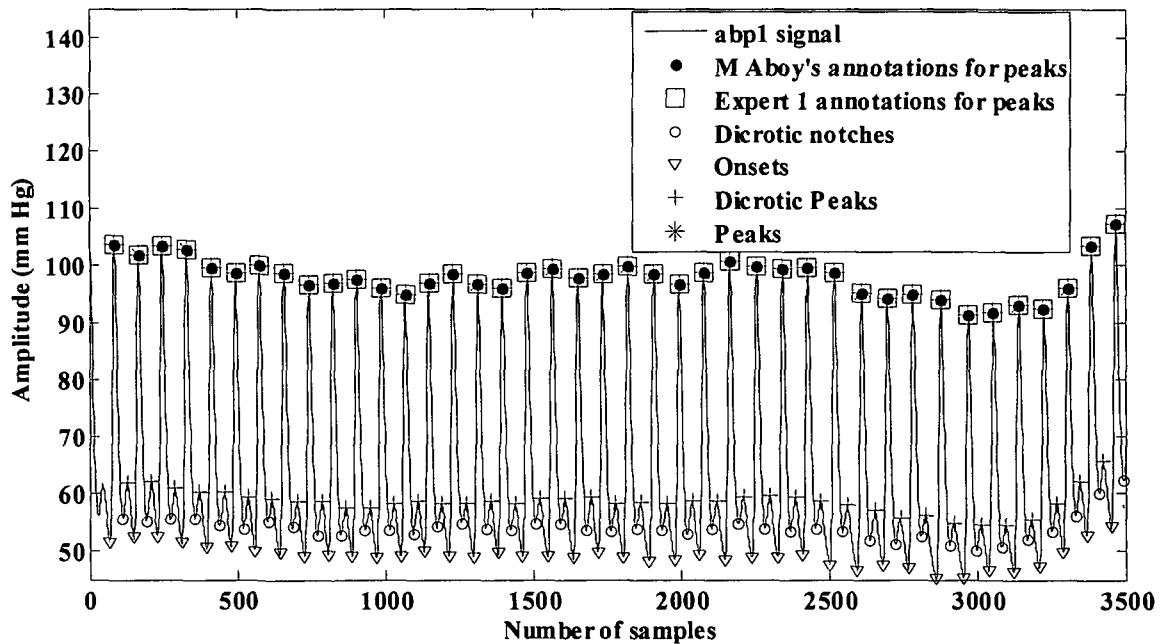
Figure 3.24 Detected peaks, onsets, dicrotic notch and dicrotic peaks in abp1 signal of CSL database

3.2. CVP feature extraction

Central Venous Pressure (CVP) is an important physiological parameter, and a clinically relevant diagnostic tool for heart failure patients. As already stated in section 3.0 and Figure 1.9, 1.10 that CVP waveform is five phasic with three upward (*'a'*, *'c'* and *'v'*) waves and two downward waves (*'x'*, *'y'*). For feature extraction of any signal, it is essential that a prominent component is identified (such as peaks in ECG and ABP signals) so that time locations of other signal components are spot out with respect to the prominent feature. Although *'a'* wave is regarded as prominent wave in CVP waveform but there is very small difference in the amplitudes of all three upward waves and consequently make it difficult to identify by thresholding methods. On the other hand, among the downward components, *'x'* descent has least amplitude. Therefore, *'x'* descent is chosen as the reference to extract other features such as *'a'*, *'c'*, *'v'* and *'y'* waves. Detection of *'x'* descent is carried out by wavelet technique. Among the available wavelet families daubechies db4 wavelet is found suitable for feature extraction of CVP signal. Block diagram and flow chart of *'x'* descent detection algorithm are given in Figure 3.25 and Figure 3.26 respectively. The flow chart for detection of *'c'*, *'v'*, *'a'* and *'y'* wave detection is shown in Figure 3.27. Detection of five phasic components in CVP waveform is carried out using wavelet technique as mentioned below –
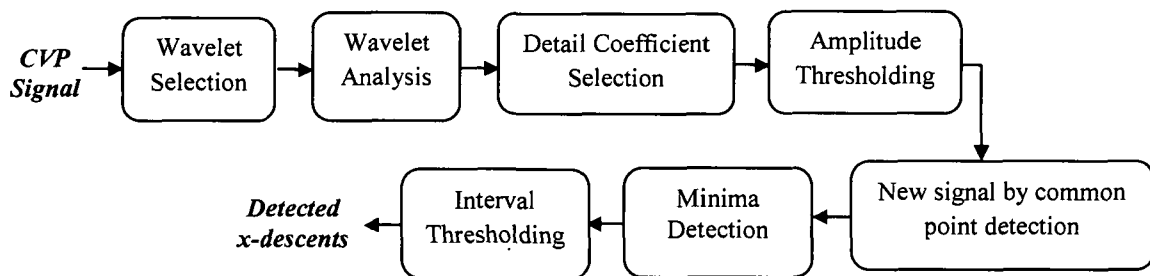
Figure 3.25 Block diagram of 'x' descent detection in CVP waveform

3.2.1. Detection of 'x' descent

i. Wavelet decomposition of the signal using 'db4' wavelet up to 9 levels (Figure 3.28 and 3.29).

ii. Selection of detail coefficient based on energy (Table 3.5), frequency (Figure 3.30, Table 3.6) and cross-correlation analysis (Table 3.7).

iii. Application of threshold value of 10% of minimum amplitude of the samples in the window. The samples in d6 and d7 with amplitude higher than threshold are discarded as shown in Figure 3.31.

iv. To select common samples with negative amplitude in selected detail coefficients to generate a new signal with the common samples and detection of the minima (Figure 3.32).

v. Application of a refractory period of 200 ms to eliminate the redundant minima to detect actual 'x' descent positions (Figure 3.33).

3.2.2. Detection of 'c' and 'v' waves

i. Selection of approximation coefficients at level 1 ('a1') and level 2 ('a2') (Figure 3.29).

ii. Application of threshold value of 5% of maximum amplitude of the signals 'a1' and 'a2' to remove lower amplitude values and generation of new signal by selecting intersecting samples.

iii. Detection of maxima in new signal.

iv. Application of an interval threshold of 20 samples to eliminate the redundant maxima and to select the maxima with higher amplitude.

v. Selection of the maxima those occur at an interval of 80 to 150 ms around 'x' descent positions. If several maxima occur within this interval, the maxima at the maximum distance around 'x' descent are selected (Figure 3.34). As seen in Table 1.6 that 'c' wave represents end of QRS complex. The onset and offset of T wave in ECG correspond to 'x' descent and 'v' wave in CVP. Therefore, it can be concluded that the segment 'c-x' is equivalent to ST segment of ECG which is generally 80 -120 ms duration. The duration of T wave is generally 160 ms. Assuming this fact, an interval threshold of 80 – 150 ms is applied before and after detection of 'x' descent position.

vi. The selected maxima before 'x' descent position is registered and marked as 'c' wave, while the maxima after 'x' descent position is registered and marked as 'v' wave (Figure 3.34).

CVP Signal

Perform wavelet decomposition using selected wavelet up to 9 levels
(Figure 3.28)

| Perform energy analysis on all the detail signals **(Table 3.5)** | Perform frequency analysis (FFT) on all the detail signals and original CVP signal **(Figure 3.30 &Table 3.6)** | Perform correlation analysis on all the detail signals **(Table 3.7)** |

| Select the detail signal with maximum energy and detail signal with next lower energy value | Select the detail signal in the frequency range of ABP signal and detail signal at next lower scale | Select the detail signal with highest value of cross-correlation coeff. and detail signal with next lower value of correlation coeff. |

Selected detail coefficients

Thresholding **(Figure 3.31)**

Select common points lower than zero in both the selected detail signals and generate a new signal **(Figure 3.32)**

Detect minima in new signal **(Figure 3.33)**

a. The detected minima has lower amplitude than previous and succeeding sample

b. Two corresponding samples possess the same amplitudes but Ist & IVth have higher amplitudes

c. Three corresponding samples possess the same amplitude values but Ist & Vth samples have higher amplitudes

d. Four corresponding samples possess the same amplitudes but Ist & VIth samples have higher amplitudes

e. Go to next sample and repeat the process till maxima is detected

Yes No   Yes No   Yes No   Yes No   Yes

Valid minima detection

Interval thresholding **(Figure 3.33)**

Detected 'x' descent
(Figure 3.33)

Figure 3.26 Flow chart of 'x' descent detection in CVP signal with relevant figures and tables marked
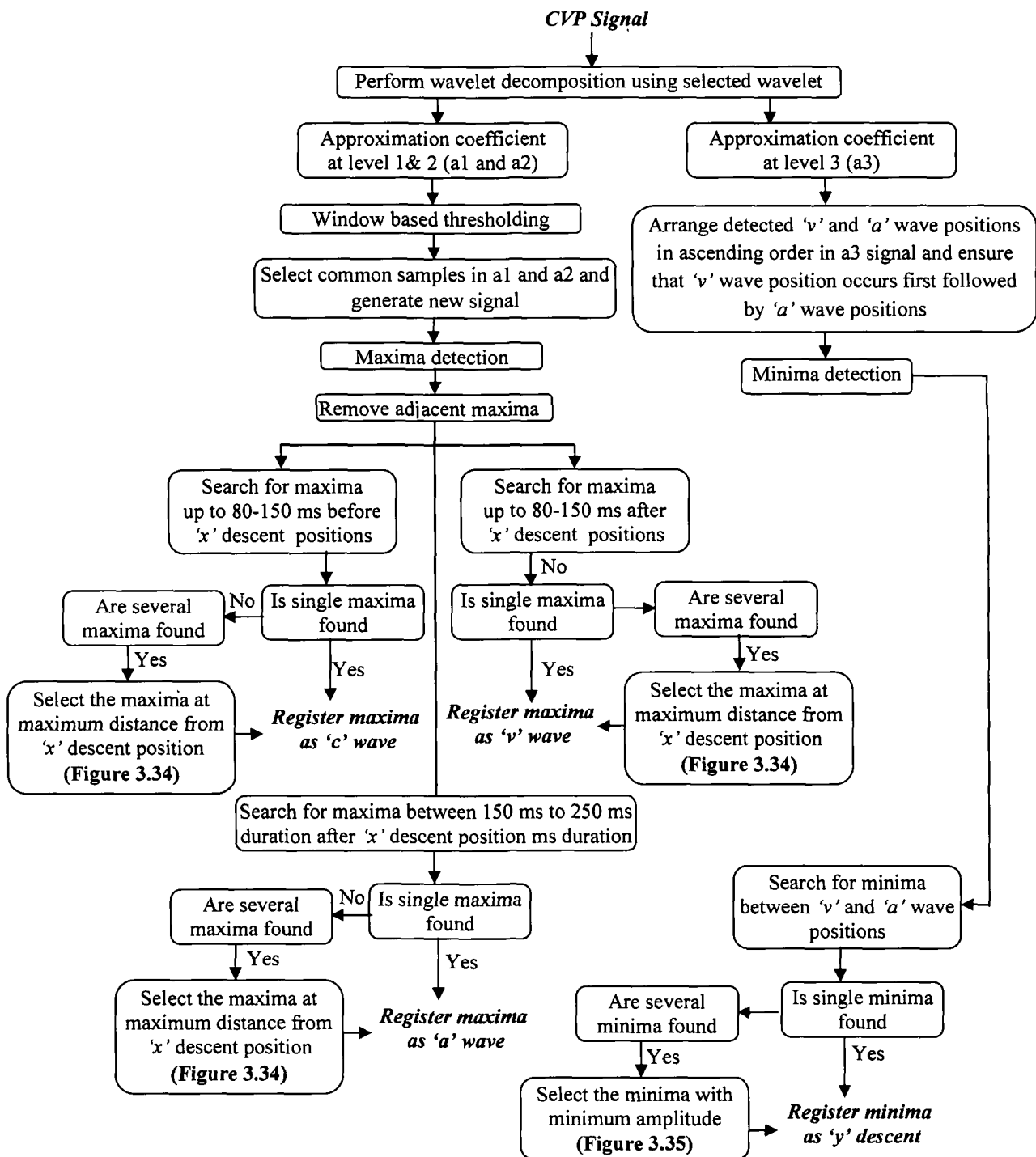
Figure 3.27 Flow chart of 'c', 'v', 'a' and 'y' wave detection in CVP signal with relevant figures and
tables marked
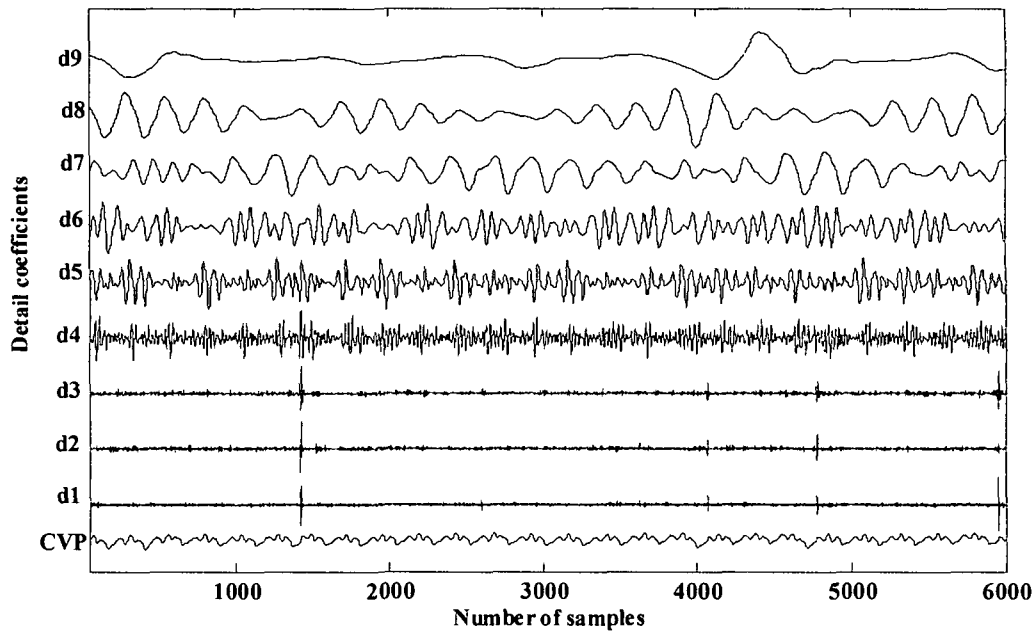
Figure 3.28 Decomposition of CVP signal of mgh007 record using db4 wavelet
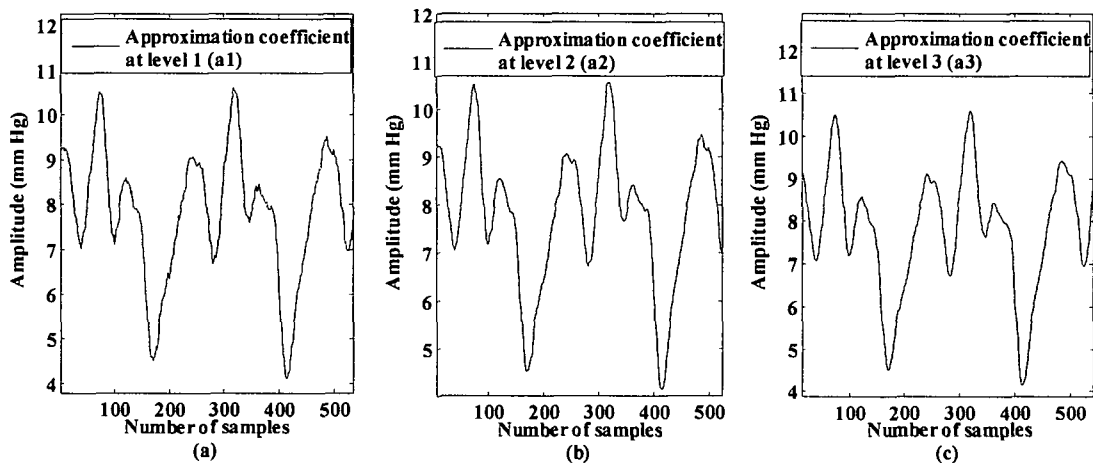
Figure 3.29 Plot of approximation coefficients at level 1, 2 and 3 obtained after decomposition of CVP signal of mgh007 record by db4 wavelet

3.2.3. Detection of 'a' waves

i.   To combine all maxima detected in generated new signal after eliminating 'c' and 'v' wave positions.

ii.   To arrange the maxima retained in (i) and 'x' descent positions.

iii.   To search for maxima within the interval of 150 ms to 250 ms after x-descent positions. If several maxima exist, maxima at the maximum distance from x-descent position is registered and marked as 'a' wave (Figure 3.34). The interval threshold of 150 -250 ms is selected because 'x' descent and 'v' wave represent onset and offset of T wave in ECG which has a duration of 160 ms. So 'a' wave would appear atleast after 160 ms interval of 'x' descent.

3.2.4. Detection of 'y' descent

i.   To arrange 'a' and 'v' waves positions in ascending order and ensure that 'v' wave position occurs first followed by 'a' wave position.

ii.   To detect minima in 'a3' and select the minima between 'v' and 'a' wave positions. If several minima exist, the minima with minimum amplitude is registered and marked as 'y' descent (Figure 3.35). Detection of 'y' descent can also be done in 'a1' and 'a2' but 'a3' is found most suitable because minima detection in 'a3' gives less number of redundant minima and consequently reduces computational complexity.

On the application of above mentioned steps, detected all components of CVP waveform are marked on the original signal and shown in Figure 3.36 and Figure 3.37 respectively.

Table 3.5 Energy content of detail coefficients and remaining approximation coefficient

Signal	Energy content (%)
d1	0.0006
d2	0.0023
d3	0.0023
d4	0.0286
d5	0.3503
d6	*0.5018*
d7	*0.8289*
d8	0.3897
d9	0.0448
a9	97.8507
Total energy content (%)	**100**
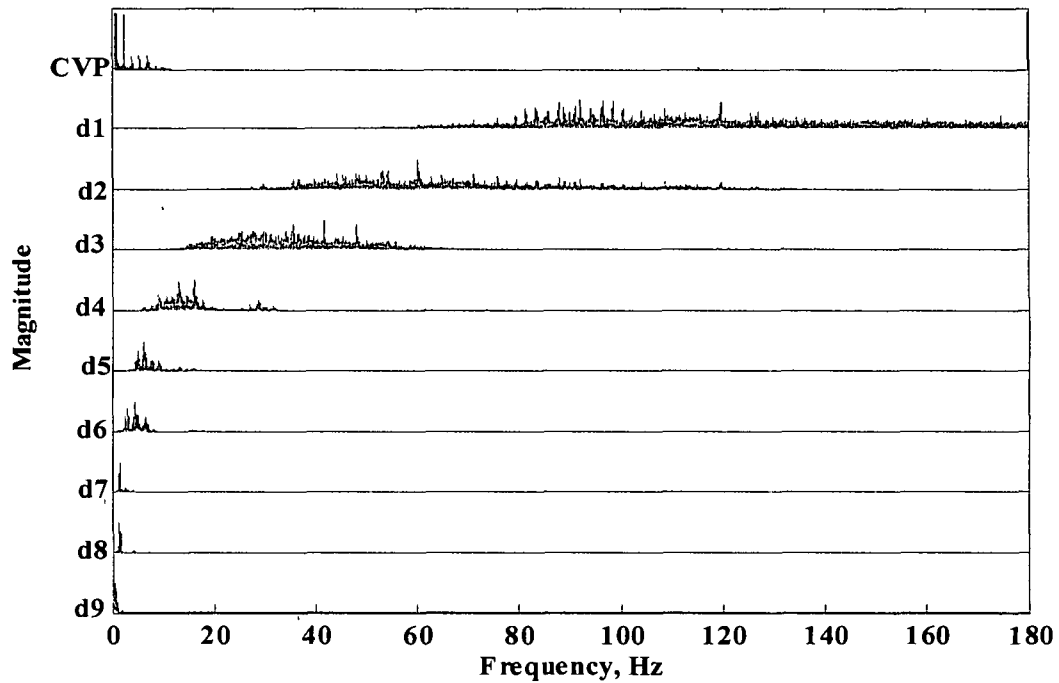
Figure 3.30 FFT of CVP signal from mgh007 record and detail coefficients (d1-d9)

Table 3.6 Frequency content of detail coefficients

Signal	Detail Coefficients	Frequency content (Hz) using db4	Frequency content (Hz) using sym4
mgh007	d1	40.99 – 180	31.22 – 179.6
	d2	27.5 – 126.9	17.6 – 141.7
	d3	10.85 – 69.68	83.09 -71.22
	d4	5.573 – 34.92	4.612 – 45.79
	d5	2.123 – 17.86	2.408 – 17.38
	d6	*0.8001 – 9.127*	*0.9352- 8.868*
	d7	*0.3533 – 4.567*	*0.6665 – 4.078*
	d8	0.16 – 1.923	0.2687 – 2.225
	d9	0.09667 – 1.123	0.07525 – 1.494

Table 3.7 Cross-correlation coefficients

Signal	Detail coefficients	Cross-correlation coefficients	Inference
mgh007	d1	0.0165	Weak
	d2	0.0315	Weak
	d3	0.0316	Weak
	d4	0.1118	Weak
	d5	0.3918	Moderate (low)
	d6	*0.4693*	Moderate (low)
	d7	*0.6014*	*Moderate (high)*
	d8	0.4091	Moderate (low)
	d9	0.1351	Weak
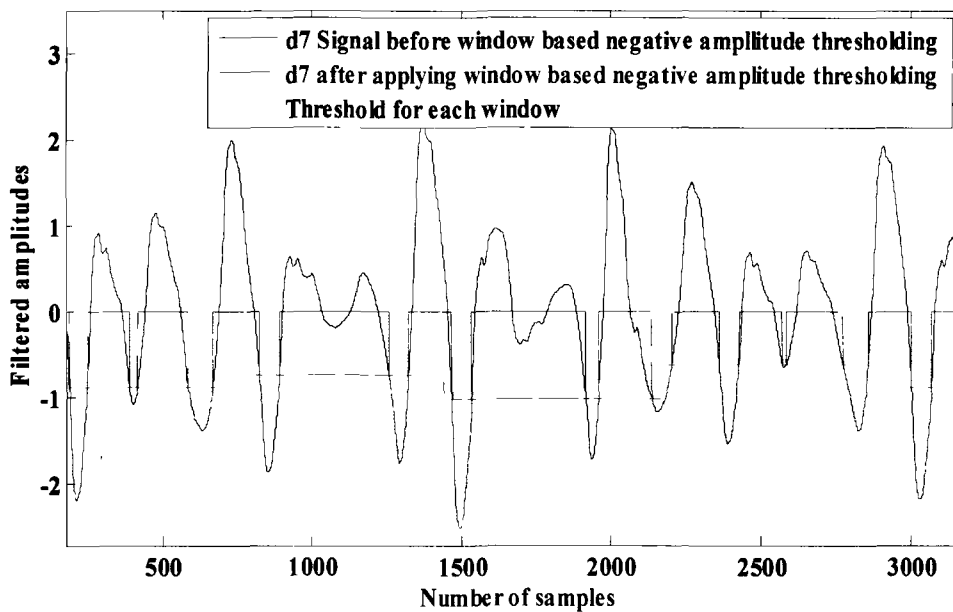
Figure 3.31 d7 signal of CVP (mgh007) signal after window based thresholding
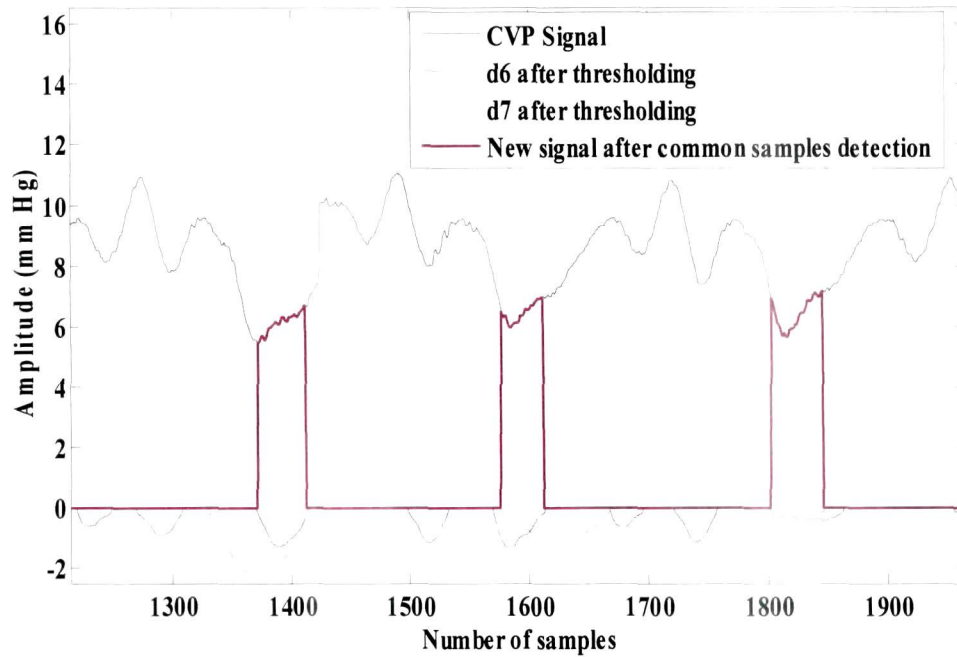
Figure 3.32 CVP detail pair for thresholding
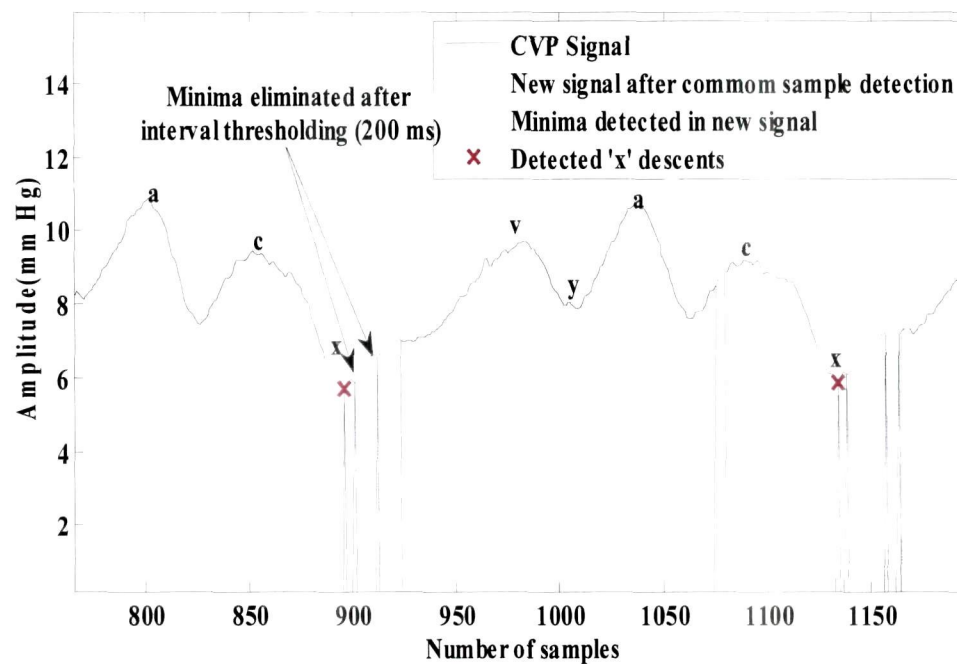
Figure 3.33 Detected 'x' descent positions of CVP signal from mgh007
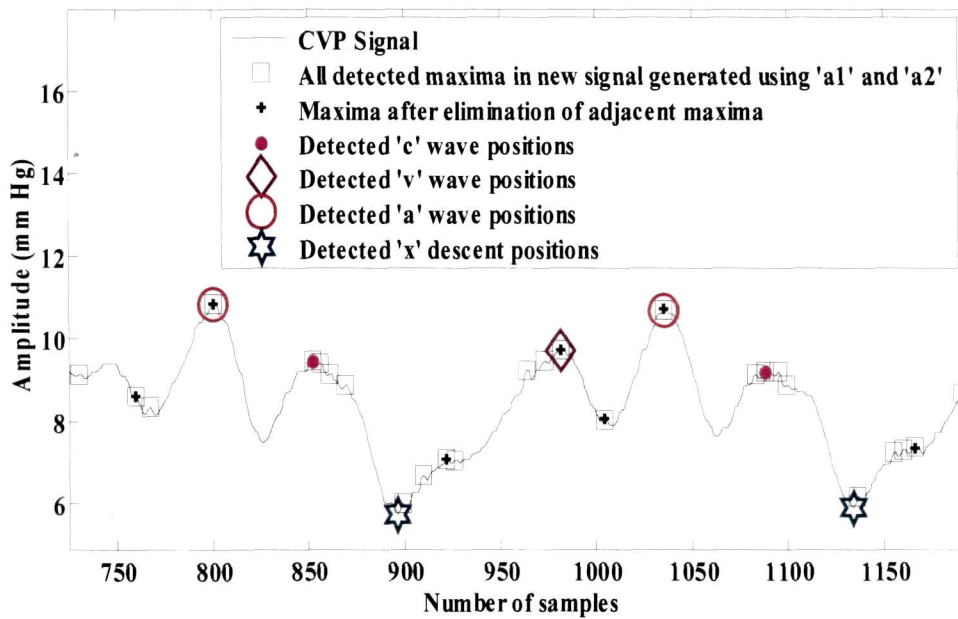
Figure 3.34 Detected *'c'*, *'v'* and *'a'* wave positions of CVP from mgh007
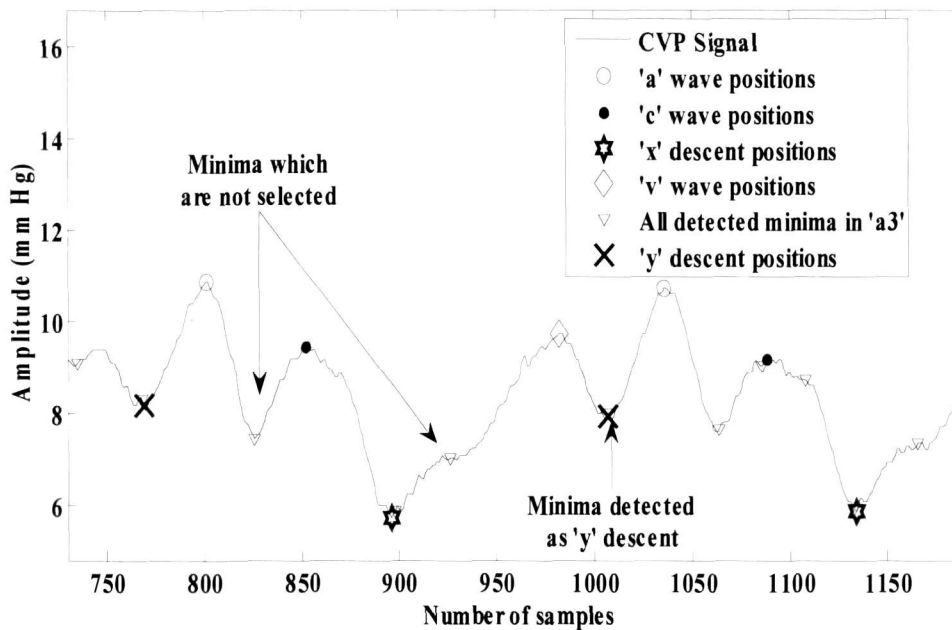
Figure 3.35 CVP signal record and all detected features along with detected minima positions in a3 for mgh007
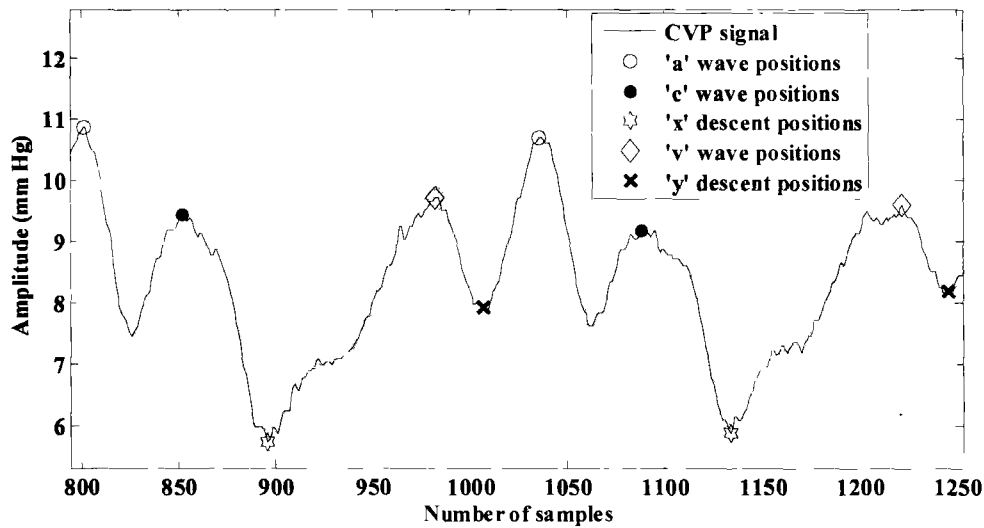
Figure 3.36 All detected components in CVP waveform of mgh007 record
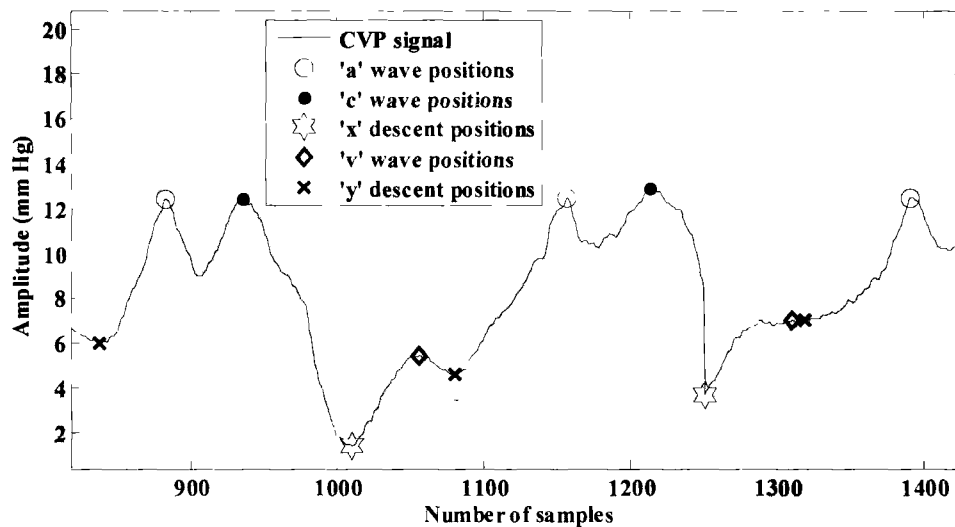
Figure 3.37 All detected components in CVP waveform of mgh008 record

3.3. Results and conclusion

There are several databases on the physionet namely Fantasia database, MIT-BIH Polysmographic database and MGH/MF waveform database etc. but the pressure signals are not annotated so far in any database. It is possible to annotate short segments of the signals manually but annotating large segments of number of signals is a tedious task. Both ABP and ECG signals are available for the same duration for the same patient in MGH/MF waveform database, Fantasia database [137] and MIT BIH Polysmographic (SLP) database [138]. ECG signals in all these databases are annotated by experts. Fantasia database

consists of non-invasive ABP waveform along with synchronously sampled ECG recording with approved beat annotations. On the contrary, SLP database is the collection of invasive ABP signals as well as synchronously sampled ECG recordings with approved beat annotations. Therefore same data segment of the ABP signal is selected for the analysis for which both ECG and ABP signals are available. Also, it has been observed that any abnormality in ECG waveform at particular instant is also available in ABP signal so the approved ECG annotations are helpful in evaluating the detector's performance to some extent [35, Figure 1.8].

Li etal [45] evaluated the performance of their algorithm on Fantasia database and Polysmographic (SLP) database. So far evaluation of beat detection algorithm on MGH/MF waveform database is not reported. Li etal [45] employed the strategy if ABP waveform is clear and corresponding ECG annotation is available, the beat annotation is considered as TP or FN based on its presence or absence. Otherwise the beat annotation is considered as FP if there is no clear ABP waveform or ECG annotation. W Zong etal [43] also employed the same technique for the performance evaluation of their algorithm based on ECG annotations.

However, it is observed practically that although recorded simultaneously, sometimes ECG annotations do not correspond to effective ABP waveforms and number of beats between both ABP and corresponding ECG differs as shown in Table (3.8 – 3.12). Table 3.11 and Table 3.12, show that there is a 23 beats difference between manual annotations of ABP peaks and ECG peak annotations for 22 signals of MGH/MF database whereas both the waveforms are clear. ECG annotations from MGH/MF database along with our manual annotations for ABP peaks for same duration of record mgh001 are shown in Figure 3.15(a-b). The selected signal segment for record mgh001 has same number of beats (12 beats) for ABP signals. This anomaly is also observed in Fantasia and MIT-BIH polysmographic database signals as shown in Table 3.12 and Table 3.13 respectively. This difference of one beat in the same duration of ABP and ECG signals is due to the fact that systolic peak in ABP waveform follows R peak of ECG.

3.3.1. Wavelet based ABP feature extraction

The algorithm has been applied for one minute segments of 22 signals from MGH/MF waveform database, first minute segments of 14 of Fantasia database, first minute segments of 15 signals of MIT-BIH polysmographic database and first 50,000 samples of abp1 signal of CSL database. CSL database is chosen because of the fact that it comprises of expert annotations for both the ABP signals. This database consists of annotations from two experts and also annotations from author are available [44]. This type of openly available database can help the researchers to validate their algorithms but ABP signals only from two patients may not be sufficient to have different artifacts and pathophysiological complexity of ABP waveform. Secondly, expert annotations only for peaks of ABP signals are available for both the recordings in CSL database whereas annotations for other features are not reported.

We have selected a certain segment of the signal records where all the features of ABP signal are prominent and can be annotated while the distorted portion of the signals are discarded for analysis.

The accuracy (A), sensitivity *(Se)* and positive predictivity *(PP)* [45] of detector is given by the relation mentioned in section 2.3.1. The formula for calculating the accuracy, sensitivity and positive predictivity are given in equations 2.14, 2.15 and 2.16 respectively. In addition to accuracy, sensitivity and positive predictivity, error is also calculated which is given by equation –

$$error = \frac{(FP + FN)}{(TP + FP)} \times 100$$

(3.5)

Where *TP* stands for the number of true positives, *FN* and *FP* denote the number of false negatives and the false positives. True positives are the beats those have been detected correctly while false positives are the beats which are detected as beats but actually do not exist. False negatives are the beats that are missed by the detector. The sensitivity depicts the percentage of true beats to overall beats those were correctly detected by the algorithm. The positive predictivity states the percentage of true beat detections to overall annotations.

The results of ABP feature extraction algorithm for each database (MGH/MF, Fantasia and MIT-BIH polysmographic) are given separately in Table 3.8 to 3.10 individually.

The results of overall performance of the algorithm on all the four database signals using 'db4' and 'sym4' wavelet is summarized in Table 3.11 and Table 3.12 respectively. It is clear from Table 3.11 and Table 3.12 that analysis with 'db4' wavelet results in more false detections for all the four components of ABP signal as compared to 'sym4' wavelet as a result reducing positive predictivity for MGH/MF and fantasia database signals. In terms of sensitivity, results by both the wavelets on both databases are comparable. On the contrary, performance of 'db4' wavelet is better on MIT-BIH polysmographic database signal in comparison to 'sym4' wavelet for all four parameters of performance evaluation such as accuracy, sensitivity, positive predictivity and error analysis of the delineator. Performance of both the wavelets is similar on CSL database signals.

3.3.2. ABP peak detection by energy analysis

The number of actual beats is counted from expert annotations for ECG signal of same duration as the ABP signal under test and beat positions are validated manually. The algorithm has been tested on the first nine records of MGH/MF waveform database. The algorithm achieved an accuracy of 99.53% for ABP signal of mgh001 record whereas the overall accuracy of detection is 98.05%. Out of total 4121 beats 4043 beats were correctly detected. The algorithm reported 78 peaks missing whereas 01 peak is detected as false beat. In addition to accuracy, two other measures of detector's performance such as sensitivity and positive predictivity are also studied. The algorithm reported sensitivity of 99.98% and positive predictive value of 98.14%.

3.3.3. CVP feature extraction

Almost all the CVP signals in MGH/MF waveform database represent abnormal waveforms and it is not possible to identify the five phasic components manually. The algorithm of CVP feature extraction is validated on two signals of MGH/MF waveform database. CVP signals of MGH/MF waveform database are not annotated by the experts and so far no other annotations from researchers are available for validation of the algorithm. Therefore, the performance of the algorithm is tested manually for CVP signals from record mgh007 and mgh008. The CVP signal from record mgh007 record has classical hemodynamic waveforms whereas CVP signal from mgh008 record has been affected by tricuspid regurgitation. The segments of CVP signal of mgh007 and mgh008

record are chosen in which all components of CVP waveform could be manually identified. The locations of 'c' and 'v' waves are identified with respect to the available ECG signal. The positive waves left to ECG peak positions are identified as 'c' waves whereas the positive waves on the right side of ECG peak positions identified as 'v' waves. The locations of 'c' and 'v' waves are easily identified with respect to ECG signal in mgh007 record whereas in mgh008, ECG signal is slightly shifted. The algorithm is applied on the selected segments of mgh007 and mgh008 records. The results of individual waves of CVP signals are presented in Table 3.13 and Table 3.14.

Table 3.8 Performance of ABP feature extraction algorithm on MGH/MF waveform database

Signal	Selected samples	Selected detail coefficie nts	No of beats ECG annotations	No of beats counted manually		Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	FP	TP	FN	FP
mgh002	395601:417200	d6-d7	79	PK	72	72	-	10	72	-	02
				OS	72	72	-	10	72	-	02
				DN	72	72	-	09	72	-	01
				DP	72	72	-	-	72	-	-
mgh006	203801:225400	d6-d7	102	PK	98	98	-	01	97	01	01
				OS	99	99	-	01	98	01	01
				DN	98	98	-	-	96	01	01
				DP	98	98	-	-	96	-	-
mgh007	266701:288300	d6-d7	93	PK	93	93	-	06	93	-	01
				OS	93	93	-	07	93	-	02
				DN	92	92	-	06	92	-	01
				DP	92	89	03	-	87	-	05
mgh008	188301:209900	d6-d7	82	PK	81	81	-	07	81	-	02
				OS	81	81	-	08	81	-	03
				DN	81	81	-	06	81	-	01
				DP	81	81	-	03	81	-	-
mgh009	154001:175600	d6-d7	115	PK	113	113	-	-	112	01	-
				OS	113	113	-	-	113	-	-
				DN	113	112	01	-	111	02	-
				DP	113	112	01	-	111	02	-
mgh010	205201:226800	d6-d7	105	PK	105	105	-	01	105	-	-
				OS	105	105	-	02	105	-	01
				DN	105	104	-	01	104	01	-
				DP	105	92	13	-	93	12	-
mgh012	183401:205000	d6-d7	91	PK	90	90	-	03	90	-	-
				OS	89	89	-	04	89	-	01
				DN	89	89	-	03	89	-	-
				DP	89	89	-	02	89	-	-
mgh014	293701:315300	d6-d7	92	PK	92	92	-	-	92	-	01
				OS	93	93	-	-	93	-	01
				DN	92	91	01	-	92	-	-
				DP	92	84	08	-	86	06	-
mgh015	338701:360300	d6-d7	94	PK	93	93	-	07	93	-	-
				OS	92	92	-	09	92	-	02
				DN	92	92	-	07	92	-	-
				DP	92	92	-	02	92	-	-
mgh016	209501:231100	d7-d8	54	PK	54	54	-	-	54	-	01
				OS	54	54	-	01	53	01	-
				DN	54	53	01	-	53	01	-
				DP	54	53	01	-	53	01	-

Contd..

Signal	Selected samples	Selected detail coefficients	No of beats ECG annotations	No of beats counted manually		Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	FP	TP	FN	FP
mgh018	1554: 23154	d6-d7	102	PK	102	97	01	05	101	01	-
				OS	101	98	-	04	101	-	-
				DN	101	97	01	04	100	01	-
				DP	101	97	-	04	100	01	-
mgh025	197501:219100	d7-d8	52	PK	53	53	-	-	53	-	-
				OS	52	52	-	01	52	-	-
				DN	52	52	-	-	52	-	-
				DP	52	51	01	-	51	01	-
mgh026	05501:27100	d6-d7	100	PK	97	95	02	02	97	02	-
				OS	96	95	01	02	95	01	-
				DN	96	95	-	01	95	01	-
				DP	96	95	-	01	95	01	-
mgh027	200501:222100	d6-d7	83	PK	82	82	-	05	82	-	02
				OS	81	81	-	06	81	-	03
				DN	81	80	01	06	81	-	02
				DP	81	80	01	04	81	-	01
mgh028	156001:177600	d7-d8	71	PK	68	68	-	-	68	-	-
				OS	67	67	-	02	67	-	02
				DN	67	67	-	-	67	-	-
				DP	67	67	-	-	67	-	-
mgh029	153601:175200	d6-d7	99	PK	98	95	-	03	97	01	-
				OS	98	95	-	03	97	01	-
				DN	97	94	-	03	96	01	-
				DP	97	85	-	12	90	07	-
mgh030	153801:175400	d6-d7	81	PK	81	81	-	19	81	-	06
				OS	82	82	-	19	82	-	06
				DN	81	81	-	18	81	-	05
				DP	81	81	-	05	81	-	02
mgh031	368001:389600	d6-d7	116	PK	116	113	-	03	113	03	-
				OS	116	114	-	02	114	02	-
				DN	115	112	-	03	112	03	-
				DP	115	112	-	03	112	03	-
mgh032	197201:218800	d6-d7	79	PK	78	78	-	14	78	-	02
				OS	79	79	-	13	79	-	02
				DN	78	78	-	13	78	-	01
				DP	78	78	-	12	78	-	01
mgh033	292901:314500	d6-d7	99	PK	99	99	-	-	99	-	-
				OS	99	99	-	-	99	-	-
				DN	98	98	-	-	98	-	-
				DP	98	97	01	-	97	01	-

Contd..

Signal	Selected samples	Selected detail coefficients	No of beats ECG annotations	No of beats counted manually		Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	TP	FN	TP	FN
mgh034	206201:227800	d7-d8	58	PK	59	59	-	-	59	-	03
				OS	58	58	-	-	58	-	03
				DN	58	58	-	-	58	-	03
				DP	58	58	-	-	58	-	-
mgh035	45911:67511	d7-d8	49	PK	49	48	01	07	49	-	-
				OS	48	47	01	09	48	-	02
				DN	48	47	01	07	48	-	-
				DP	48	47	01	02	48	-	-
MGH signals.	Total samples		Total ECG annotations for peaks	PK	1873	1862	03	93	1866	09	42
				OS	1868	1858	03	112	1862	07	31
22	4,75,200		1896	DN	1860	1843	06	87	1848	12	15
				DP	1860	1810	31	50	1818	36	09

Table 3.8 concluded

Where -

- PK – Peaks
- ON – Onsets
- DN – Dicrotic Notches
- DP – Dicrotic Peaks

Table 3.9 Performance of ABP feature extraction algorithm on Fantasia database

Signal	Selected samples	Selected detail coefficients	No of beats ECG annotations		No of beats counted manually	Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	FP	TP	FN	FP
f2o03	00001:15000	d7-d8	61	PK	57	56	01	03	57	-	-
				OS	57	56	01	03	57	-	-
				DN	57	56	01	02	56	01	-
				DP	57	56	01	01	55	02	-
f2o04	00001:15000	d6-d7	62	PK	61	60	01	07	61	-	04
				OS	62	61	01	07	62	-	04
				DN	61	60	01	06	61	-	03
				DP	61	60	01	-	61	- ı	01
f2o05	00001:15000	d6-d7	77	PK	76	76	-	-	76	-	-
				OS	77	77	-	-	77	-	-
				DN	76	75	01	-	75	01	-
				DP	76	75	01	-	75	01	-
f2o08	00001:15000	d6-d7	66	PK	59	59	-	01	54	05	-
				OS	59	59	-	01	55	04	-
				DN	58	32	26	-	53	05	-
				DP	58	32	-	13	53	05	-
f2o09	00001:15000	d7-d8	63	PK	53	50	03	-	47	06	-
				OS	54	51	03	-	47	07	-
				DN	53	48	05	-	46	07	-
				DP	53	48	05	-	46	07	-
f2o10	00001:15000	d6-d7	82	PK	78	78	-	-	78	-	-
				OS	78	78	-	-	78	-	-
				DN	78	77	01	-	77	01	-
				DP	78	77	01	-	77	01	-
f2y01	00001:15000	d6-d7	72	PK	67	67	-	-	67	-	01
				OS	68	68	-	-	68	-	01
				DN	67	66	01	-	67	-	-
				DP	67	65	02	-	63	-	04
f2y02	00001:15000	d6-d7	45	PK	40	40	-	01	38	-	02
				OS	40	40	-	02	39	-	01
				DN	40	40	-	-	37	-	03
				DP	40	39	01	-	37	-	03
f2y04	00001:15000	d6-d7	80	PK	76	76	-	03	76	-	-
				OS	77	77	-	02	77	-	-
				DN	76	76	-	02	75	01	-
				DP	76	76	-	-	73	03	-
f2y05	00001:15000	d6-d7	81	PK	76	76	-	07	76	-	-
				OS	76	76	-	07	76	-	-
				DN	76	76	-	06	75	01	-
				DP	76	76	-	02	75	01	-

Contd...

Signal	Selected samples	Selected detail coefficients	No of beats ECG annotations	No of beats counted manually		Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	FP	TP	FN	FP
f2y06	00001:15000	d6-d7	64	PK	62	62	-	01	62	-	03
				OS	62	62	-	02	62	-	04
				DN	62	62	-	-	62	-	02
				DP	62	59	03	-	62	-	-
f2y07	00001:15000	d5-d6	63	PK	55	55	-	03	50	05	-
				OS	55	55	-	04	51	04	-
				DN	55	55	-	02	49	06	-
				DP	55	51	04	-	47	08	-
f2y08	00001:15000	d6-d7	66	PK	63	63	-	04	61	02	-
				OS	63	63	-	05	62	01	-
				DN	62	62	-	04	60	02	-
				DP	62	62	-	-	57	05	-
f2y10	00001:15000	d6-d7	67	PK	61	61	-	03	61	-	04
				OS	61	61	-	03	61	-	04
				DN	60	60	-	03	60	-	04
				DP	60	60	-	02	60	-	02
Fantasia signals	Total samples		Total ECG annotations for peaks	PK	884	879	05	33	864	18	14
				OS	889	884	05	36	872	16	13
14	2,10,000		949	DN	881	845	36	25	853	25	12
				DP	881	836	18	18	841	33	06

Table 3.9 concluded

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

Table 3.10 Performance of ABP feature extraction algorithm on MIT-BIH polysmographic database

Signal	Selected samples	Selected detail coefficients	No of beats ECG annotations	No of beats counted manually		Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	FP	TP	FN	FP
slp01am	00001:15000	d6-d7	67	PK	66	65	01	-	63	03	-
				OS	65	64	01	-	62	03	-
				DN	65	64	01	-	62	03	-
				DP	65	64	01	-	62	03	-
slp01bm	00001:15000	d6-d7	60	PK	61	61	-	01	61	-	05
				OS	61	61	-	-	61	-	04
				DN	60	60	-	01	60	-	05
				DP	60	60	-	01	60	-	05
slp02am	00001:15000	d6-d7	93	PK	90	90	-	01	88	02	-
				OS	89	89	-	01	87	02	-
				DN	89	89	-	01	87	02	-
				DP	89	89	-	01	87	02	-
slp02bm	00001:15000	d6-d7	79	PK	79	79	-	-	77	02	-
				OS	78	78	-	-	76	02	-
				DN	78	78	-	-	76	02	-
				DP	78	78	-	-	76	02	-
slp03m	00001:15000	d6-d7	73	PK	72	72	-	-	72	-	-
				OS	73	71	02	-	71	'02	-
				DN	71	71	-	-	71	-	-
				DP	71	70	01	-	71	-	-
slp04m	00001:15000	d6-d7	74	PK	75	74	01	-	73	02	01
				OS	75	74	01	-	75	-	-
				DN	74	73	01	-	73	01	-
				DP	74	73	01	-	73	01	-
slp14m	00001:15000	d6-d7	68	PK	68	65	03	-	67	01	-
				OS	68	66	02	-	67	01	-
				DN	67	64	03	-	66	01	-
				DP	67	63	04	-	64	03	-
slp16m	00001:15000	d6-d7	68	PK	87	84	03	-	84	03	-
				OS	87	85	02	-	85	02	-
				DN	86	83	03	-	83	03	-
				DP	86	83	03	-	83	03	-
slp32m	00001:15000	d6-d7	71	PK	71	67	04	-	69	02	-
				OS	71	67	04	-	70	01	-
				DN	70	66	04	-	68	02	-
				DP	70	65	05	-	68	02	-
slp41m	00001:15000	d6-d7	69	PK	68	68	-	-	67	01	-
				OS	68	68	-	-	68	-	-
				DN	67	37	30	-	66	01	-
				DP	67	37	30	13	65	02	-

Contd..

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

Signal	Selected samples	Selected detail coefficients	No of beats ECG annotations	No of beats counted manually		Analysis using db4 wavelet			Analysis using Sym4 wavelet		
						TP	FN	FP	TP	FN	FP
slp45m	00001:15000	d6-d7	83	PK	81	78	03	01	79	02	01
				OS	81	79	02	01	81	-	-
				DN	80	78	02	-	79	01	-
				DP	80	78	02	-	77	03	-
slp60m	00001:15000	d6-d7	80	PK	79	79	-	-	77	02	01
				OS	79	79	-	01	78	01	01
				DN	78	78	-	-	77	-	-
				DP	78	78	-	-	77	-	-
slp61m	00001:15000	d6-d7	73	PK	72	72	-	01	70	02	-
				OS	72	72	-	02	71	01	-
				DN	71	71	-	01	69	02	-
				DP	71	71	-	01	69	02	-
slp66m	00001:15000	d6-d7	74	PK	73	73	-	-	69	04	-
				OS	74	74	-	-	70	04	-
				DN	72	72	-	-	68	04	-
				DP	72	70	02	-	68	04	-
sp67xm	00001:15000	d6-d7	69	PK	68	66	02	-	67	01	-
				OS	69	67	02	-	68	01	-
				DN	68	65	03	-	66	02	-
				DP	68	65	03	-	66	02	-
Signals	Total samples		Total ECG annotations for peaks	PK	1110	1093	17	04	1083	27	08
				OS	1110	1094	16	05	1090	20	05
15	2,25,000		1101	DN	1096	1049	47	03	1071	24	05
				DP	1096	1044	52	16	1066	29	05

Table 3.10 concluded

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

Table 3.11 Overall performance of the ABP feature extraction algorithm with db4 wavelet

Database	Features	ECG annotations		Manual Annotations	TP	FN	FP	A (%)	Se (%)	PP (%)	Error (%)
MGH/MF (22 Signals)	PK	1896		1873	1862	03	93	94.87	99.83	95.24	4.9
	OS			1868	1858	03	11	93.84	99.83	94.31	5.8
	DN			1860	1843	06	2	95.00	99.67	95.49	4.8
	DP			1860	1810	31	87 50	95.64	98.31	97.31	4.35
Fantasia (14 Signals)	PK	949		884	879	05	33	95.70	99.43	96.38	4.1
	OS			889	884	05	36	95.38	99.43	96.08	4.45
	DN			881	845	36	25	93.07	95.91	97.12	6.9
	DP			881	836	18	18	95.91	97.89	97.89	4.21
MIT-BIH Polysmograp hic (15 Signals)	PK	1101		1110	1093	17	04	98.10	98.46	99.63	1.91
	OS			1110	1094	16	05	98.10	98.55	99.54	1.91
	DN			1096	1049	47	03	99.53	95.71	99.71	4.75
	DP			1096	1044	52	16	93.79	95.25	98.49	6.41
CSL (01 Signal) (1:50000) samples	PK	M Aboy	Expert	602	602	-	-	100	100	100	0.00
	OS			602	602	-	-	100	100	100	0.00
	DN	602	602	601	600	01	01	99.67	99.84	99.84	0.34
	DP			601	600	01	01	99.67	99.84	99.84	0.34

Table 3.12 Overall performance of the ABP feature extraction algorithm with sym4 wavelet

Database	Features	ECG annotations		Manual Annotations	TP	FN	FP	A (%)	Se (%)	PP (%)	Error (%)
MGH/MF (22 Signals)	PK	1896		1873	1866	09	42	97.27	99.52	97.79	2.67
	OS			1868	1862	07	31	97.96	99.62	98.36	2.007
	DN			1860	1848	12	15	98.54	99.35	99.19	1.449
	DP			1860	1818	36	09	97.58	98.05	99.50	2.46
Fantasia (14 Signals)	PK	949		884	864	18	14	96.38	97.95	98.40	3.64
	OS			889	872	16	13	96.73	98.19	98.53	3.27
	DN			881	853	25	12	96.93	97.15	98.61	3.12
	DP			881	841	33	06	95.57	96.23	99.29	4.6
MIT-BIH Polysmograp hic (15 Signals)	PK	1101		1110	1083	27	08	96.84	97.56	97.56	3.2
	OS			1110	1090	20	05	97.74	98.19	99.54	2.28
	DN			1096	1071	24	05	97.35	97.80	99.53	2.69
	DP			1096	1066	29	05	96.89	97.35	99.53	3.17
CSL (01 Signal) (1:50000) samples	PK	M Aboy	Expert	602	602	-	-	100	100	100	0.00
	OS			602	602	-	-	100	100	100	0.00
	DN	602	602	601	600	01	01	99.67	99.84	99.84	0.34
	DP			601	600	01	01	99.67	99.84	99.84	0.34

Table 3.13 Performance of CVP feature extraction algorithm

MGH/MF Database Signal No.	Selected detail coefficients	No of beats counted manually		Analysis using db4 wavelet		
				TP	MP	FP
Mgh007	d6-d7	a	142	123	19	16
		c	142	126	15	16
		x	142	131	11	11
		v	142	121	02	21
		y	142	115	27	15
Mgh008	d6-d7	a	59	54	05	09
		c	59	52	07	11
		x	59	57	02	06
		v	59	50	09	10
		y	59	48	11	12

Table 3.14 Summary of results for CVP waveform components

CVP waves	Manual Annotations	True wave detected	Missed waves detection	False wave detection	Accuracy (%)	Sensitivity (%)	Positive predictivity(%)
a	201	177	24	25	75.62	88.05	87.62
c	201	178	22	27	75.62	89	86.82
x	201	188	13	17	85.07	93.53	91.70
v	201	171	11	31	79.10	93.95	84.65
y	201	163	38	27	67.66	81.09	85.78

Publication on this chapter

[1] Pachauri, A., & Bhuyan, M. ABP peak detection using energy analysis technique, IEEE International Conference on Multi-media, Signal Processing and Communication Technologies (IMPACT' 2011), AMU, Aligarh, India.

[2] Pachauri, A., & Bhuyan, M. Wavelet transform based arterial blood pressure waveform delineator, *International Journal of Biology and Biomedical Engineering*, 1 (6), 15-25, 2012.

CHAPTER 4

Modeling and Synthesis of ECG using System Identification Technique

4.0. Introduction

The human physiological processes are naturally interesting phenomenon when they produce biomedical signals that describe their nature and actions. These signals may be in the form of secretion of hormones, electrical in the form of potential or current, and physical in the form of pressure or temperature etc. One of the most common biomedical signals is the Electrocardiogram (ECG). Although the electro-physiological phenomenon of the heart is well known and mostly assessed by cardiologists with the help of ECG, there are few other biomedical signals generated by the blood circulatory system such as ABP, CVP, PAP signals etc., those provide sufficient understanding in critical and complex heart situations. If a system is well known, it is possible to assess the system dynamics by studying the corresponding input-output signals.

Experienced physicians can outline the cardiac condition by the examination of physiological signals such as ECG, ABP and CVP of a patient. As stated in section 1.7, ABP, CVP and ECG signals represent the heart function and possess sufficient correspondence among their constituent waveforms components. Therefore, it is possible to synthesize one signal from the other two signals if all the signals are synchronously sampled and acquired from the same patient. In intensive care units, where these signals are monitored continuously, there may be instances when actual ECG signal is missed or corrupted due to errors in sensors or due to external interruption. These interruptions in actual signals result in a great difficulty for precise diagnosis. At times, acquisition of ECG may not be possible due to surgical dressing of patients. So the synthesis of ECG using ABP and CVP signals can be used to supplement the information when actual ECG is

either missing or corrupted. Also, a single lead ECG information can be expanded to multilead information using the information derived from other cardiovascular signals such as ABP, CVP etc.

A combined investigation on ECG, CVP and ABP signals have prognostic significance as well as help to pinpoint the offending lesion when multiple obstructions are present and can thus be used to refine the treatment approach.

Although ECG, ABP and CVP are believed to be faithful independent signatures of the cardiac system, many unseen and hidden correlating facts of the relevant signals are yet to be revealed. This has led to a new area of cardiovascular research - the '*modeling of the heart*'. A model is used to replicate a system in order to give best possible results in the same way as the original system [161]. Of late the cardiovascular system has been modeled by researchers in various ways – mathematical modeling, parametric modeling, knowledge based modeling and so on. Mathematical models of heart have proved useful in understanding the heart from different point of view [162].

There are various techniques applied for modeling of ECG as stated in section 1.9.12. So far in the work, [113-131], authors have suggested ECG generation that mimics real ECGs. In physionet challenge 2010 [132], reconstruction of biomedical signals using ANN [96-99, 133] and wavelet based approach [134] has been suggested. But the main drawback of these approaches is that missing ECG signal reconstruction in one lead needs at least ECG of another lead as input [96-99, 133].

Although a considerable amount of research has been carried out for ECG reconstruction and synthesis, however, the techniques based on measured ABP and CVP signals has not been reported literature. Synthesis of ECG from measured ABP and CVP signal has been shown to be a potential modeling approach undertaken in this research. This approach to ECG modeling is inspired in view of the fact that there are sufficient correlation of features of ABP and CVP signals to that of ECGs and efficacy of system identification technique that encompasses a unique input – output relationship.

Apart from the modeling techniques described above, there is a strong tool for modeling of dynamic systems – the system identification. System Identification is the field of building

mathematical models of dynamic system based on measured data [72]. A model can be used to explain the system completely using input-output relationship. It is achieved by adjusting the parameters of a developed model until the model output best matches with the measured output System identification techniques can exemplify the physiologic mechanisms by studying measured phenomenological input -output data of a system with less physical insight. The study involves conversion of measured data to the mathematical representation of a system including analysis and transformation of developed model. This method can also be used for the prediction of future output by means of the developed model.

System identification helps in simulation of the output of a system for a given input and analyzes the system's response. System identification has been used in wide range of linear and nonlinear models such as linear non-parametric models, polynomial models, state-space models and non-linear ARX structures etc. The nonlinear system identification of dynamical systems involves structure selection, input sequence selection, noise modeling, parameter estimation and validation of model using independent dataset.

For physiological interpretation of the ECG generation method, it can be perceived that the generated ECG from the depolarization and repolarization of the potential waves in atria and ventricles stimulate the ventricle and aorta to produce pressure signals. In this chapter, we present a system identification based approach for modeling of ECG using ABP and CVP signals by autoregressive model using least square method and state space model using prediction error minimization (PEM) and subspace algorithms.

This approach for modeling and synthesis of ECG has been applied on 7 records – mgh003, mgh004, mgh005, mgh007, mgh008, mgh029 and mgh031 of MGH/MF waveform database where three signals (ABP, CVP and ECG) are available for system identification. The inputs to the model are simultaneously acquired ABP and CVP signals and synthesized output is ECG signal. The input and output signals of the model are preprocessed before testing and validation of the model. The model generates two transfer functions related to each input signal. Initially, a higher order model is obtained which is reduced to a significant order. The accuracy of synthesized ECG with measured ECG is given in terms of best fit percentage. Further, the modeled transfer functions are analyzed for stability

using pole zero plots and step responses. An interesting inference has been drawn to relate the ECG condition (i.e. normal or having a PVC or SVPB (Supraventricular premature beats)) to the stability of the model transfer function. A stable transfer function is obtained if the ECG cycles of the model are normal whereas an unstable transfer function is obtained for an abnormal ECG signal. The abnormal cycles of ECG comprise of PVC and SVPB. A supraventricular premature beat is an ectopic beat that appears in the upper chambers such as atrium, atrioventricular node or atrioventricular junction in contrast to rhythms arising in the ventricles themselves. In community prospective studies, atrial premature complexes are not related to sudden death, as are ventricular premature beats in coronary disease. Normal and SVPB annotated by cardiologist in ECG II signal of mgh004 record are shown in Figure 4.1. The normal beats are shown by a dot (.) sign whereas the supraventricular beats are labeled as 'S'.
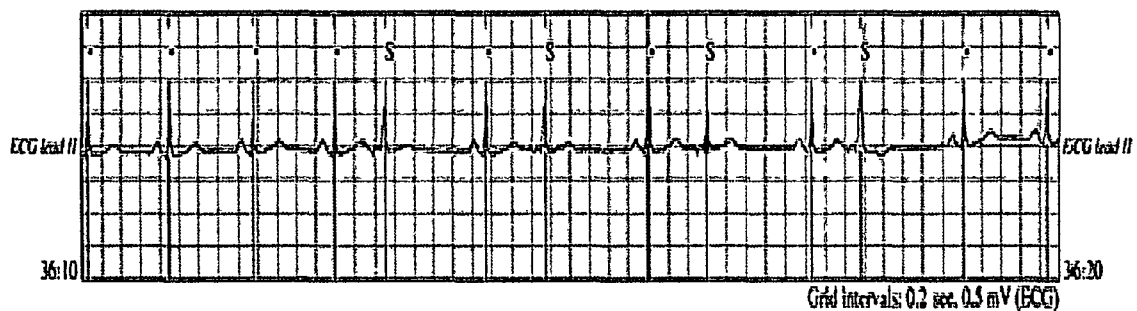
Figure 4.1 Supraventricular premature beats in ECG (lead II) signal of mgh004 record [139]

4.1. Dynamic cardiovascular model

We have used electronic recordings database (MGH/MF) of hemodynamic and electrocardiographic waveforms of patients in critical care units. The typical recordings of our interest include - ECG (lead-II), ABP and CVP signals with a sampling frequency of 360 Hz. The baseline shift and gain of each record is used for converting the signals to real engineering unit using equation 3.1. Table 4.1 shows an example of mgh004 record with ABP, CVP and ECG II signals of a patient that comprises of normal, PVC and SVPB.

Table 4.1. Description of record mgh004 of MGH/MF waveform database [139]

Record	Signals	Gain	Base	Units
mgh004	ECG lead II	1225	-964	mV
	ABP	11.98	-1213	mm Hg
	CVP	19.94	-1999	mm Hg

The input-output sample datasets are selected such that all three signals – ABP, CVP and ECG are available. There are various model structures which can be used such as - autoregressive model, autoregressive moving average models, state space models etc which are discussed in detail in section 1.9.7. In addition to these models, transfer function models are well suited for single input and single output or multiple input and single output systems. We have used autoregressive and state space models for ECG modeling and synthesis. Among these model structures, it is possible to acquire a reduced order model directly using subspace method however we have opted model order reduction to a significant order for all type of models which is useful to examine the performance of model at each order level.

Considering the cardiac system as a state space model, we can take a two-input/one-output model with the following notations-

$$Input \; 1 = u_1 = P_{ABP} \tag{4.1}$$

$$Input \; 2 = u_2 = P_{CVP} \tag{4.2}$$

$$Output \; 1 = y_1 = V_{ECG} \tag{4.3}$$

The general form of discrete state-space model can be written as-

$$x(k+1) = Ax(k) + Bu(k) \tag{4.4}$$

$$y(k+1) = Cx(k) + Du(k) \tag{4.5}$$

Where, x is the state, u is the input and y is the output; A and C are state matrices; B and D are the output matrices. The state-space representation of the two-input/single-output cardiac model can be written as –

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} \begin{bmatrix} P_{ABP}(k) \\ P_{CVP}(k) \end{bmatrix} \tag{4.6}$$

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

$$V_{ECG}(k+1) = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} P_{ABP}(k) \\ P_{CVP}(k) \end{bmatrix} \qquad (4.7)$$

Considering the cardiac system as two distinct discrete ARX models relating to each input, equivalent ARX models can be written as -

$$V_{ECG}(1) + a_1 V_{ECG}(2) + a_2 V_{ECG}(3)........ a_k V_{ECG}(k) = b_1 P_{ABP}(1) + ... b_m P_{ABP}(m) \quad (4.8)$$

$$V_{ECG}(1) + a_1 V_{ECG}(2) + a_2 V_{ECG}(3)........ a_k V_{ECG}(k) = b_1 P_{CVP}(1) + ... b_m P_{CVP}(m) \quad (4.9)$$

Where k and m are sampling instants of ECG and CVP/ABP respectively, however, in our case we have taken k = m.

Before system identification of the parametric models and simulation, ABP, CVP and ECG signals in the input-output datasets are preprocessed by 20 point moving average filter to reduce the noise and preserve a sharp step response. Then the offset levels in each dataset are removed to bring the signal to equilibrium by subtracting the offsets of input and output signals from the actual input-output values using equations 1.91, 1.92.

The offset nullification operation removes the linear trend in the data and thus helps to estimate the non-linear system near to the point of linearization in the selected operating range. A common flowchart of system identification based modeling using PEM and subspace method is shown in Figure 4.2. The structure of state space model is selected depending upon the number of inputs and outputs in the test dataset. Initially a 6th order state space model is obtained. This estimated higher order model captures most of the information in the data. The model reduction step then extracts the most significant states of the model. The model order can be reduced by the visual examination of poles and zeros of estimated state space model or the response to stimulated inputs. Poles and zeros that practically cancel each other are due to over parameterization in the high-order model whereas the poles and zeros that do not overlap are assumed to correspond to actual system dynamics.
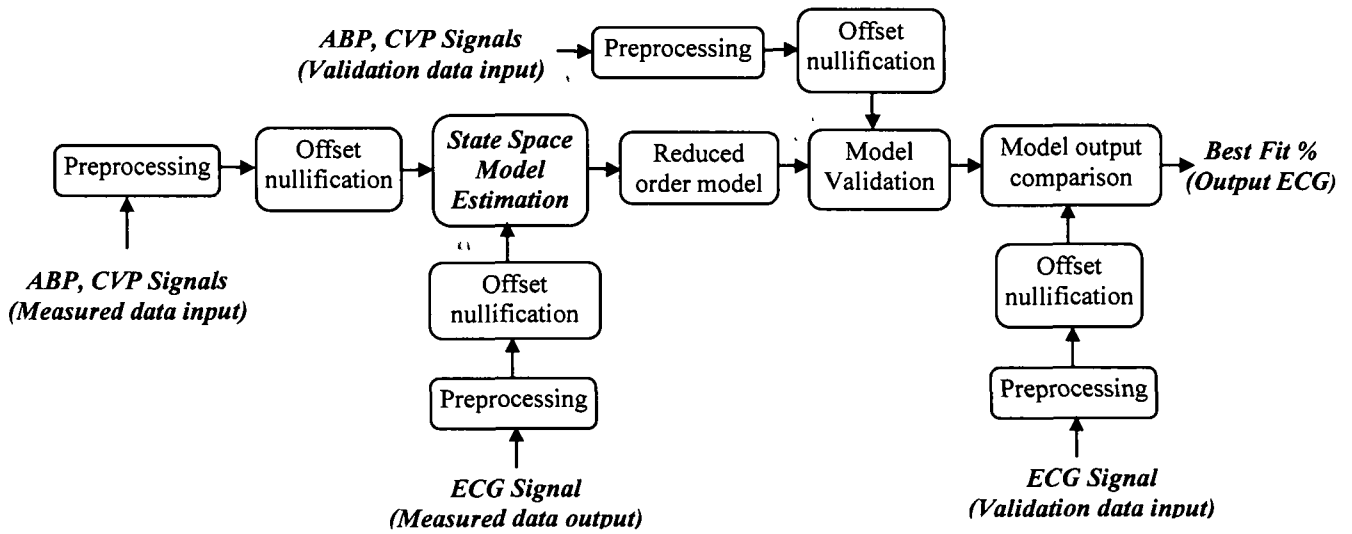
Figure 4.2. Flow chart of ECG modeling using system identification

Modeling of ECG from ABP and CVP signals in performed for normal, PVC and supra-ventricular premature beats (SVPB) as stated below –

4.1.1. Modeling with normal data

In this part, a state space model of ECG has been simulated using expert annotated data of a normal subject. The segments of ABP and CVP signals those represent a normal sinus cycle of ECG are applied as input to the model whereas the corresponding normal ECG data are applied as output for simulating the model. Using the record mgh004 with normal data, the following 4th order state space model is obtained applying subspace algorithm and model order reduction with an error (E) value of 6.59×10^{-6}.

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 0.9908 & 0.16379 & -0.015092 & -0.0011222 \\ -0.17269 & 0.91385 & -0.17428 & -0.032923 \\ -0.022546 & 0.078189 & 0.85418 & -0.39799 \\ 0.040978 & 0.013527 & 0.46235 & 0.22968 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} -0.18387 & 1.3201 \\ 0.47547 & -5.832 \\ 0.65775 & -10.09 \\ -0.070118 & -2.1175 \end{bmatrix} \begin{bmatrix} P_{ABP}(k) \\ P_{CVP}(k) \end{bmatrix}$$

$$(4.10)$$

$$V_{ECG}(k+1) = \begin{bmatrix} -2.0502 & -0.16245 & 0.015888 & 0.0010386 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix}$$

$$(4.11)$$

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

Where, x, u, y have the same meaning as defined in equation (4.1 – 4.5). In the study of linear time-invariant systems, it is of great advantage to use the Laplace transform and z-transform techniques [163]. The state space equation represented in equations (4.10-4.11) is the matrix of two transfer functions relating each input to output signal. These transfer functions are denoted as V_{ECG}/ P_{ABP} and V_{ECG}/ P_{CVP}.

The obtained transfer functions for modeling of normal beats in z-domain are given by –

V_{ECG}/ P_{ABP}:

$$G_1(z) = \frac{V_{ECG}}{P_{ABP}} = \frac{0.37234(z-1.419)(z^2 - 2.142z + 1.154)}{(z^2 - 1.929z + 0.9469)(z^2 - 1.483z + 0.8173)} \tag{4.12}$$

V_{ECG}/ P_{CVP}:

$$G_2(z) = \frac{V_{ECG}}{P_{CVP}} = \frac{0.86292(z+0.1587)(z^2 - 2.084z + 1.113)}{(z^2 - 1.873z + 0.912)(z^2 - 1.573z + 0.8515)} \tag{4.13}$$

Corresponding transfer function in Laplace domain is given by –

V_{ECG}/ P_{ABP}:

$$G_1(s) = \frac{V_{ECG}}{P_{ABP}} = \frac{2148545(s-1144)(s^2 - 53.27s + 1481)}{(s^2 + 19.65s + 2413)(s^2 + 72.62s + 49410)} \tag{4.14}$$

V_{ECG}/ P_{CVP}:

$$G_2(s) = \frac{V_{ECG}}{P_{CVP}} = \frac{-152.1646(s+1048)(s^2 - 38.59s + 3580)}{(s^2 + 33.15s + 5325)(s^2 + 57.85s + 40040)} \tag{4.15}$$

ECG and ABP are the functions of time (t). If we represent ECG as V_E and ABP as P_A then as time (t) increases the amplitude of both ECG and ABP change, therefore the differential equation corresponding to transfer function (4.14) for input ABP to output ECG is given as –

$$\frac{d^4V_E}{dt^4} + 92.27\frac{d^3V_E}{dt^3} + 53250\frac{d^2V_E}{dt^2} + 1.14\times10^6\frac{dV_E}{dt} + 12\times10^8 V_E = 214\ 85\frac{d^3P_A}{dt^3} - 36024\frac{d^2P_A}{dt^2} + 1.62\times10^6\frac{dP_A}{dt} - 363\times10^7 P_A \tag{4.16}$$

Similarly, differential equation corresponding to transfer function (4.15) for input CVP to output ECG is given as –

$$\frac{d^4 V_E}{dt^4} + 90\frac{d^3 V_E}{dt^3} + 4.72\times10^4 \frac{d^2 V_E}{dt^2} + 1.63\times10^6 \frac{dV_E}{dt} + 2.13\times10^8 V_E = -152.16\frac{d^3 P_C}{dt^3} + 153590\frac{d^2 P_C}{dt^2} - 5.6\times10^6 \frac{dP_C}{dt} - 5.7\times10^8 P_C$$

$$(4.17)$$

From equations (4.16) & (4.17), it is clear that all the coefficients in L.H.S of the equation are present and have the positive values. Therefore, according to Routh Hurwitz stability criterion described in section 1.9.10(c), it can be concluded that model of normal heart is stable. Further, stability analysis of this model is done using pole zero plots and step responses described in section 4.2.2.

4.1.2. Modeling with PVC and SVPB beats

Like any other arrhythmia condition of the heart, PVC and SVPB are also an indication of an abnormal heart conditions. We have chosen expert annotated ECG signals having PVC and SVPB for modeling an abnormal heart condition in state space domain.

4.1.2.1. Modeling with PVC beat

Using the PVC beat of mgh004, the following 4th order state space model is obtained using subspace algorithm and model order reduction with an error (E) value of 1.288×10^{-6} –

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 0.98943 & -0.11671 & 0.006669 & 0.0001612 \\ 0.11924 & 0.98444 & -0.10416 & 0.0063682 \\ 0.0003631 & 0.089774 & 0.91861 & 0.15804 \\ -0.055251 & -0.061119 & -0.58496 & -0.077513 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} -0.32163 & 0.5006 \\ -6.8394 & 5.08592 \\ -48.59 & 34.267 \\ 280.26 & -195.82 \end{bmatrix} \begin{bmatrix} P_{ABP}(k) \\ P_{CVP}(k) \end{bmatrix}$$

$$(4.18)$$

$$V_{ECG}(k+1) = \begin{bmatrix} 3.5453 & -0.19165 & 0.011216 & 0.00034256 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix}$$

$$(4.19)$$

V_{ECG}/ P_{ABP}:

$$G_1(z) = \frac{V_{ECG}}{P_{ABP}} = \frac{0.93508(z - 0.5546)(z^2 - 1.97z + 1.02)}{(z^2 - 1.95z + 0.9924)(z^2 - 1.808z + 1.023)}$$

$$(4.20)$$

V_{ECG}/ P_{CVP}:

$$G_2(z) = \frac{V_{ECG}}{P_{CVP}} = \frac{0.41941(z - 0.9078)(z^2 - 1.881z + 1.08)}{(z^2 - 1.949z + 0.9915)(z^2 - 1.808z + 1.023)}$$

$$(4.21)$$

Corresponding transfer functions in Laplace domain are given by

V_{ECG}/P_{ABP}:

$$G_1(s) = \frac{V_{ECG}}{P_{ABP}} = \frac{271.6585(s+202.9)(s^2-7.265s+6431)}{(s^2+2.75s+5568)(s^2-8.207s+28080)}$$

(4.22)

V_{ECG}/P_{CVP}:

$$G_2(s) = \frac{V_{ECG}}{P_{CVP}} = \frac{150.2721(s+34.67)(s^2-27.96s+25170)}{(s^2+3.081s+5612)(s^2-8.207s+28080)}$$

(4.23)

Therefore, the differential equation corresponding to transfer function (4.22) for input ABP to output ECG in case of modeling of PVC beat is given by –

$$\frac{d^4V_E}{dt^4} - 5.45\frac{d^3V_E}{dt^3} - 33625\frac{d^2V_E}{dt^2} - 45696\frac{dV_E}{dt} + 1.56\times10^8 V_E = 271.6\frac{d^3P_A}{dt^3} - 53146\frac{d^2P_A}{dt^2} + 4957\frac{dP_A}{dt} + 1.30\times10^6 P_A$$

(4.24)

Similarly, differential equation corresponding to transfer function (4.23) for input CVP to output ECG in case of modeling of PVC beat is given by –

$$\frac{d^4V_E}{dt^4} - 5.12\frac{d^3V_E}{dt^3} + 33667\frac{d^2V_E}{dt^2} + 40457\frac{dV_E}{dt} + 1.57\times10^8 V_E = 150.2\frac{d^3P_C}{dt^3} - 1008\frac{d^2P_C}{dt^2} - 3.63\times10^6 \frac{dP_C}{dt} + 1.31\times10^8 P_C$$

(4.25)

From equation (4.24) and equation (4.25), it is clear that all coefficients in L.H.S. of the equation are present, but in equation (4.24), three coefficients are negative and similarly in equation (4.25), one coefficient is negative, therefore, following the same criteria of stability, it can be stated that model of abnormal heart with PVC beat is unstable.

4.1.2.2. Modeling with SVPB

Using the SVPB of mgh004, the following 4th order state space model using subspace algorithm is obtained after model order reduction with an error (E) value of 1.00924 $\times10^{-6}$ -

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 0.95402 & -0.15783 & 0.004918 & 0.000424 \\ 0.1585 & 1.0041 & -0.066568 & -0.00224 \\ -0.01944 & 0.092555 & 0.80828 & -0.25101 \\ -0.00511 & -0.00336 & 0.14092 & 0.59891 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} -0.34027 & 0.12906 \\ -0.66074 & 0.072927 \\ -0.2206 & 0.5472 \\ 3.9743 & -0.15361 \end{bmatrix} \begin{bmatrix} P_{ABP}(k) \\ P_{CVP}(k) \end{bmatrix}$$

(4.26)

$$V_{ECG}(k+1) = \begin{bmatrix} 1.6866 & -0.12317 & 0.004007 & 1.3222 \times 10^{-5} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{4.27}$$

The obtained transfer functions for modeling of SVPB beats in z-domain are given by –

V_{ECG}/ P_{ABP} :

$$G_1(z) = \frac{V_{ECG}}{P_{ABP}} = \frac{0.708(z - 3.706)(z^2 - 2.12z + 1.126)}{(z^2 - 1.886z + 0.9232)(z^2 - 1.783z + 1.187)} \tag{4.28}$$

V_{ECG}/ P_{CVP}:

$$G_2(z) = \frac{V_{ECG}}{P_{CVP}} = \frac{-0.46924(z - 3.106)(z^2 - 2.001z + 1.023)}{(z^2 - 1.871z + 0.9087)(z^2 - 1.783z + 1.187)} \tag{4.29}$$

Corresponding transfer functions in Laplace domain are given by

V_{ECG}/ P_{ABP}:

$$G_1(s) = \frac{V_{ECG}}{P_{ABP}} = \frac{652.065(s - 396.4)(s^2 - 43.27s + 698.9)}{(s^2 + 28.78s + 5043)(s^2 - 61.68s + 49500)} \tag{4.30}$$

V_{ECG}/ P_{CVP}:

$$G_2(s) = \frac{V_{ECG}}{P_{CVP}} = \frac{-358.4381(s - 357.8)(s^2 - 8.40s + 2801)}{(s^2 + 34.45s + 5180)(s^2 - 61.68s + 49500)} \tag{4.31}$$

Therefore, the differential equation corresponding for transfer function (4.30) for input ABP to output ECG in case of modeling of SVPB beat is given by -

$$\frac{d^4 V_E}{dt^4} - 27.23\frac{d^3 V_E}{dt^3} - 52555\frac{d^2 V_E}{dt^2} - 1.38 \times 10^6 \frac{dV_E}{dt} + 1.38 \times 10^6 V_E = 652\frac{d^3 P_A}{dt^3} - 286693\frac{d^2 P_A}{dt^2} + 1.18 \times 10^7 \frac{dP_A}{dt} - 4.36 \times 10^8 P_A \tag{4.32}$$

Similarly, the differential equation corresponding to transfer function (4.31) for input CVP to output ECG in case of modeling of SVPB beat is given by –

$$\frac{d^4 V_E}{dt^4} - 27.23\frac{d^3 V_E}{dt^3} + 3055\frac{d^2 V_E}{dt^2} + 1.38 \times 10^6 \frac{dV_E}{dt} - 2.56 \times 10^8 V_E = -3.58 \times 10^6 \frac{d^3 P_C}{dt^3} - 1.31 \times 10^9 \frac{d^2 P_C}{dt^2} + 3005\frac{dP_C}{dt} - 1.002 \times 10^6 P_C \tag{4.33}$$

Similarly, in case of modeling with SVPB beat, following the equations (4.32) and (4.33), it is evident that the model of abnormal heart in case of SVPB is also unstable.

4.2. Results and validation

The modeling approach is applied on the seven records - mgh003, mgh004, mgh005, mgh007, mgh008, mgh029 and mgh031 of MGH/MF waveform database. System identification tries to approximate the model by minimizing the error. This error (E) is a measure of model quality obtained by simulating the model by different inputs. The error (E) is given by the following relation –

$$E = V \left[\frac{1 + \dfrac{d}{N}}{1 - \dfrac{d}{N}} \right] \tag{4.34}$$

Where, 'V' is the loss function, 'd' denotes the number of estimated parameters and 'N' is the number of values in the estimated dataset. The errors (e) between the measured and the modeled outputs are weighed at specific frequencies (λ_i) during the minimization of the following loss function which is given by –

$$\textit{Loss Function } (V) = \sum_i \lambda_i \, e_i^2 \tag{4.35}$$

The obtained 4th order reduced model is validated using other cycles of ABP and CVP signals of the same patient.

4.2.1. Accuracy of synthesized ECG cycles

A medical application algorithm for disease diagnosis needs full proof detection criterion which will ensure that the identification algorithm provides an acceptance level of confidence to doctors. An universally accepted metric is used to compare the synthesized and actual signals given by 'Best fit' in percentage (%) –

$$\textit{Best Fit } (\%) = \left(1 - \frac{|y - \hat{y}|}{|y - \bar{y}|} \right) \times 100 \tag{4.36}$$

Where,

$\hat{y}$ = Estimated model output

y = Measured output

$\bar{y}$ = Mean of measured output

It is observed that a better fit percentage is obtained when a developed ABP/CVP of normal heart model is validated using normal dataset. On the other hand,, the fit percentage goes down when the model of a normal heart is validated with abnormal beat (PVC/SVPB). Similarly, a model of an abnormal heart results a better fit percentage if it is validated by abnormal dataset.

First , we have used seven number of records for developing transfer function models using the signals – ABP, CVP and ECG. From each record, both normal and abnormal cycles were used for modeling as stated in Table 4.2 –

Table 4.2. Records used for modeling and validation

Record No.	No of model of normal cycle			No of model of abnormal cycle		
	Model	Validation normal	Validation abnormal	Model	Validation normal	Validation abnormal
mgh003	1	2	1	1	×	1 (PVC)
mgh004	1	1	1 (PVC)	1 (PVC)	×	1 (PVC)
			1 (SVPB)	1 (SVPB)	×	1 (SVPB)
mgh005	1	3	×	×	×	×
mgh007	1	3	×	×	×	×
mgh008	1	3	×	×	×	×
mgh029	1	5	×	×	×	×
mgh031	1	16	1 (PVC)	×	×	×

The estimated transfer functions of normal and abnormal models are shown in Table 4.3. The synthesized ECG using normal and abnormal model of mgh004 are shown in Figure 4.3 (a-e).

The best fit for synthesized ECG cycles for a record is defined as the average of best fit of each synthesized cycle. The average best fit obtained for five cycles of mgh029 record and seventeen cycles of mgh031 record are 75.3557% and 63.3793% respectively. Using this method any length of ECG signal can be generated if corresponding ABP and CVP signals of same duration are known. Synthesized ECG cycles for records mg029 and mgh031 are shown in Figure 4.4 (a-b) respectively.
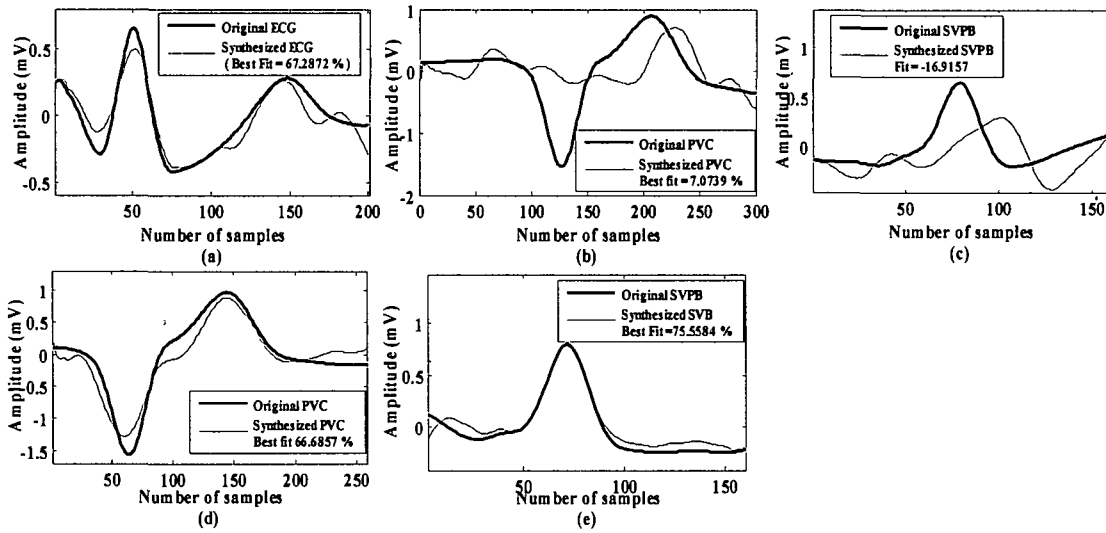
Figure 4.3 Original and synthesized ECG for mgh004 record, modeling with normal and validating with (a) normal (b) PVC (c) SVPB, (d) modeling with PVC and testing with PVC and (e) modeling with SVB and testing with SVPB
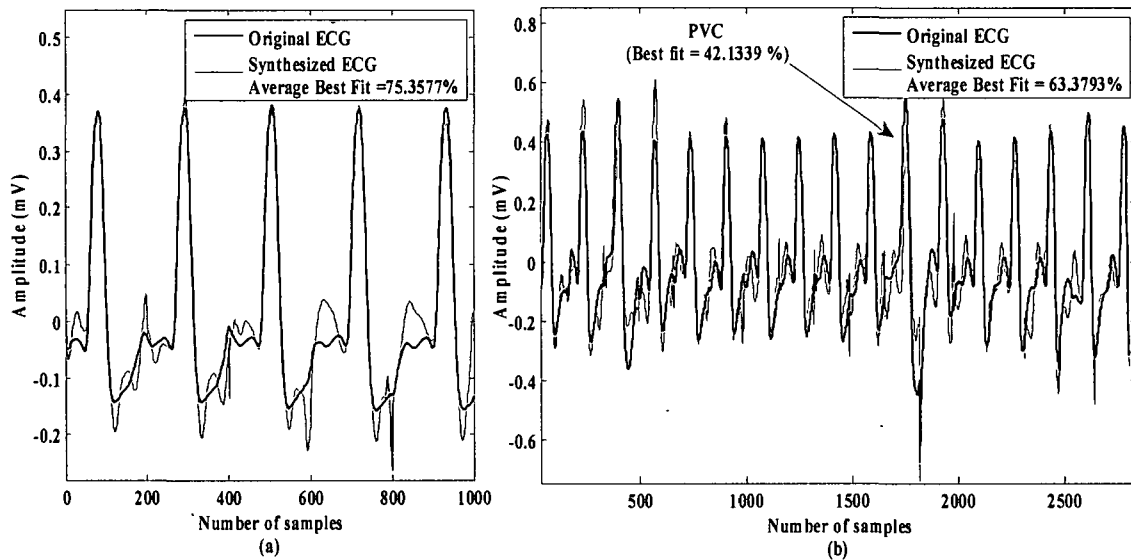
Figure 4.4 Original and synthesized ECG for (a) mgh029 record (5 cycles) and (b) mgh031 record (17 cycles)

Table 4.3 Results of ECG modeling by subspace method

Record No	Modeling			Validation			
	Beat type	Transfer function (z-domain)	Stability	Beat Type	Transfer function (z-domain)	Fit (%)	Stability
Mgh003	Normal	$V_{ECG}/P_{ABP}:$ $$G_1(z)=\frac{-1\,7594(z-1\,042)(z^2-2\,2z+1\,489)}{(z^2-1\,805z+0\,8675)(z^2-1\,652z+0\,865)}$$ $V_{ECG}/P_{CVP}:$ $$G_2(z)=\frac{-3\,2181(z-1\,093)(z^2-1\,427z+0\,8157)}{(z^2-1\,852z+0\,9038)(z^2-1\,697z+0\,8633)}$$	Stable	Normal	$V_{ECG}/P_{ABP}:$ $$G_3(z)=\frac{-3\,7911(z-2\,787)(z^2-2\,065z+1\,075)}{(z^2-1\,438z+0\,5382)(z^2-1\,777z+0\,9132)}$$ $V_{ECG}/P_{CVP}:$ $$G_4(z)=\frac{-0\,70586(z+0\,4163)(z^2-2\,139z+1\,165)}{(z^2-1\,753\,z+0\,819)(z^2-1\,773z+0\,9137)}$$	63 6651	Stable
				Normal	$V_{ECG}/P_{ABP}:$ $$G_5(z)=\frac{-6\,0556(z-1\,021)(z^2-2\,379z+1\,482)}{(z^2-1\,811z+0\,8561)(z^2-1\,683z+0\,9374)}$$ $V_{ECG}/P_{CVP}:$ $$G_6(z)=\frac{-0\,063075\ \ (z+1\,306\ )(z-1\,081)(z+0\,3207)}{(z^2-1\,857\,z+0\,8924)(z^2-1\,676z+0\,9201)}$$	61 5860	Stable
				PVC	$V_{ECG}/P_{ABP}:$ $$G_7(z)=\frac{0\,98282(z-0\,9966)(z^2-2\,121z+1\,129)}{(z-1\,008)(z-0\,7104)(z^2-2\,022z+1\,046)}$$ $V_{ECG}/P_{CVP}:$ $$G_8(z)=\frac{0\,045032\ (z-3\,416\ )(z^2-1\,997\,z+1\,001)}{(z-1\,008\ )(z-0\,8779\ )(z^2-2\,022\,z+1\,046)}$$	-53 55	Unstable
	PVC	$V_{ECG}/P_{ABP}:$ $$G_9(z)=\frac{0\,21313(z+2\,779)\ (z^2-2\,059\,z+1\,062)}{(z-1\,053)\ (z-0\,4555)(z^2-2\,173z+1\,19)}$$ $V_{ECG}/P_{CVP}:$ $$G_{10}(z)=\frac{0\,64683(z-1\,192)(z^2-2\,028z+1\,031)}{(z-1\,053)\ (z-0\,9027)\ (z^2-2\,173z+1\,19)}$$	Unstable	PVC	$V_{ECG}/P_{ABP}:$ $$G_{11}(z)=\frac{5\,8797(z-0\,344)\ (z^2-1\,85z+0\,8962)}{(z^2-2\,056z+1\,083)(z^2-0\,9401z+1\,025)}$$ $V_{ECG}/P_{CVP}:$ $$G_{12}(z)=\frac{-0\,029921\,(z+0\,5035\ )(z^2-1\,677z+0\,8035)}{(z^2-2\,056\,z+1\,083)(z^2-0\,9401z+1\,025)}$$	61 0222	Unstable

Mgh004	Type	Transfer Functions	Stability	Type	Transfer Functions	Value	Stability
	Normal	V_{ECG}/P_{ABP}: $$G_1(z)=\frac{0\,37234(z-1\,419)(z^2-2\,142z+1\,154)}{(z^2-1\,929z+0\,9469)(z^2-1\,483z+0\,8173)}$$ V_{ECG}/P_{CVP}: $$G_2(z)=\frac{0\,86292(z+0\,1587)(z^2-2\,084z+1\,113)}{(z^2-1\,873z+0\,912)(z^2-1\,573z+0\,8515)}$$	Stable	Normal	V_{ECG}/P_{ABP}: $$G_3(z)=\frac{-0.33927(z-5.77)(z-0.8924)(z-0.6722)}{(z^2-1.808z+0.879)(z^2-1.523z+0.7038)}$$ V_{ECG}/P_{CVP}: $$G_4(z)=\frac{0\,23706\,(z-1\,687)(z-1\,059)(z-0\,5241)}{(z^2-1\,807\,z+0\,87)(z^2-1\,728\,z+0\,8502)}$$	67 2842	Stable
				PVC	V_{ECG}/P_{ABP}: $$G_5(z)=\frac{-0.064572(z-1.052)(z^2-1.87z+0.8905)}{(z-0.9763)(z-1.57)(z^2-1.97z+0.9777)}$$ V_{ECG}/P_{CVP}: $$G_6(z)=\frac{-0.014666(z-9.967)(z^2-1.955z+0.957)}{(z-0.9765)(z-1.57)(z^2-1.907z+0.9217)}$$	7 0739	Unstable
				SVPB	V_{ECG}/P_{ABP}: $$G_7(z)=\frac{0.14315(z-1.427)(z^2-2.018z+1.031)}{(z-1.102)(z-0.7415)(z^2-1.963z+1.1002)}$$ V_{ECG}/P_{CVP}: $$G_8(z)=\frac{4.0495(z-1.088)(z^2-1.961z+0.9977)}{(z-1.102)(z-0.7427)(z^2-1.963z+1.002)}$$	-16 9157	Unstable
	PVC	V_{ECG}/P_{ABP}: $$G_9(z)=\frac{0\,93508(z-0\,5546)(z^2-1\,97z+1\,02)}{(z^2-1\,95z+0\,9924)(z^2-1\,808z+1\,023)}$$ V_{ECG}/P_{CVP}: $$G_{10}(z)=\frac{0\,41941(z-0\,9078)(z^2-1\,881z+1\,08)}{(z^2-1\,949z+0\,9915)(z^2-1\,808z+1\,023)}$$	Unstable	PVC	V_{ECG}/P_{ABP}: $$G_{11}(z)=\frac{1.9247(z-1.063)(z^2-1.952z+0.9535)}{(z-0.3918)(z-1.007)(z^2-1.917z+0.9239)}$$ V_{ECG}/P_{CVP}: $$G_{12}(z)=\frac{0.93425(z-1.024)(z^2-2.001z+1.004)}{(z-0.7568)(z-1.007)(z^2-1.906z+0.919)}$$	68 6857	Unstable
	SVPB	V_{ECG}/P_{ABP}: $$G_{13}(z)=\frac{0\,708(z-3\,706)(z^2-2\,12z+1\,126)}{(z^2-1\,886z+0\,9232)(z^2-1\,783z+1\,187)}$$ V_{ECG}/P_{CVP}: $$G_{14}(z)=\frac{-0\,46924(z-3\,106)(z^2-2\,001z+1\,023)}{(z^2-1\,871z+0\,9087)(z^2-1\,783z+1\,187)}$$	Unstable	SVPB	V_{ECG}/P_{ABP}: $$G_{15}(z)=\frac{-0.70751(z-1.417)(z^2-2.086z+1.133)}{(z^2-1.991z+1.032)(z^2-1.756z+0.9542)}$$ V_{ECG}/P_{CVP}: $$G_{16}(z)=\frac{0.66185(z+0.1105)(z^2-1.97z+0.9938)}{(z^2-1.991z+1.032)(z^2-1.756z+0.9534)}$$	75 5584	Unstable

		V_{ECG}/P_{ABP} and V_{ECG}/P_{CVP}			V_{ECG}/P_{ABP} and V_{ECG}/P_{CVP}		
Mgh005	Normal	V_{ECG}/P_{ABP}: $$G_1(z)=\frac{-0.65081(z-2.445)(z-0.9441)(z-0.9365)}{(z-0.2618)(z-0.6754)(z^2-1.889z+0.9165)}$$ V_{ECG}/P_{CVP}: $$G_2(z)=\frac{-0.38231(z-1.104)(z^2-2.053z+1.081)}{(z-0.3238)(z-0.9832)(z^2-1.895z+0.9211)}$$	Stable	Normal	V_{ECG}/P_{ABP}: $$G_3(z)=\frac{-0.89891(z-0.5071)(z-1.05)(z-1.179)}{(z^2-1.918z+0.945)(z^2-1.668z+0.8678)}$$ V_{ECG}/P_{CVP}: $$G_4(z)=\frac{-0.0054126(z-0.5125)(z^2-2.048z+1.076)}{(z^2-1.908z+0.9381)(z^2-1.676z+0.8815)}$$	60 95	Stable
				Normal	V_{ECG}/P_{ABP}: $$G_5(z)=\frac{-0.34614(z-1.864)(z-0.9196)(z-0.8799)}{(z^2-1.946z+0.9739)(z^2-1.523z+0.7761)}$$ V_{ECG}/P_{CVP}: $$G_6(z)=\frac{-0.046274(z-1.169)(z^2-2.219z+1.248)}{(z-0.8158)(z-0.9955)(z^2-1.943z+0.9727)}$$	50 1025	Stable
				Normal	V_{ECG}/P_{ABP}: $$G_7(z)=\frac{-1.2848(z-1.251)(z-0.9205)(z-0.8861)}{(z^2-1.931z+0.9603)(z^2-1.633z+0.8236)}$$ V_{ECG}/P_{CVP}: $$G_8(z)=\frac{-0.076017(z-1.112)(z^2-1.981z+1.004)}{(z^2-1.937z+0.9382)(z^2-1.894z+0.9523)}$$	42 77	Stable
Mgh007	Normal	V_{ECG}/P_{ABP}: $$G_1(z)=\frac{0.60709(z-1.344)(z^2-1.111z+0.6645)}{(z^2-1.889z+0.9601)(z^2-1.041z+0.9568)}$$ V_{ECG}/P_{CVP}: $$G_2(z)=\frac{0.036813(z-0.8606)(z^2-3.084z+3.924)}{(z^2-1.906z+0.971)(z^2-1.041z+0.9591)}$$	Stable	Normal	V_{ECG}/P_{ABP}: $$G_3(z)=\frac{-0.037107(z-3.235)(z^2-2.213z+1.226)}{(z^2-1.912z+0.9683)(z^2-1.769z+0.8672)}$$ V_{ECG}/P_{CVP}: $$G_4(z)=\frac{0.00089884(z-123)(z^2-2.202z+1.251)}{(z^2-1.913z+0.9693)(z^2-1.775z+0.8763)}$$	61 9850	Stable
				Normal	V_{ECG}/P_{ABP}: $$G_5(z)=\frac{3.603(z-0.2122)(z-0.9495)(z-1.098)}{(z^2-1.85z+0.9303)(z^2-1.078z+0.7482)}$$ V_{ECG}/P_{CVP}: $$G_6(z)=\frac{0.71334(z+0.764)(z^2-1.823z+1.16)}{(z^2-1.842z+0.9407)(z^2-1.172z+0.9127)}$$	79 2377	Stable

				Normal	V_{ECG}/P_{ABP} : $G_7(z) = \dfrac{0\,22381(z-0\,6185)(z^2-2\,31z+1\,335)}{(z^2-1\,32z+0\,4629)(z^2-1\,892z+0\,9597)}$ V_{ECG}/P_{CVP} : $G_8(z) = \dfrac{0\,11432(z-1\,083)(z^2-2\,512z+1\,754)}{(z^2-1\,904\,z+0\,9705)(z^2-1\,642z+0\,7967)}$	57 36	Stable
Mgh008	Normal	V_{ECG}/P_{ABP} : $G_1(z) = \dfrac{0\,8552(z+1\,178)(z^2-2\,063z+1\,067)}{(z-0\,922)(z-0\,615)(z^2-1\,683z+0\,853)}$ V_{ECG}/P_{CVP} : $G_2(z) = \dfrac{0\,59323(z-0\,9891)(z^2-2\,069z+1\,077)}{(z-0\,9621)(z-0\,9972)(z^2-1\,749z+0\,8761)}$	Stable	Normal	V_{ECG}/P_{ABP} : $G_3(z) = \dfrac{-1\,0605(z-0\,2858)(z^2-1\,942\,z+0\,9479)}{(z^2-1\,773z+0\,811)(z^2-1\,333z+0\,5957)}$ V_{ECG}/P_{CVP} : $G_4(z) = \dfrac{-0\,14312(z-2\,871)(z^2-2\,116z+1\,133)}{(z^2-1\,85z+0\,8685)(z^2-1\,364z+0\,5969)}$	76 2298	Stable
				Normal	V_{ECG}/P_{ABP} : $G_5(z) = \dfrac{1\,9247(z-1\,116)(z^2-1\,795z+0\,8072)}{(z^2-1\,315z+0\,4715)(z^2-1\,877z+0\,9261)}$ V_{ECG}/P_{CVP} : $G_6(z) = \dfrac{-1\,6658(z-1\,008)(z^2-2\,124z+1\,226)}{(z^2-1\,867z+0\,9177)(z^2-0\,8578z+0\,2639)}$	75 3023	Stable
				Normal	V_{ECG}/P_{ABP} : $G_7(z) = \dfrac{5\,322(z-1\,052)(z^2-1\,836z+0\,8575)}{(z^2-1\,851z+0\,8949)(z^2-1\,631z+0\,8194)}$ V_{ECG}/P_{CVP} : $G_8(z) = \dfrac{0\,43017(z-1\,12)(z^2-1\,801z+0\,9906)}{(z^2-1\,84z+0\,9007)(z^2-1\,622z+0\,8429)}$	80 4852	Stable
Mgh029	Normal	V_{ECG}/P_{ABP} : $G_1(z) = \dfrac{0\,00025613(z-1\,13)(z^2-1\,909z+0\,9239)}{(z^2-1\,966z+0\,9836)(z^2-1\,823z+0\,8909)}$ V_{ECG}/P_{CVP} : $G_2(z) = \dfrac{0\,63259(z-0\,9981)(z^2-2\,007z+1\,031)}{(z^2-1\,968z+0\,9848)(z^2-1\,9z+0\,9491)}$		Normal	V_{ECG}/P_{ABP} : $G_3(z) = \dfrac{0\,0014319(z-0\,8464)(z^2-2\,023z+1\,026)}{(z^2-1\,928z+0\,9383)(z^2-1\,919z+0\,9692)}$ V_{ECG}/P_{CVP} : $G_4(z) = \dfrac{0\,59618(z-1\,03)(z^2-1\,877z+0\,906)}{(z^2-1\,928z+0\,9383)(z^2-1\,919z+0\,9692)}$	78 3040	Stable

				Normal	V_{ECG}/P_{ABP}: $$G_5(z)=\frac{000016991(z-1271)(z-113)(z-09588)}{(z^2-1948z+09516)(z^2-1947z+09947)}$$ V_{ECG}/P_{CVP}: $$G_6(z)=\frac{044177(z-103)(z-09533)(z-07571)}{(z^2-1948z+09516)(z^2-1947z+09947)}$$	75 9092	Stable
			Stable	Normal	V_{ECG}/P_{ABP}: $$G_7(z)=\frac{000046518(z-09574)(z^2-2116z+113)}{(z^2-1833z+089)(z^2-1948z+09946)}$$ V_{ECG}/P_{CVP}: $$G_8(z)=\frac{049038(z-1017)(z-09086)(z-07299)}{(z^2-1833z+089)(z^2-1948z+09946)}$$	74 0555	Stable
Mgh031	Normal	V_{ECG}/P_{ABP}: $$G_1(z)=\frac{-0012627(z-1066)(z^2-2022z+1035)}{(z^2-1912z+09467)(z^2-1475z+06025)}$$ V_{ECG}/P_{CVP}: $$G_2(z)=\frac{-200837(z-1067)(z^2-2049z+108)}{(z^2-1919z+09557)(z^2-1591z+07069)}$$	Stable	Normal	V_{ECG}/P_{ABP}: $$G_3(z)=\frac{-55\times10^{-3}(z-1511)(z-1031)(z^2-1969z+09741)}{(z-09651)(z^2-1934z+09785)(z^2-1418z+06544)}$$ V_{ECG}/P_{CVP}: $$G_4(z)=\frac{-1855(z-1005)(z-09671)(z^2-1969z+1028)}{(z-09651)(z^2-1934z+09785)(z^2-1418z+06544)}$$	76 0562	Stable
				Normal	V_{ECG}/P_{ABP}: $$G_5(z)=\frac{-607\times10^{-3}(z^2-1998z+1002)(z^2-2223z+131)}{(z-09827)(z-06638)(z-05938)(z^2-1886z+09291)}$$ V_{ECG}/P_{CVP}: $$G_6(z)=\frac{-158(z-1008)(z-05866)(z^2-1896z+09103)}{(z-09827)(z-06638)(z-05938)(z^2-1886z+09291)}$$	65 1145	Stable
				Normal	V_{ECG}/P_{ABP}: $$G_7(z)=\frac{-504\times10^{-4}(z+2717)(z-09951)(z^2-1987z+09897)}{(z-09782)(z^2-149z+05598)(z^2-1931z+09787)}$$ V_{ECG}/P_{CVP}: $$G_8(z)=\frac{-11\times10^{-2}(z-1065)(z-1)(z^2-2028z+1417)}{(z-09782)(z^2-149z+05598)(z^2-1931z+09787)}$$	62 1158	Stable

| | | | | PVC (XIth cycle) | V_{ECG}/P_{ABP} : $$G_{21}(z) = \frac{617 \times 10^{-3}(z-0.8934)(z-0.194)(z^2-1.963z+0.9667)}{(z-0.416)(z^2-1.899z+0.9153)(z^2-1.989z+1.045)}$$ V_{ECG}/P_{CVP} : $$G_{22}(z) = \frac{4.57(z-1.02)(z-0.2414)(z^2-2.175z+1.185)}{(z-0.416)(z^2-1.899z+0.9153)(z^2-1.989z+1.045)}$$ | 42.1339 | Unstable |

Table 4.3 concluded

4.2.2. Stability analysis

Stability of a system implies that a small change in input does not result in large change in system behaviour. If a system is applied with bounded input and produces an unbounded response, the system is said to be unstable.

The normal flow of blood in the heart is maintained by the contraction of highly organized sequence of its four chambers. During this normal blood flow, electrical conduction system of heart is also normal. This is reflected as a normal ECG cycle. When any of the chambers of heart does not function properly or its organized sequence of contraction is disturbed such as there is block or obstruction in blood flow, normal sinus rhythm of ECG is also lost and the output pressure signals also reflect a large change in the behaviour of the heart. Therefore, electrocardiographic manifestations reveal the abnormal behaviour of the heart. These changes are also revealed in terms of variations in blood pressure. The obstruction of the blood flow affect the pumping action of the heart. Therefore, we can say that consistent behaviour of heart is also affected.

As stated in section 1.9.10(a-b) that an LTI system is stable if its output attains a steady state value if the system is subjected to an initial condition or disturbance. To analyze the stability, the poles of the transfer function must lie in the left half of s-plane or inside the unit circle in z-domain. As an example, if transfer function of a system with complex conjugate poles in z-plane is –

$$G(z) = \frac{-0.064772\,(z-0.9812)}{(z^2 - 1.91z + 0.9833)} \qquad (4.38)$$

This transfer function has complex conjugate poles i.e. (0.955+0.176i) and (0.955-0.176i) in z-domain. The above transfer function in s-domain is written as –

$$G(s) = \frac{-21.9132\,(s+6.801)}{(s^2 + 21.01s + 4427)} \qquad (4.39)$$

This transfer function has complex conjugate poles i.e. (-10.5+65.7i) and (-10.5 - 65.7i) in s-domain. The pole zero plots of G(z) and G(s) are shown in Figure 4.5 (a) and Figure 4.5(b) respectively. From Figure 4.5(a), it is clear that poles of z-domain transfer
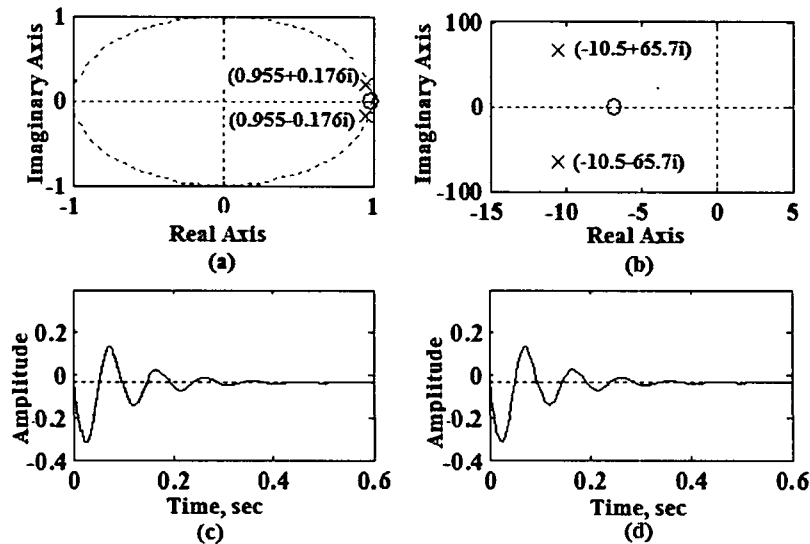
Figure 4.5 Pole zero plots of second order transfer function in (a) z-plane (b) s-plane with their corresponding step responses in (c) step response of Figure (a) and (d) step response of Figure (b)

function lie in right side of unit circle but are inside the unit circle and the poles of s-domain transfer function lie in lefthalf of s-plane indicating that both transfer functions represent the stable response. Any system can be tested for its stability by applying a step input and observing the response. An unstable system produces transients and oscillates while a stable system produces a stable output response. So we have also tested the transfer functions with step response to examine the steady state performance. The step response (Figures 4.5 (c-d)) of both transfer functions are same and represent a stable response.

We attempt to model the cardiac system of both healthy subjects (i.e. normal) and having PVC and SVPB (Supraventricular premature beat i.e. abnormal) and then we simulate the models using cross-validation inputs (ABP, CVP) i.e. normal to abnormal.

In this process, the model generated with normal data was validated with (a) model of normal data (b) model of abnormal data. Similarly, the model generated with abnormal data was validated with (c) model of abnormal data. It will be proved that in case of (a) the model will be stable and fit percentage will be obviously high, however in case of (b) although the fit percentage is high but the model is unstable.

Considering this fact, the generated transfer functions for test and validation data consisting of normal, PVC and SVPB beats mentioned in Table 4.3 for mgh004 record are studied using pole zero plot and step response as stated below –

a)   When both test and validation data are normal

First we form the models (G1, G2) with normal data and then validate the model with another set of models (G3, G4) of normal data. The pole zero and step response plots of the transfer functions are shown in Figure 4.6 (a-b) respectively. It is observed from Figure 4.6(a) that the poles of both model and validated transfer functions lie inside the unit circle stating that the estimated model is stable when both model and validation data are normal. It is also clear from step response plot of both transfer functions shown in Figure 4.6(b) that both model and validated transfer functions are stable.
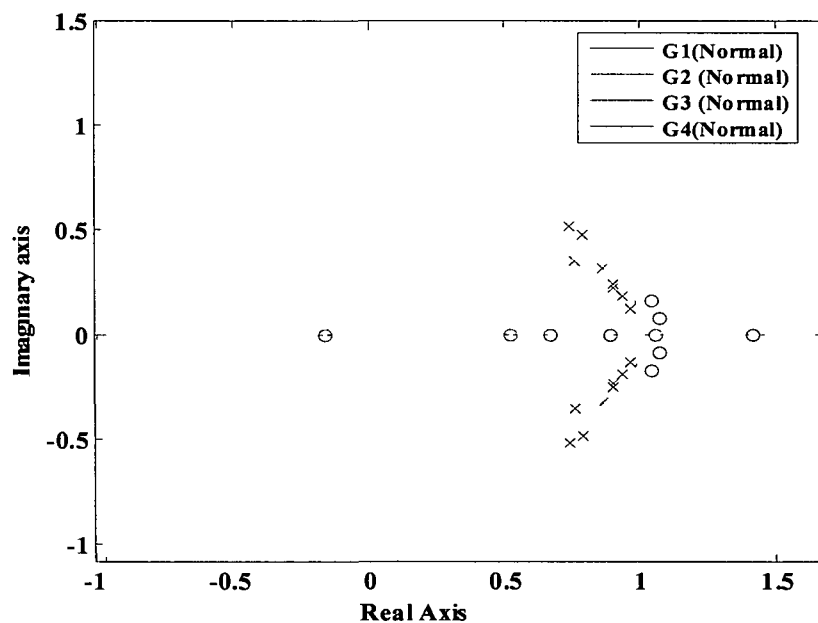

Figure 4.6(a) Pole zero plot of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh004 record
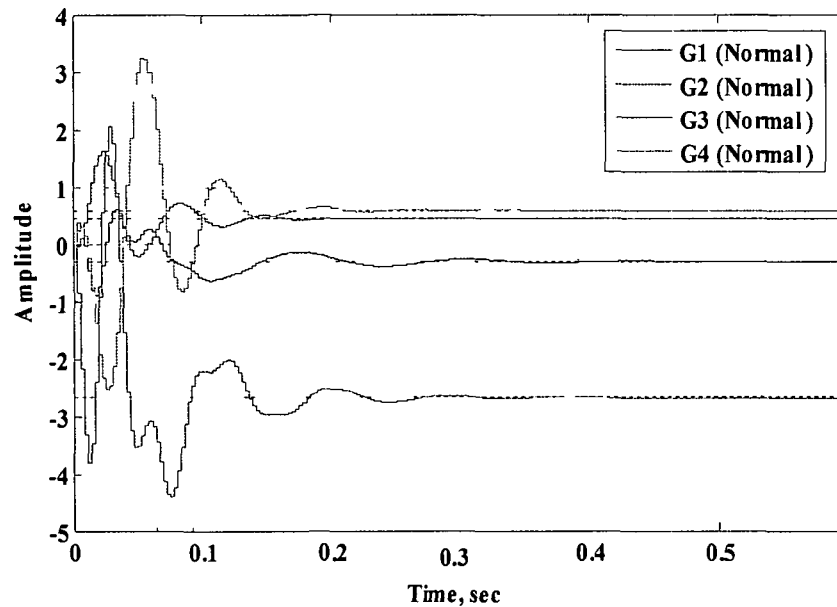
Figure 4.6(b) Step response of transfer functions - G1, G2 for normal model and G3, G4 of normal validation data of mgh004 record

b) When model data are normal and validation data consists of PVC and SVPB beats

When the model transfer function is developed with normal input-output data and validated using abnormal input (PVC), it generates two stable transfer functions (G1, G2) for normal test data and two transfer functions (G5 and G6) for validation data as PVC beat. The pole zero plots of transfer functions of normal model and validation data are shown in Figure 4.7 and step response are shown in Figure 4.8 (a-b). It is observed from Figure 4.7 that the system simulated with PVC beat comprise of two overlapping poles associated with each transfer function lying outside the unit circle which indicates that the system simulated with PVC beat is unstable. It is also clear from step response plot of transfer functions of test and validation data shown in Figure 4.8 (a-b) that the model output comes back to equilibrium state for test data whereas the system output is continuously decreasing and never attains a steady state value for PVC beat. It further confirms the fact that the estimated model for PVC beat is unstable.

Figure 4.7 Pole zero plot of transfer functions - G1, G2 for normal model and (G5, G6) of abnormal (PVC) validation data of mgh004 record

Pole zero and step response plots of obtained transfer functions (G7 and G8) in case of SVPB beats are also analyzed for stability. It is observed from the pole zero plots (Figure 4.9) and step response plots (Figure 4.8(c)) that estimated model for the ECG with SVPB beat is also unstable. Hence it can be concluded that modeling of normal beats of ECG using ABP and CVP signals gives a stable model and model estimated using abnormal data are unstable.
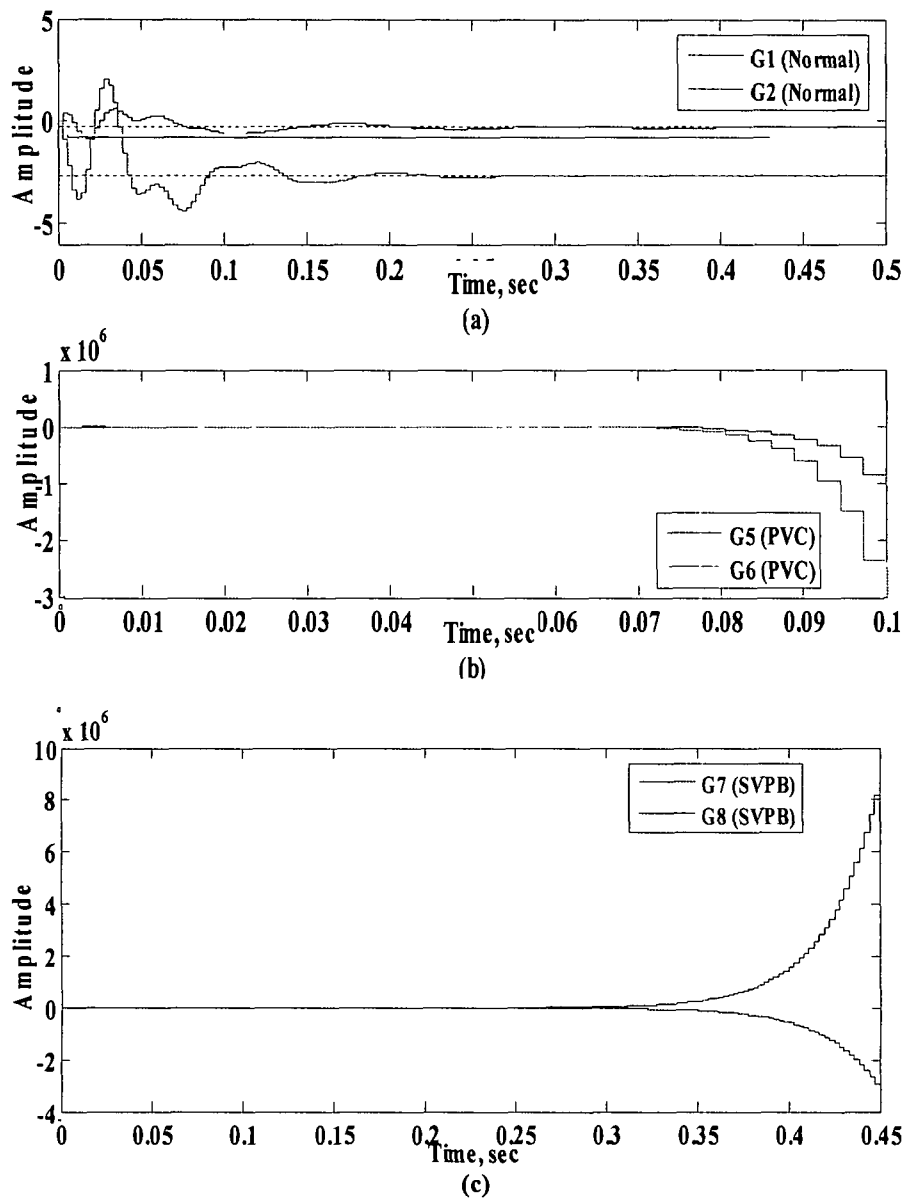
Figure 4.8 Step response plot of transfer functions for (a) normal model (G1, G2) and abnormal validation data (b) PVC (G5,G6) (c) SVPB (G7,G8) of mgh004 record
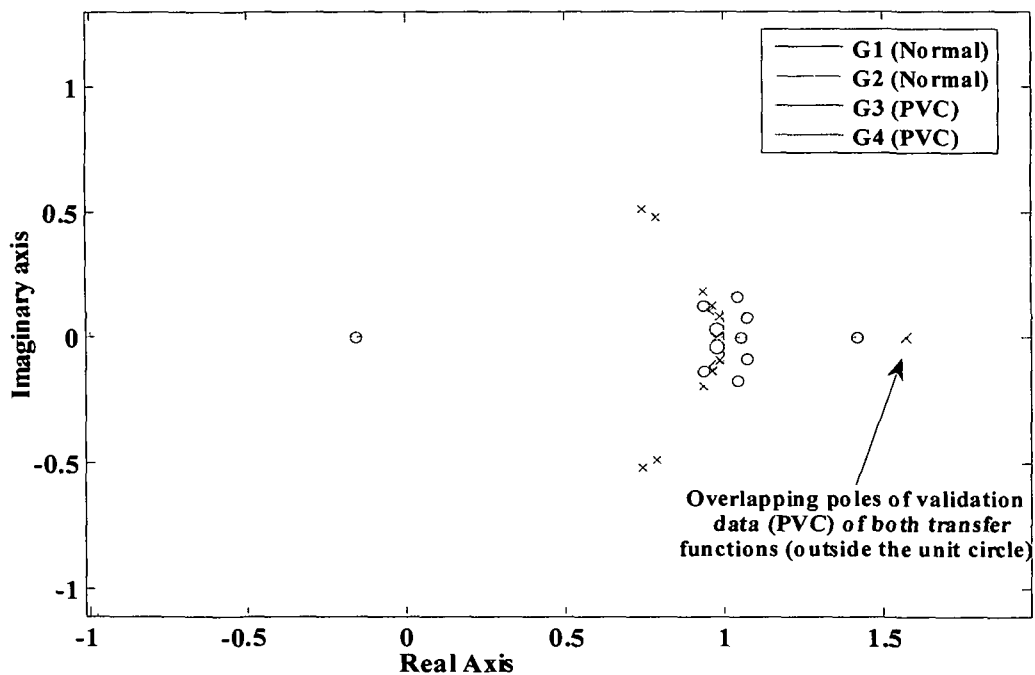
Figure 4.9 Pole zero plot of transfer functions - G1, G2 for normal model and (G7, G8) of abnormal (SVPB) validation data of mgh004 record

4.2.3. Comparative study of ECG modeling

For comparison of the models and identification techniques, we have used only normal beats of seven records – mgh003, mgh004, mgh005, mgh007, mgh008, mgh029, mgh031 records. We present here least square technique for autoregressive and PEM and subspace technique for state space models of ECG. The results for comparative study of these models for seven records are given in Table 4.4. It is found that state space model shows better modeling with higher fit percentage compared to autoregressive model. For example, in terms of highest fit percentage in mgh029, state space model shows 78.3040 %, while autoregressive model shows 62.0156 % of fit percentage. Similarly, for the minimum fit percentage condition in mgh008, state space model shows 76.2298% while autoregressive shows 1.9167% of fit percentage. ,

Similarly, if we compare the identification technique, subspace technique is found to be better than least square and PEM except for mgh005, in all records, subspace technique outperforms the least square and PEM technique. For example – for highest fit percentage in mgh029, the fit percentage is 78.3040 %, 74.5805 % and 62.0156 % for subspace, PEM

and least square respectively. Similarly, for minimum fit percentage case in mgh008, the fit percentage76.2298 %, 66.7466 % and 1.9167 % respectively for subspace, PEM and least square respectively. We conclude that subspace identification technique is better than PEM and least square technique.

The synthesized ECG cycles using normal data of mgh007 record for above mentioned identification techniques are shown in Figure 4.10(a-c). The pole zero and step response plots of generated transfer functions using autoregressive modeling and state space modeling using PEM and subspace methods are shown in Figures 4.11(a-b), 4.12(a-b) and 4.13(a-,b) respectively.

Table 4.4 Results of comparative study of ECG modeling by autoregressive model and state space models using PEM and Subspace method

Record No.	Type Model and validation data	Validation					
		Autoregressive Model (least squares)		State space Model (PEM method)		State space Model (Subspace method)	
		Best Fit (%)	Stability	Best Fit (%)	Stability	Best Fit (%)	Stability
Mgh003	Normal	21.2506	Stable	61.0302	Stable	63.6651	Stable
Mgh004	Normal	49.23	Stable	60.7404	Stable	67.2842	Stable
Mgh005	Normal	24.44	Stable	63.6515	Stable	60.95	Stable
Mgh007	Normal	53.4849	Stable	59.4684	Stable	61.9850	Stable
Mgh008	Normal	1.9167	Stable	66.7466	Stable	76.2298	Stable
Mgh029	Normal	62.0156	Stable	74.5805	Stable	78.3040	Stable
Mgh031	Normal	76.2032	Stable	72.7455	Stable	76.0562	Stable
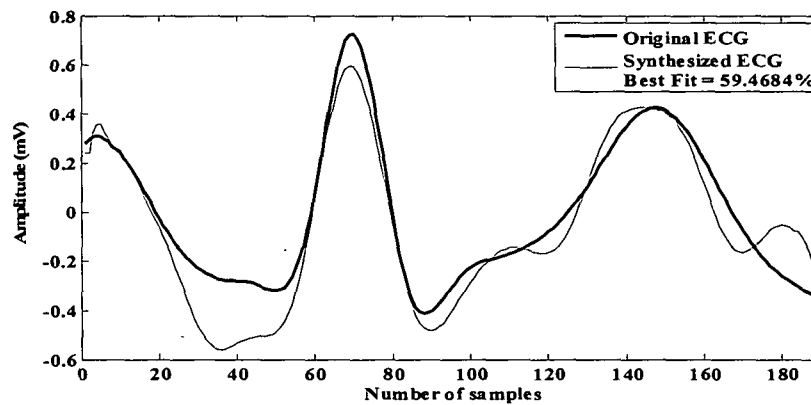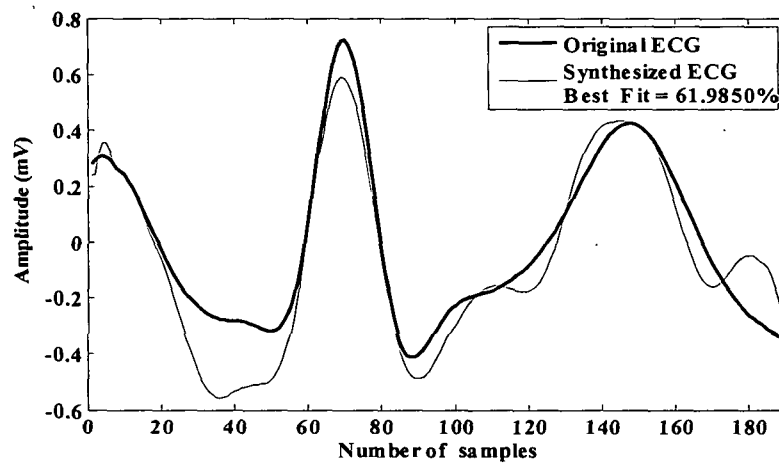
Figure 4.10 Original and synthesized ECG for mgh007 record for both normal model and validation data by (a) Least square (b) PEM and (c) subspace method
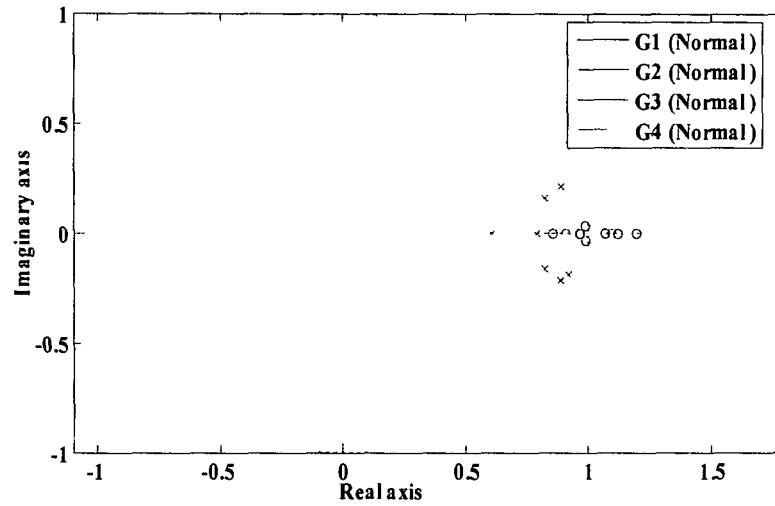
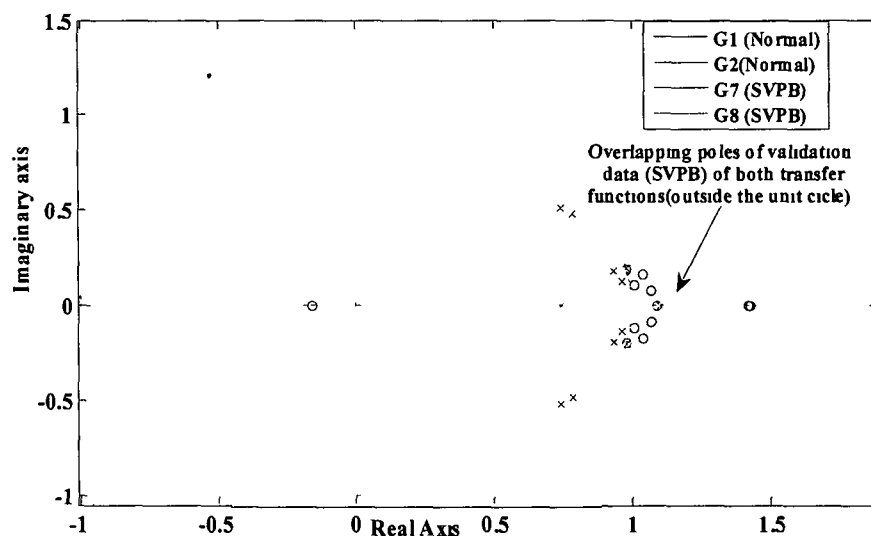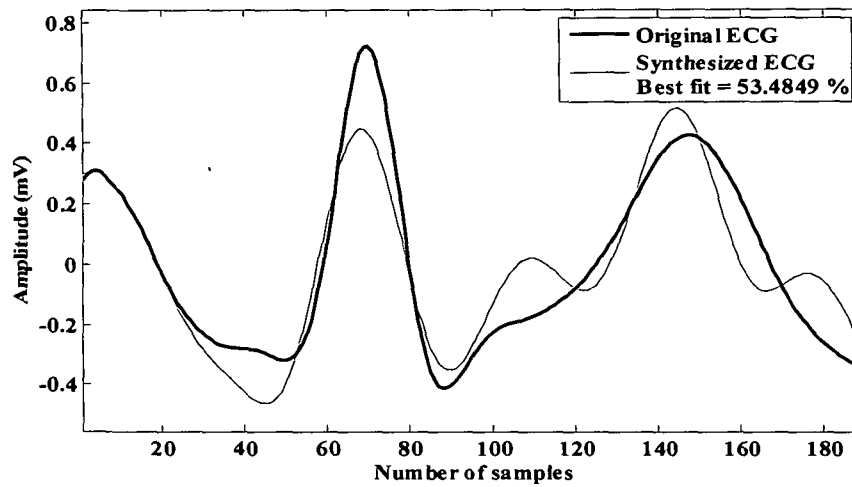Figure 4.11(a) Pole zero plot of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh007 record by least square method
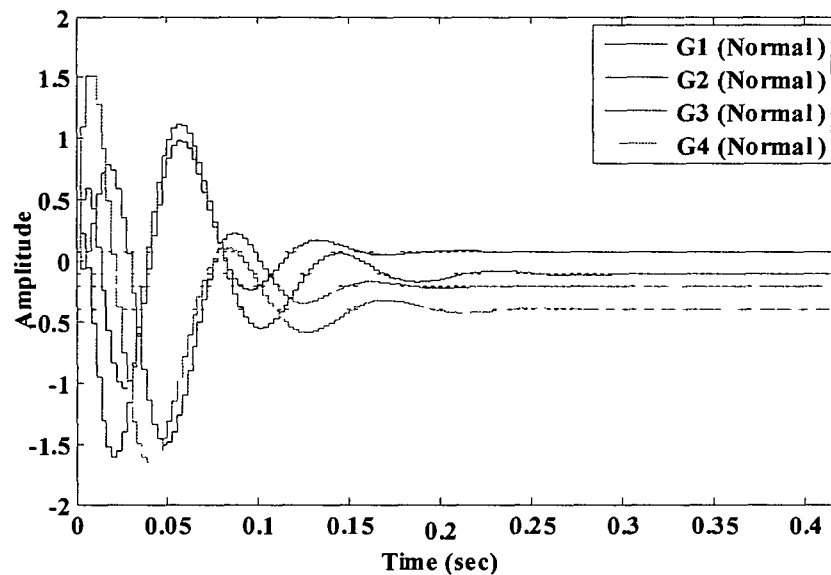
Figure 4.11(b) Step response of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh007 record by least square method
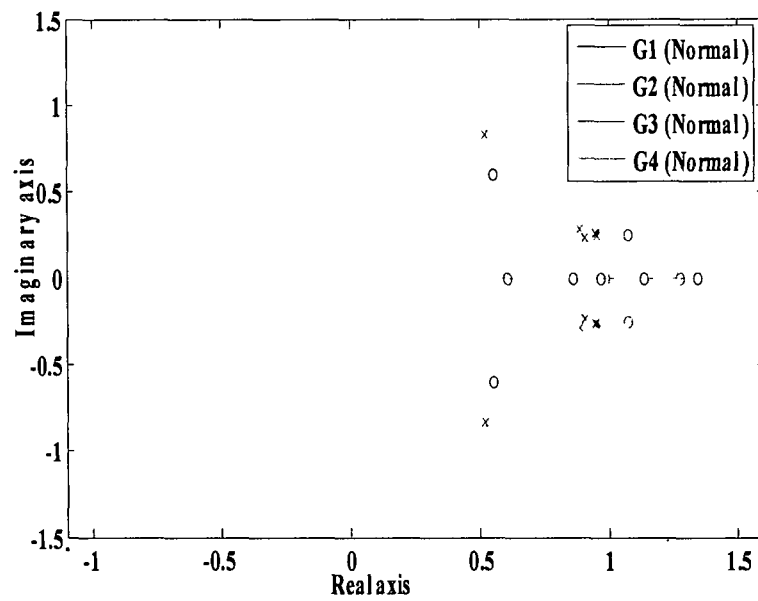
Figure 4.12(a) Pole zero plot of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh007 record by PEM method
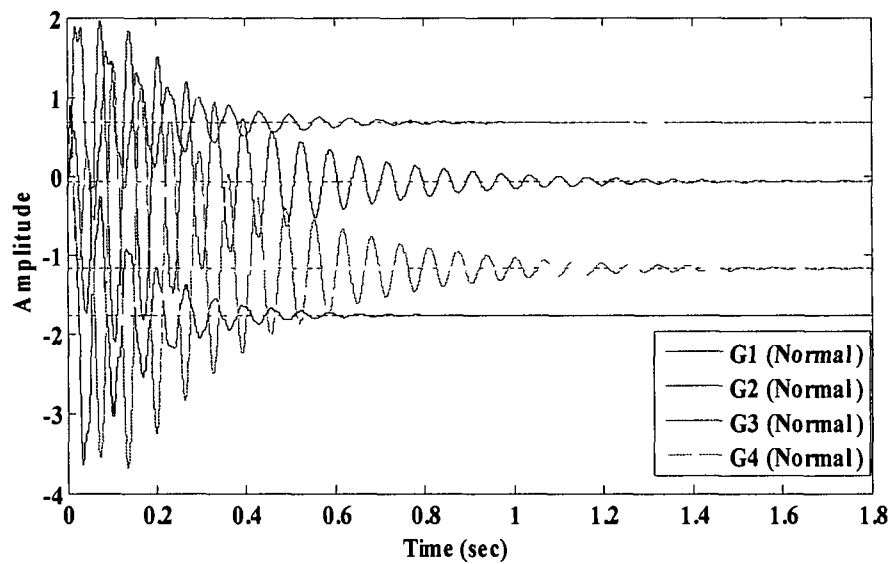
Figure 4.12(b) Step response of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh007 record by PEM method
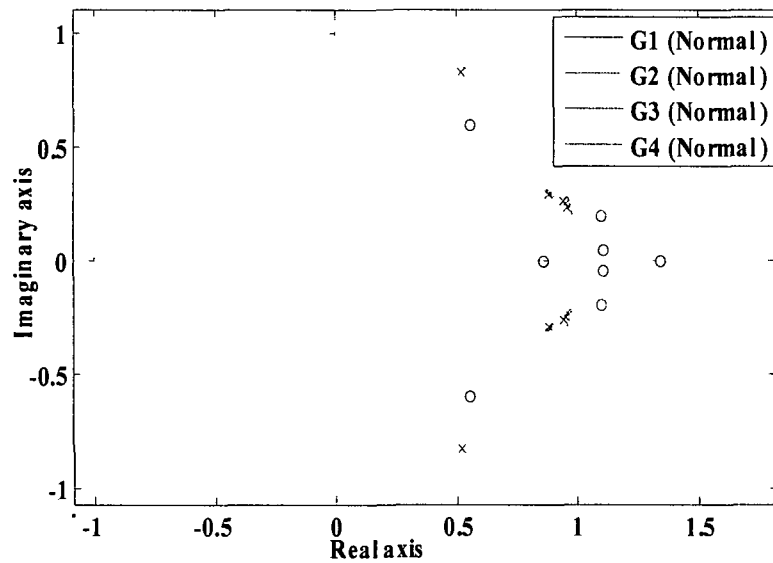
Figure 4.13(a) Pole zero plot of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh007 record by subspace method
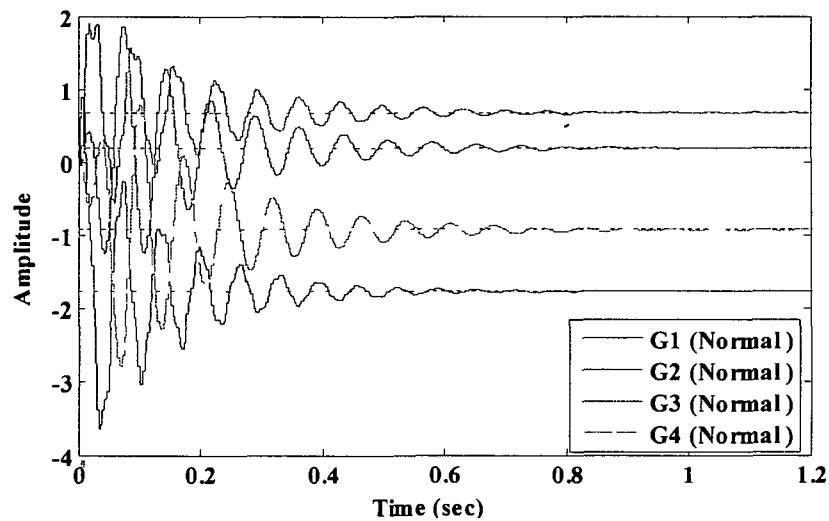
Figure 4.13(b) Step response of transfer functions - G1, G2 for normal model and (G3, G4) of normal validation data of mgh007 record by subspace method

4.3. Conclusion

A system identification based approach for ECG modeling and synthesis using the combination of normal ABP and CVP signals is discussed in this chapter. This approach of ECG modeling may be applied in intensive care units where ECG signals are monitored along with hemodynamic signals such as ABP and CVP signals and thus may enable monitoring of ECG without the positioning of leads for ECG acquisition. There may be a case when patient suffers from severe injury and surgical dressing of patient does not permit to position ECG acquisition leads at the desired places. This approach of modeling ECG may also be regarded as soft sensor acquisition of ECG. The generated transfer functions mentioned in Table 4.3 are also analyzed for stability analysis using pole zero and step response plots. It is concluded that the system generates a stable transfer function when model is simulated using normal data. On the other hand, system generates an unstable transfer function by simulating using abnormal (PVC and SVPB) beats. It is observed in terms of model stcuctures, state space model performs better than autoregressive models. Also, it is observed that subspace identification technique is better than PEM and least square technique.

Publication on this chapter

[1] Pachauri, A., & Bhuyan, M. Modeling of ECG using ABP and CVP signals : A system identification based approach, *International Journal of Engineering, Science and Innovative Technology*, 2(6), 321 -330, 2013.

Publication under review

[1] Pachauri, A., & Bhuyan, M. System identification based modeling and synthesis of electrocardiograms, *International Journal of Modeling and Simulation*.

CHAPTER 5

Modeling and Synthesis of ECG using Artificial Neural Network

5.0.   Introduction

The introduction to the basis and background of modelling and synthesis of ECG from ABP and CVP has been discussed in section 4.0. In this section, it was discussed that modeling of the cardiac system has been done by various researchers [113-131], while many have done synthesis [128] and reconstruction of ECG [96-99, 133-134]. It was discussed that modeling or reconstruction of ECG signal has been performed by various techniques such as wavelet based approach [134], ANN [96-99, 133] etc. In these techniques, ANN based reconstruction has shown promising results.

In physionet challenge 2010 [G B Moody], reconstruction of 3750 samples of biomedical signals using ANN [96-99, 133] and wavelet based approach [134] has been suggested. The signal reconstruction was followed only for last 3750 samples of the cardiovascular signals those appear as last segment in the signal to be reconstructed. For reconstruction of ECG signal, some authors have taken ECG signals from other two leads as input to ANN [98] while some have considered all the available signals such as respiration (RESP), fingertip plethysmogram outputs (PLETH) etc and along with ECG signals from two other leads [97,99].Thus, at least one input ECG signal from other lead is used for reconstruction of ECG signal [96-99, 133].

So far in the work [96-99, 133], authors have suggested ECG reconstruction using different approaches and ECG signal reconstruction using ECG from other leads [96-99, 133].

In all the above mentioned literature, generation of ECG is considered as 'reconstruction', since the aim is to generate missing ECGs of one of the leads. However, this research aims at synthesis of ECG from ABP and CVP without knowing a priori; the ECG pattern in any of the leads. The reason is that we intend to prove that there is obvious correlation between

ECG and pressure signals – ABP and CVP. In this chapter, we propose an ANN based approach for modeling and synthesis of ECG using ABP and CVP signals.

ANN is the most commonly used pattern recognition technique. There are many different types of ANN structures that have been applied to resolve complex problems such as multi-layered perceptron (MLP), radial basis function (RBF), learning vector quantization (LVQ) etc. Out of these ANN structures, RBF is capable of fast learning than back propagation networks, MLP etc. RBF networks are less susceptible to problems with non-stationary inputs and these networks have been successfully used in the prediction of time series data [110-111]. Our goal is to model and synthesize the ECG using ABP and CVP signals; hence RBF networks are better suited for our application.

The artificial neural network used for modeling and synthesis is radial basis network. The proposed method utilizes synchronously sampled ABP and CVP cycles of a patient for the synthesis of ECG cycles of that patient.

ECG synthesis using ABP and CVP signals as inputs by ANN modeling has been performed to validate that our synthesized ECG is faithfully conforms the actual ECG. To do that, we have adopted two validation procedures –

i)   Validation by peak detection

ii)   Validation by similarity measures

A total number of 16 records – mgh003, mgh004, mgh006, mgh007, mgh009, mgh011, mgh013, mgh015, mgh016, mgh020, mgh022, mgh025, mgh029, mgh032, mgh034 and mgh035 from MGH/MF waveform database are chosen for ECG modeling and synthesis. These records are chosen because they have 100000 samples length of each of simultaneously recorded ABP, CVP and ECG signals without any missing segment in this duration. All these signals are sampled at 360 samples/ sec and ECG signals are annotated by experts. As an example, details of record mgh003 are given below in Table 5.1 –

At first, the radial basis neural network is trained using large segments consisting of 10000 samples of ABP and CVP signals as input and ECG lead II signal as target signal for 16 records of the database mentioned above. Other segments of ABP and CVP signals of the same duration are used for testing the trained ANN. In this process, the trained ANN

Table 5.1. Description of record mgh003 of MGH/MF waveform database [139]

Record	Age/Sex	Signals	Gain	Base	Units
mgh003	47/ Female	ECG lead II	1170	-163	mV
		ABP	12.08	-1227	mm Hg
		CVP	20.96	-992	mm Hg

synthesizes ECG which has clear ECG peaks whereas other components of ECG waveform such as P and T waves are not clear for certain records. The number of peaks available in original ECG signal from database and ECG signal synthesized from ANN output are compared to determine accuracy, sensitivity and positive predictivity of synthesized ECG peaks.

In the second approach, a total number of 100000 samples of ABP, CVP and ECG signals of each of the 16 records mentioned above are considered for modeling and synthesis of ECG. The length of each signal is divided into 40 datasets comprising of 2500 samples of each signal. First 34 segments of ABP and CVP signals are given as input to radial basis network and corresponding ECG signals as target signals. The trained ANN outputs synthesized ECG while testing with remaining 06 segments of ABP and CVP signals which is compared with actual ECG signal available from the database. The synthesized ECG signals possess resemblance with actual ECG signals available from the database. The accuracy of this synthesized ECG is given in terms of cosine measure and cross correlation coefficient with respect to original ECG.

5.1.   Synthesis of ECG and peak detection

A radial basis network is used to synthesize ECG from ABP and CVP signals. The network architecture is the combination of an input layer, a hidden layer fully connected with the input layer and an output layer.

In general hidden layer provides the network its ability to generalize. Theoretically, a neural network with one hidden layer comprising of sufficient number of hidden neurons is capable to approximate any continuous function. In practice, neural networks having two hidden layers are widely used and have performed well. Increase in the number of hidden layers also increase the computation time and may lead to over fitting and results in poor

prediction of output performance. The ANN based algorithm suggested here utilizes one hidden layer.

There is no way to select the optimum number of hidden neurons. The choice of desired number of hidden neurons depends on the experiment being performed. Generally, more number of neurons is given to ensure better training. The number of output neuron is one in this case. Schematic illustration of used network architecture is shown in Figure 5.1. Parameters of radial basis function network used for training are shown in Table 5.2. Flow chart of ANN based ECG peak detection algorithm is shown in Figure 5.2. In this part, synthesis of ECG and peak detection from ABP and CVP signals is accomplished in the following steps –
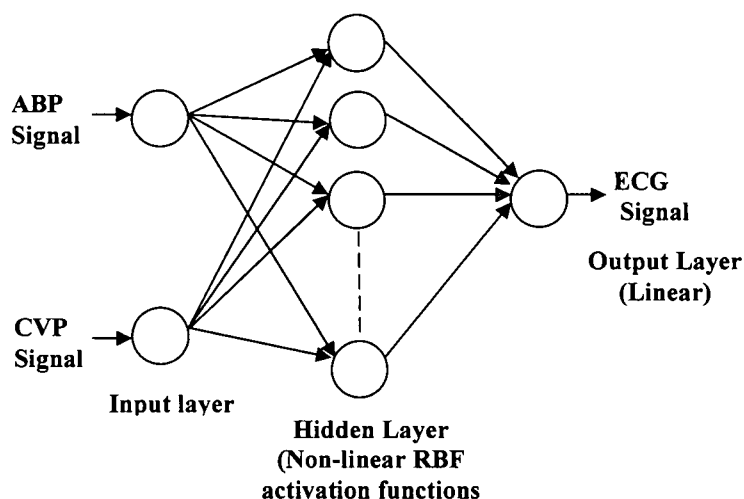
Figure 5.1 Architecture of radial basis network

5.1.1. Data preprocessing

Data preprocessing or normalization refers to analyzing and transforming the input and output signals to minimize the noise, highlight the important relationships, detect trends and flatten the distribution of the variable to assist the neural network in learning the relevant patterns. At the very least, the raw data must be normalized between the upper and lower bounds usually between -1 to +1. The signals under study are pressure (ABP, CVP) signals and ECG II signals those have different noise properties. ABP, CVP signals are chosen as input to the neural network and ECG as target. An example of normalized input (ABP, CVP) signals and target (ECG) signals are shown in Figure 5.3. A total number of 20000 samples of ABP, CVP and ECG signals are considered for the purpose of synthesis

of ECG and peak detection. The network is trained with 10000 samples and remaining 10000 samples of ABP and CVP signals are given as input to test the trained ANN.

Table 5.2. Parameters of proposed RBF network for peak detection in synthesized ECG

Number of input neuron	Number of output neuron	Number of hidden layers	Performance function	Ephos	Goal	Spread constant
02	01	01	MSE	250	10^{-7}	110

5.1.2. Training and testing

ANN training is generally accomplished by using transfer functions which are also called activation or threshold functions. These activation functions are mathematical formulae that decide the output of a neuron. Most of the neural networks use sigmoid (S-shaped) function while some use step, ramping etc functions. These activation functions limit the output of neural network to attain a very high value which can paralyze neural network and thus inhibit training. We have used $f(n) = e^{-n^2}$ as the activation function which is Gaussian function. In radial basis function networks, the activation of hidden layer is determined by the distance between input vector and a vector using non-linear radial basis function. The maximum number of neurons to be added in the hidden layer is equal to the dimension of input vector. Choice of spread constant plays a very important role in the training and testing of the radial basis network. If the value of spread constant is large, the function approximation is smooth and more numbers of neurons are required to fit a fast changing function. If the value of spread is very small, more numbers of neurons are required for fit a smooth changing function.

5.1.3. Performance function

The performance of ANN training is judged by calculating mean square error (MSE). The least mean square error (LMS) algorithm is an example of supervised training, in which the learning rule is provided with a set of examples of desired network behaviour such as –

$$\{p_1, t_1\}, \{p_2, t_2\}, \cdots \cdots \cdots \cdots \cdots \cdots \{p_q, t_q\},$$

$$(5.1)$$

Here, p_q is an input to the network, and t_q is the corresponding target output. As inputs are applied to the network, the network output is compared with the target iteratively. The error is calculated as the difference between the target output $y(k)$ and the network output $\hat{y}(k)$ and MSE is evaluated as –

$$mse = \frac{1}{Q}\sum_{k=1}^{Q}e(k)^2 = \frac{1}{Q}\sum_{k=1}^{Q}(y(k) - \hat{y}(k))^2$$

(5.2)

Where, ' Q ' denotes the number of predictions.
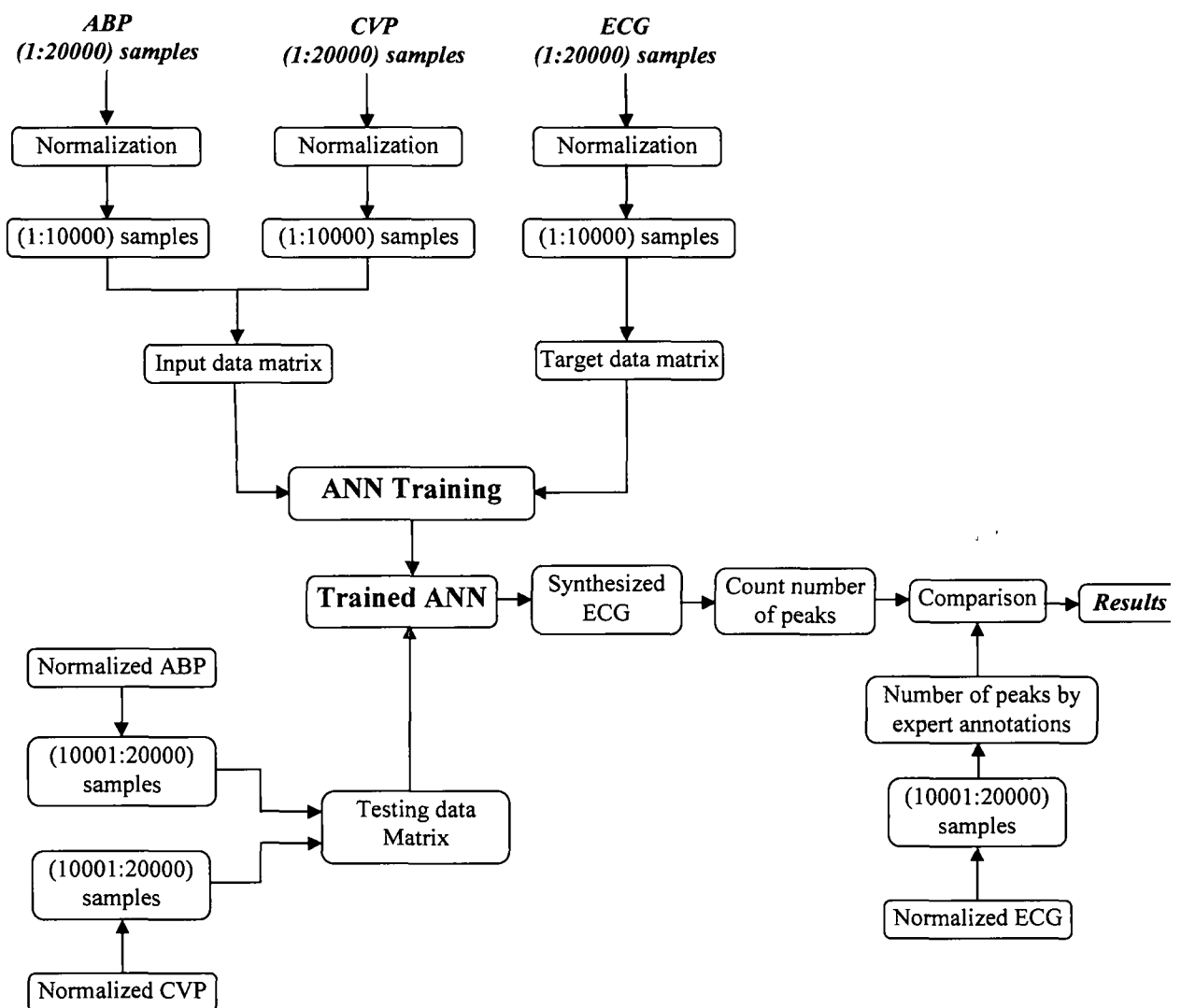
Figure 5.2 Flow chart of ANN based ECG synthesis and peak detection algorithm

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques
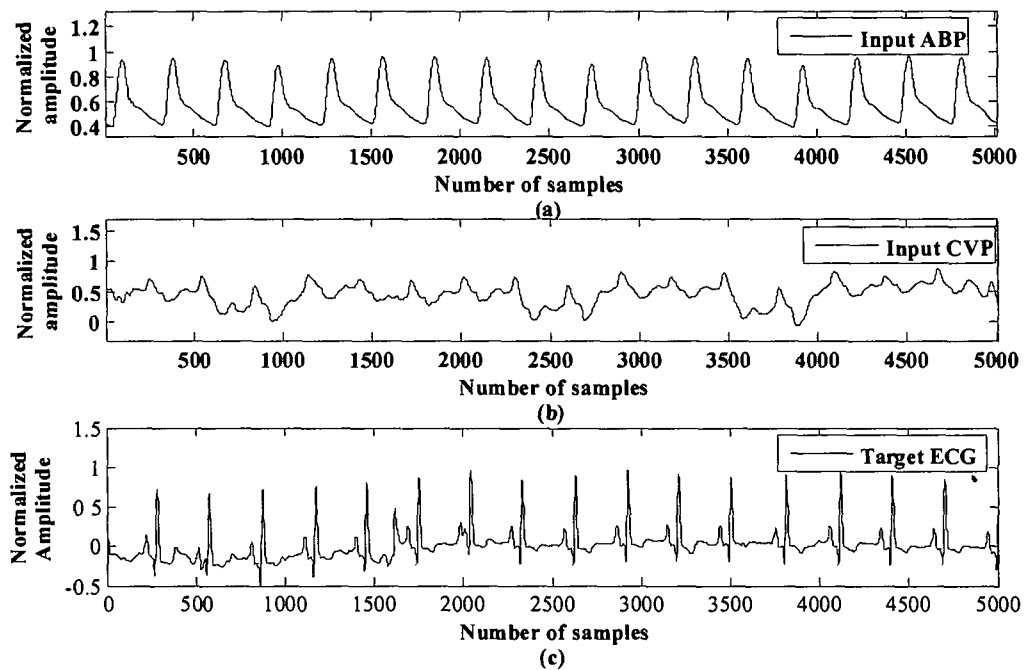
Figure 5.3 An example of input ABP, CVP and target ECG signals of mgh011 record used for training ANN

The goal is to minimize the average of the sum of these errors. The training algorithm adjusts the weights and biases of the network so as to minimize this mean square error. Fortunately, the mean square error performance index for the linear network is a quadratic function. Thus, the performance index will either have one global minimum, a weak minimum, or no minimum, depending on the characteristics of the input vectors. Specifically, the characteristics of the input vectors determine whether or not a unique solution exists.

In our case, we have chosen spread constant value of 110 by observing the ANN performance using different spread constants. The number of epochs given for training is 250. The goal for training is 0.0000001 but the network is trained by 22 epochs with a performance value of 1.115337×10^{-28} in 31.476068 seconds. The training performance plot of ANN is shown in Figure 5.4.
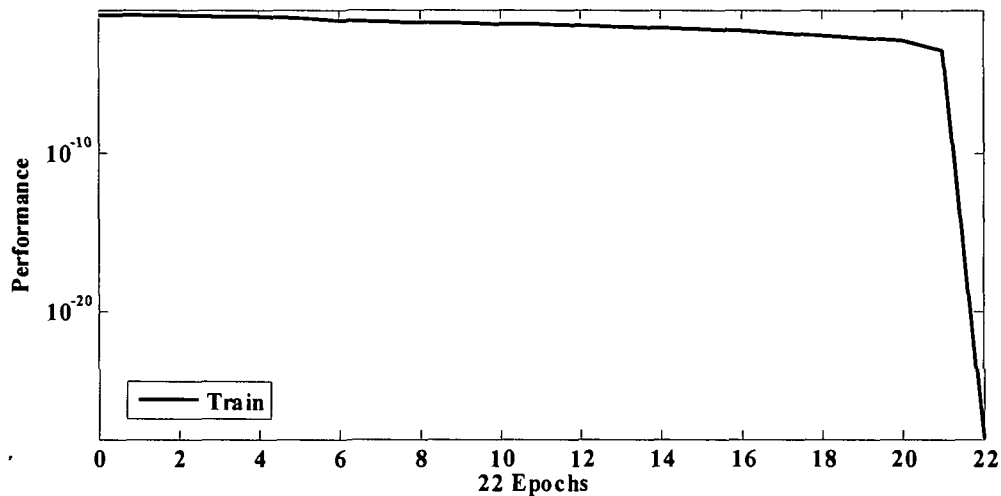
Figure 5.4 Training performance of ANN for ECG synthesis and peak detection algorithm

5.1.4. Results and discussion

The synthesized ECG signals obtained by this method has prominent R peaks whereas other components of ECG waveform such as P and T waves are not clear as shown in Figure 5.5 (b). Also the amplitude of synthesized ECG signal is varied from the original signal amplitude. The number of peaks obtained in the predicted ECG are compared with expert annotations for ECG peaks available from the database. The accuracy, sensitivity and positive predictivity are calculated using the equations 2.14, 2.15 and 2.16 respectively. The results of peak detection on synthesized ECG are presented in Table 5.3. The overall accuracy, sensitivity and positive predictivity for 16 records shown in Table 5.3 are 95.96%, 97.05% and 98.99% respectively. Original and synthesized ECG signals for record mgh011 and mgh09 records are shown in Figure 5.5 (a-b) and Figure 5.6 (a-b) respectively.
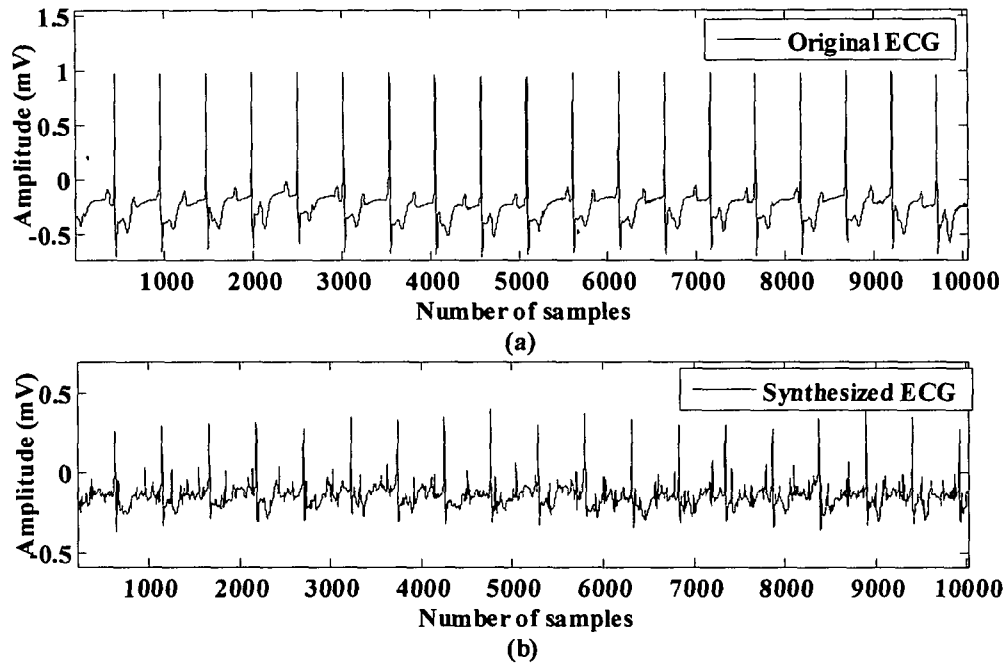
Figure 5.5 (a) Original and (b) synthesized ECG signals of mgh011 record
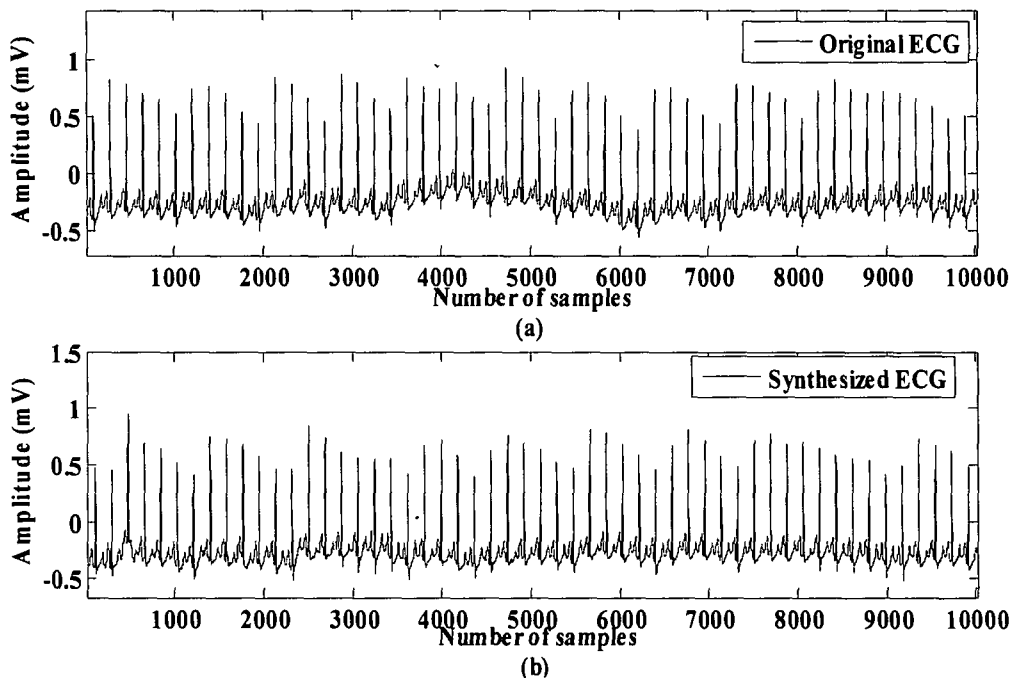
Figure 5.6 (a) Original and (b) synthesized ECG signals of mgh009 record

Table 5.3 Results of ECG peak detection of synthesized ECG (10,000 samples)

Record No	Age/Sex	Expert annotations	Peaks counted		
			TP*	FP*	FN*.
Mgh003	47 F	60	60	01	-
Mgh004	64 F	34	33	-	01
Mgh006	72 M	48	46	-	02
Mgh007	60 F	44	44	-	-
Mgh009	56 M	54	54	-	-
Mgh011	70 F	19	19	-	-
Mgh013	73 M	35	33	-	02
Mgh015	73 M	43	43	01	-
Mgh016	82 M	26	25	-	01
Mgh020	60 F	38	38	02	-
Mgh022	74 M	40	40	01	-
Mgh025	69 M	50	40	-	10
Mgh029	84 M	37	36	-	01
Mgh032	78 F	38	38	02	-
Mgh034	66 M	55	54	-	01
Mgh035	68 M	23	22	-	01
Total		644	625	07	19

*Where, TP, FP and FN denote true positives, false positives and false negatives as discussed in section 2.3.1.

5.2. Synthesis of ECG and similarity measures

In this part of ECG synthesis, same radial basis function network as shown in Figure 5.1 is used to synthesize ECG from ABP and CVP signals by segmenting the input and target signals. The aim is to test the synthesized ECG by using two similarity measures. The parameters of proposed RBF network are listed in Table 5.4 and the flow chart of ECG synthesis for calculation of similarity measures is shown in Figure 5.7.

Table 5.4. Parameters of proposed RBF network for ECG synthesis and similarity measures

Number of input neuron	Number of output neuron	Number of hidden layers	Performance function	Ephos	Goal	Spread constant
02	01	01	MSE	500	10^{-7}	10

5.2.1. Preprocessing

The input and target signals of all the 16 records mentioned above are segmented into 40 datasets after data normalization. The network is trained with 34 datasets and remaining 6 datasets from each record are considered for the purpose of testing the model. More

number of datasets are used to ensure better training. Each input dataset consist of 2500 samples of ABP signal, 2500 samples of CVP signal and 2500 samples of ECG II signal are taken as target. First one segment of input and target signals of mgh007 record used for training ANN is shown in Figure 5.8.
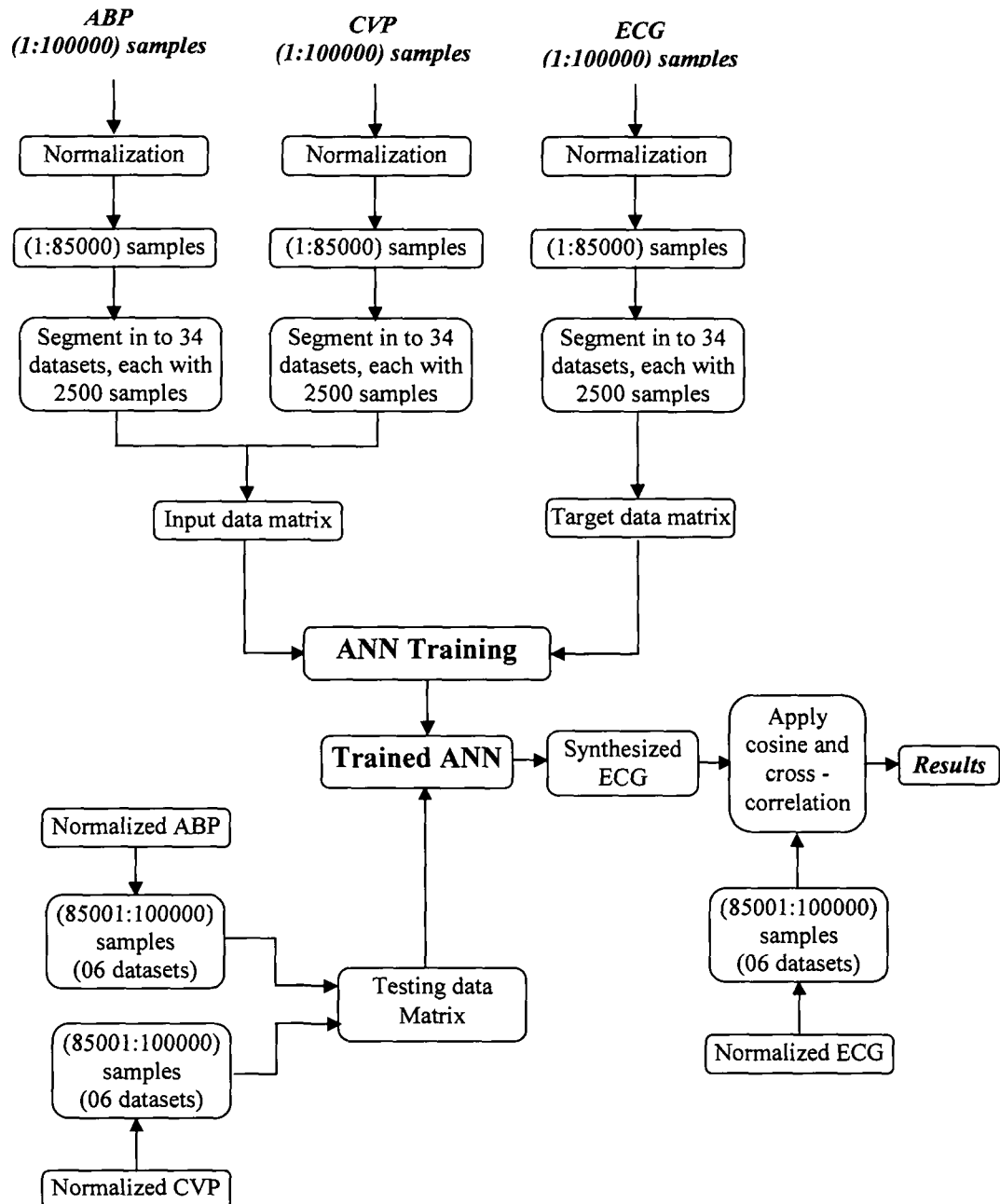
Figure 5.7 Flow chart of ANN based ECG synthesis and similarity measures algorithm

---

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques

The mean square error given for training is 0.0000001 and the obtained value for MSE are 0.00716627. A total of 500 epochs are given for training with a spread constant value of 10. The training is completed in 5 hrs 41 min 36.24 seconds. The training performance plot of ANN is shown in Figure 5.9.
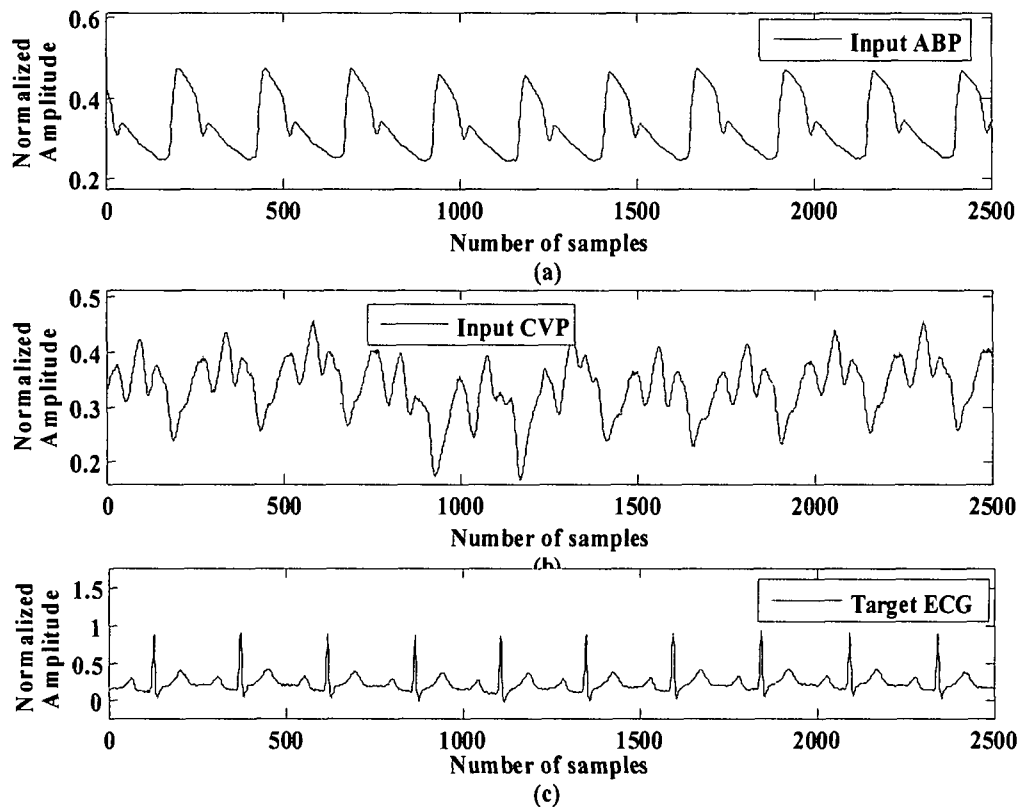
Figure 5.8 An example of input ABP, CVP and target ECG signals of mgh007 record used for training ANN

5.2.2. Similarity measures

The accuracy of the synthesized ECG with actual ECG is evaluated given in terms of cosine measure and cross-correlation of predicted ECG with respect to actual ECG signals for 16 patients available from the database.

5.2.2.1. Cosine measure

Cosine angle distance is used for similarity measurement of signal records in time domain. The distance between two vectors X and Y is measured by cosine of angle 'θ' between them. To make quantitative measurement between the actual ECG and predicted ECG

signals, cosine measure [164] between actual and predicted ECG signals for each of 16 records is computed. The similarity using cosine measure between actual ECG (A_{ECG}) and predicted ECG (P_{ECG}) for each record is given by the following relation –

$$sim\ (A_{ECG}, P_{ECG}) = \frac{\vec{A}_{ECG} \cdot \vec{P}_{ECG}}{\sqrt{\sum_{i=1}^{n} (A_{ECG})^2}\ \sqrt{\sum_{i=1}^{n} (P_{ECG})^2}}$$

(5.3)

Here, n=15000,. The larger the value of $sim(A_{ECG}, P_{ECG})$ in above equation, more the similarity between actual and predicted ECG signals. The accuracies of synthesized ECG using radial basis network in terms of cosine measure between two signals are given in Table 5.5. It is clear from Table 5.5 that predicted ECG signal for mgh006 has the highest value of cosine measure.
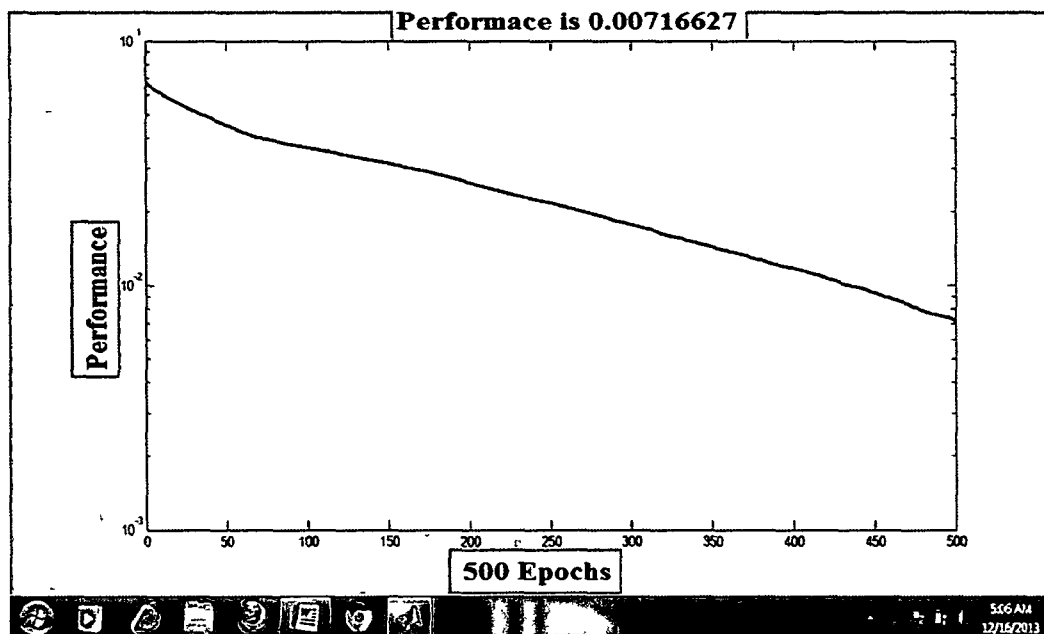
Figure 5.9 Training performance of ANN used for ECG synthesis after segmenting the input and target signals

5.2.2.2. Cross-correlation

As another method for validation of synthesized ECG with original ECG signals, the cross-correlations between the predicted and actual ECG signals for 16 records from the database are computed. It is stated in section 1.9.4, that the correlation coefficient takes on values

ranging between -1 and +1 which denote the perfect negative and perfect positive linear relationship between two signals.

The value of cross-correlation coefficient of the predicted ECGs with respect to actual ECG signals for 16 records available from the database is given in Table 5.5 along with the values of cosine measures. It is further confirmed wherever cosine measure is high, the value of cross-correlation coefficient is also high in most of the cases or in some cases these are either equal (mgh015, mgh025) or approximately equal (mgh006, mgh022). Several combinations of training parameters such as number of epochs, spread constant, are employed for training the RBF network. The better results are obtained using number of epochs = 500, spread constant = 10 which are shown in Table 5.5. The actual and predicted ECG signals for records mgh003, mgh007 and mgh025 are given in Figure 5.10, 5.11 and 5.12 respectively.

Table 5.5 Results of ECG synthesis using radial basis network

Record No.	Age/Sex	Cosine Measure	Cross-correlation
Mgh003	47 F	0.8391	0.8408
Mgh004	64 F	0.8855	0.9562
Mgh006	72 M	0.9493	0.9498
Mgh007	60 F	0.5560	0.6108
Mgh009	56 M	0.8682	0.8789
Mgh011	70 F	0.7814	0.8409
Mgh013	73 M	0.6614	0.7530
Mgh015	73 M	0.7294	0.7294
Mgh016	82 M	0.5005	0.7824
Mgh020	60 F	0.7824	0.7863
Mgh022	74 M	0.8050	0.8090
Mgh025	69 M	0.9021	0.9021
Mgh029	84 M	0.4854	0.5542
Mgh032	78 F	0.6215	0.6245
Mgh034	66 M	0.7152	0.7182
Mgh035	68 M	0.3250	0.5223
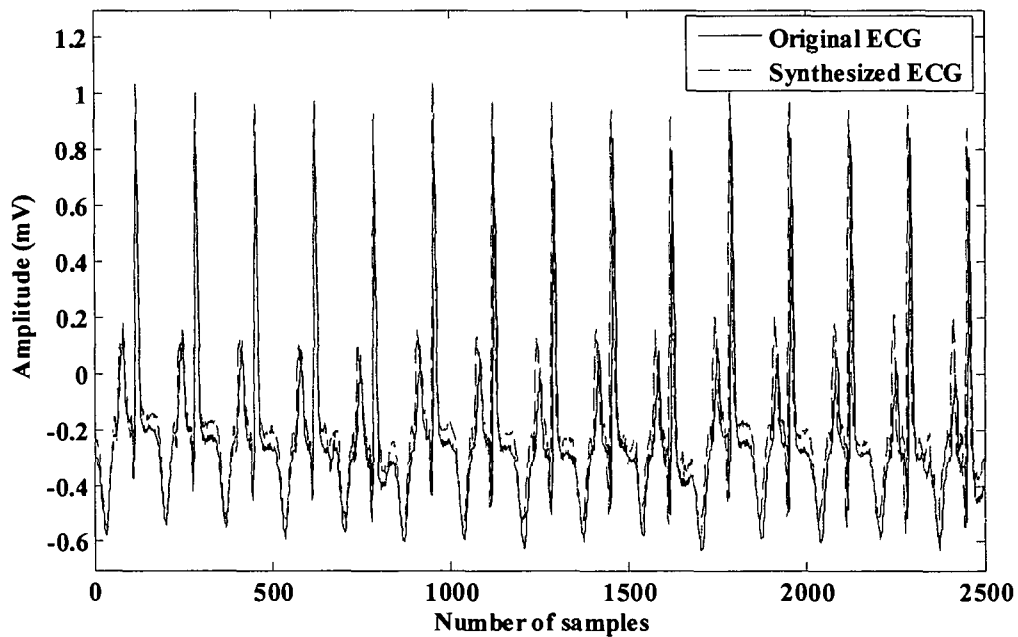
Figure 5.10 Original and synthesized ECG signals for mgh003 record
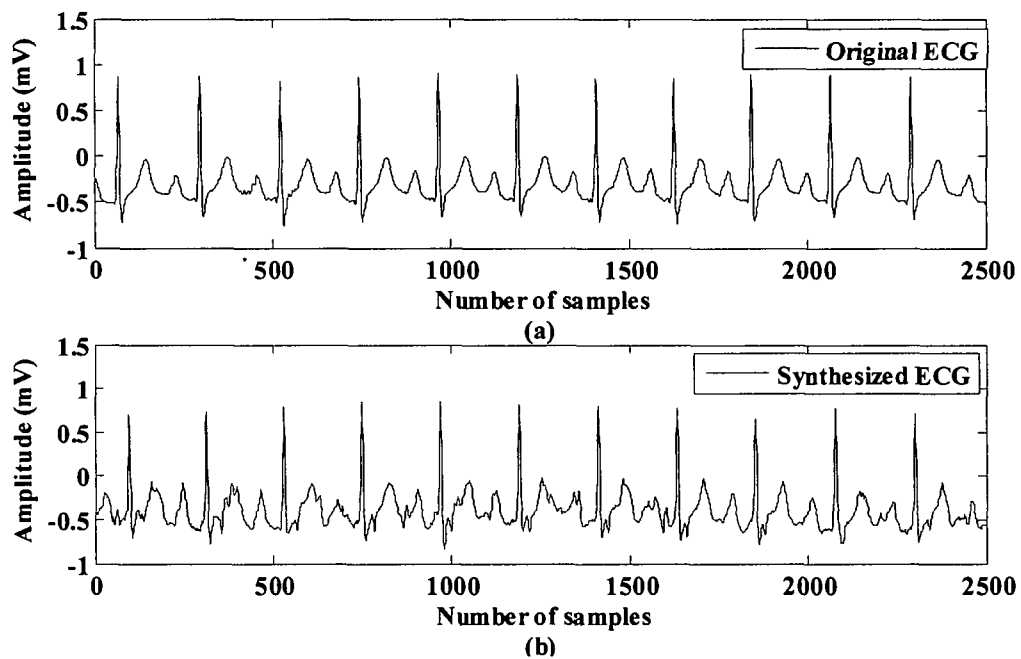
Figure 5.11 (a) Original and (b) synthesized ECG signals for mgh007 record

Feature Extraction, Modeling and Synthesis of ECG from Arterial Blood Pressure and Central Venous Pressure Signals by Signal Processing Techniques
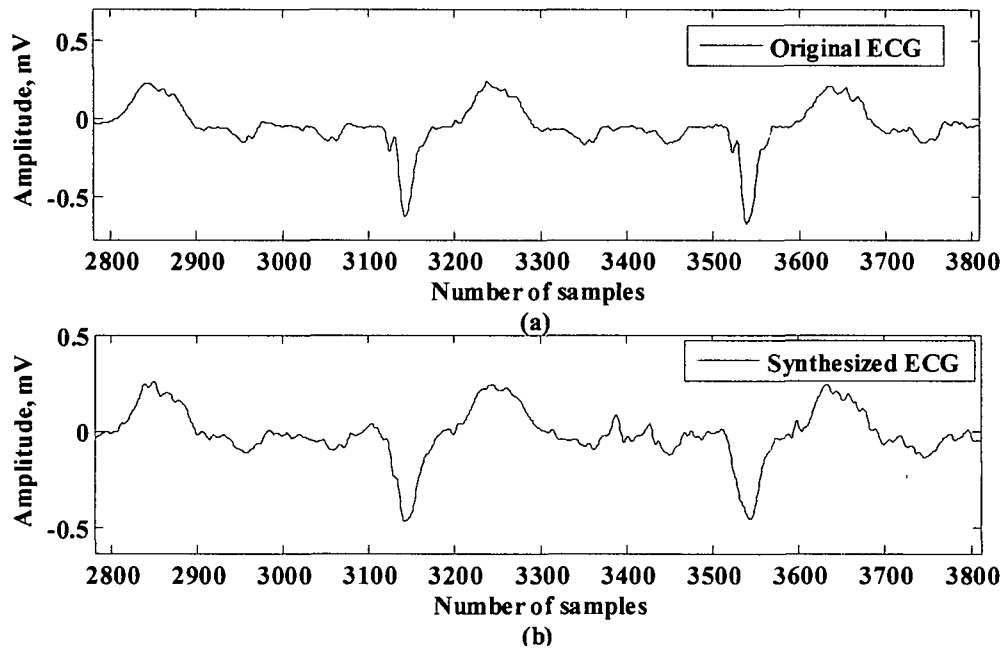
Figure 5.12 (a) Original and (b) synthesized ECG signals for mgh025 record

5.3. Conclusion and discussion

An algorithm for modeling and synthesis of ECG using the combination of ABP and CVP is developed. The algorithm can be used for modeling of other invasively acquired biomedical signals and even for the signals comprising of random fluctuations. So far authors have discussed modeling of ECG with different approaches but for patient specific reconstruction of ECG signals, ECG signals from other leads are used as input to ANN [96-99, 133]. Our algorithm highlights the importance of relevant features of ABP and CVP with that of ECG signals. The radial basis network trained with fewer neurons outputs ECG signals with clear ECG peaks only that can be employed for estimation of heart rate. The algorithm performance is further enhanced with the segments of inputs and target signals used for training. This trained ANN results in clear ECG signals. In general, better accuracy of predicted ECG signal was obtained if noise free ABP and CVP signals are input to ANN for validation. This method may be useful to derive ECG signals in critical care units when surgical dressing of patients may not allow placing the required leads at the desired position. In our case, we have obtained better results with radial basis network. However, other approaches such as genetic algorithm may also be applied to obtain further better results.

Publication on this chapter

[1] Pachauri, A., & Bhuyan, M. Modeling of ECG from arterial blood pressure and central venous pressure using artificial neural network, in IEEE International conference on recent advances and innovations in engineering (ICRAIE' 2014), Jaipur, India.

CHAPTER

Conclusion and Future Scope

6

6.0. Introduction

The prevalence of chronic diseases such as heart disease and stroke are the major cause of death in almost all countries, and such cardiovascular diseases present challenging problems in early diagnosis and subsequent treatment. Engineering has played a vital role in understanding the cardiac electrodynamics and in the development of diagnostic instruments. However, there is a requirement for a comprehensive computational and systematic approach to understand the cardiovascular systems for applications in the diagnosis [3]. It has already been mentioned that in intensive care units, where, continuously monitoring of physiological signals is required, there are various cardiac conditions and features which need to be detected on a large data of signal record and validation of abnormal ECG alarms using other cardiac signals. There may be instances when actual ECG signal is missed or corrupted due to errors in sensors or due to external interruption. These interruptions in actual signals result in a great difficulty for precise diagnosis. At times, acquisition of ECG may not be possible due to surgical dressing of patients. So the synthesis of ECG using ABP and CVP signals can be used to supplement the information when actual ECG is either missing or corrupted. Also, a single lead ECG information need to be expanded to multilead information using the information derived from other cardiovascular signals such as ABP, CVP, PAP etc.

The problems faced by cardiologists such as detection of features in ECG and the pressure signals (ABP and CVP), synthesis of ECG using ABP and CVP signals when actual ECG is either missing or corrupted have emerged as a great motivation to research in engineering. The objectives of this research have been achieved with satisfactory results as discussed in the following sections –

6.1. Feature extraction of ECG

We have proposed ECG peak detection techniques by two methods - wavelet based and energy analysis of ECG signal. The wavelet based approach is robust and simple to implement with no requirement of preprocessing. The selection of detail wavelet component has been justified by energy, frequency and correlation analysis. Since, there are wide variations in amplitudes of wavelet decomposed signals; a fixed threshold does not work for R peak detection. Therefore, we have adopted a 'window based threshold' where the threshold value is adjusted depending upon the signal amplitude over a certain duration. The selected detail signal is first thresholded then the maximum amplitude levels of all the peaks are detected. The signal is then filtered by applying a refractory period to select the R peaks. The R peaks detected by wavelet method are used for the detection of remaining features of ECG signal such as P, Q, S and T waves.

In the energy analysis technique for R wave detection, energy calculation of ECG signal under test has been performed by dividing the signal records into a number of windows. The techniques used include window shifted by window size and window shifted by one sample. Energy analysis of detail coefficients shows that d4 signal containing highest energy content comprise of maximum information of QRS complexes. This concept is proposed to detect ECG peaks.

We have also proposed a technique to detect PVC (premature ventricular contraction) beats in ECG. The method for detecting the abnormal PVC complexes is based on the calculation of RR interval of detected R peaks. We have utilized a combined method for PVC detection where, RR interval calculation by energy method is supported by intersection of energy analysis on the ECG signal with extended window size. The algorithm proposed for PVC detection includes detection of R peaks using window based energy analysis of ECG signal using a window of 100 ms duration that incorporate window shift by one sample and further energy analysis of ECG signal using a window of 600 ms duration where window is shifted by window size.

In wavelet based feature extraction method, the overall sensitivity, positive predictivity and accuracy obtained on 44 records of MIT-BIH database are 99.62%, 99.87% and 99.48%

respectively. The accuracies obtained for records 102 and 104 are 99.86% and 99.77% respectively which are higher than the result of [30] (94% and 92% respectively).

In energy based peak detection algorithms, on the said forty four records we have achieved an accuracy, sensitivity and positive predictivity of 98.17%, 98.82%, 99.36% respectively using window shifted by window size and 98.63%, 99.36%,99.28% respectively using window shifted by one sample. This method reported accuracies of 99.73% and 100% in using window shifted by window size technique; 98.93% and 98.15% in window shifted by one sample technique for records 102 and 104 respectively. Both the techniques resulted higher accuracies than the result of [30] (94% and 92% respectively).

In case of PVC detection on 37 records of MIT-BIH database, the accuracy, sensitivity and positive predictivity achieved are 96.79%, 98.31% and 98.48% respectively.

6.2. Feature extraction of ABP and CVP

Since the wavelet based method applied for ECG feature extraction resulted promising results we have applied this technique for the analysis of ABP and CVP signal also. In this part of our research, we have performed - (i) Extraction of all features of ABP and CVP signals and (ii) energy based approach for peak detection of ABP signal.

6.2.1. Feature extraction of ABP

The wavelets used for decomposition of ABP signal are daubechies (db4) and symmetric (sym4). Since justification of selection of detail coefficient by energy, frequency and cross-correlation analysis is found to be successful in ECG feature extraction, in this part also selection of wavelet detail coefficients has been justified by energy, frequency and cross-correlation analysis. Further, it is found that application of '*window based threshold*' overcomes the setback of missing peaks due to large variations in the signal amplitude at any particular instant.

This algorithm of ABP waveform delineation has been applied on 1 minute segment of 22 signals of MGH/MF waveform database, 14 signals of Fantasia database, 15 signals of MIT-BIH polysmographic database and 1 signal of CSL database. We have selected a certain segment of the signal records where all features of ABP signals are prominent and can be annotated while the distorted portion of the signals are discarded for analysis. The

performance of the algorithm is expressed in terms of percentage of accuracy (Ac), sensitivity (Se) and positive predictivity (PP) and error.

In terms of type of wavelet used - sym4 wavelet is found to result better accuracy (97-98%) and positive predictivity (97-99%) as compared to db4 (93-99%) for all the four components of ABP signal, however the sensitivity in both cases are comparable.

In terms of the database used for the analysis – accuracy and sensitivity is found to be highest in MGH/MF database while MIT-BIH polysmographic database outperforms the other database in positive predictivity. Performance of both the wavelets for all the features on CSL database is found to be better than other three database since we have used only one signal from CSL database.

We have also developed a framework for the detection of systolic peaks in the ABP signal using energy analysis. The energy calculation of the ABP signal is performed by taking a window size of 100 ms duration. Similar to ECG peak detection, the area in ABP signal where systolic peaks are available appear as high energy zones. Window based amplitude threshold and interval threshold are applied to reject the unwanted peaks.

Energy based peak detection algorithm has been applied on a 5 minute segment of ABP signals for the duration for which ECG signal is also available. The number of actual ECG beats is counted from expert annotations of same duration as the ABP signal and beats positions are validated manually. The algorithm has been tested on the first 9 records of MGH/MF waveform database. We have achieved an accuracy of 99.53% for ABP signal of mgh001 record whereas the overall accuracy of detection is 98.05%. Out of total 4121 beats 4043 beats were correctly detected. The result shows that 78 peaks were missing whereas 01 peak is detected as false beat. In addition to accuracy, we have achieved a sensitivity of 99.98% and positive predictive value of 98.14%.

6.2.2. Feature extraction of CVP

The detection of 'x' descent in CVP signal is carried out using db4 wavelet and selection of relevant detail coefficient has been validated based on energy, frequency and correlation technique. Here, negative amplitude thresholding is used to detect 'x' descent precisely which is further used to mark remaining peaks ('a', 'c', 'v') and descent ('y') waves of CVP signal.

Since most of the CVP signals in MGH/MF waveform database represent abnormal waveforms, therefore, we have used only two records – mgh007, mgh008 for analysis since all components are available and can be annotated. Further, CVP signals of this database are not annotated by experts and researchers so far for validation of the algorithm. The CVP signal from record mgh007 record has classical hemodynamic waveforms whereas CVP signal from mgh008 record has been affected by tricuspid regurgitation.

The result shows accuracies of 75.62% for both 'a' and 'c' waves detection; 85.07%, 79.10% and 67.66% for 'x', 'v' and 'y' waves respectively. On the other hand, the sensitivities obtained for 'a', 'c', 'x', 'v' and 'y' waves are 88.05%, 89%, 93.53%, 93.95% and 81.09% respectively. In terms of positive predictivity, the algorithm reported 87.62%, 86.82%, 91.70%, 84.65% and 85.78% respectively for the five mentioned features.

6.3.   Modeling and synthesis of ECG

ECG modeling and synthesis has been performed by other researchers, however, a parametric ECG model based on measured phenomenological cardiac data such as ABP and CVP has not been attempted so far. We have used system identification technique to develop the linear time invariant (LTI) model of the cardiac system which can synthesize ECG from ABP and CVP. Two different models were attempted for modeling- Autoregressive and State Space models.

6.3.1.   Modeling and synthesis using system identification technique

We have used system identification based modeling and synthesis of ECG for 7 records – mgh003, mgh004, mgh005, mgh007, mgh008, mgh029 and mgh031 of MGH/MF waveform database where all three signals (ABP, CVP and ECG) are available for system identification. The models of the cardiac system of both healthy subjects (i.e. normal) and having PVC and SVPB (Supraventricular premature beat i.e. abnormal) were developed and then the models were cross-validated using inputs (ABP, CVP).

We present a system identification framework for modeling of ECG using ABP and CVP signals in autoregressive and state space models using prediction error minimization (PEM) and subspace algorithms of system identification. One cycle of ABP, CVP and ECG from 7 records are used for estimating the state space models.

The accuracy of the synthesized and original ECG is evaluated by a universally accepted metric given by *'best fit'* percentage. An interesting inference was drawn to relate the condition of heart (i.e. a normal or having abnormal PVC/ SVPB) to the stability of the model transfer function. A stable transfer function is obtained if the ECG cycles of the model are normal, whereas an unstable transfer function is obtained from an abnormal ECG cycle. Therefore, the generated transfer functions are analyzed for stability using pole zero plots and step responses.

By using subspace method of identification in state space model, we have achieved a maximum best fit percentage of 80.4852% for normal ECG model validated by normal ABP & CVP data of mgh008 record, while a maximum fit percentage of 75.5584% was achieved when an abnormal ECG model (having SVPB) validated by ABP and CVP data having SVPB in mgh004 record.

Moreover, in the state space modeling by PEM method, the technique reported a highest fit percentage of 74.5805 % when a normal ECG model was validated by normal data in mgh029 signal.

In case of autoregressive modeling, we have achieved a maximum of 76.2032% fit percentage while validating a normal ECG model by normal data of mgh031 signal.

Further, comparative study of all the three models shows that subspace method of system identification results in higher fit percentage for all the seven records. Therefore, it can be concluded that the system generates a stable transfer function when model is simulated using normal data, while an unstable transfer function is generated when simulated using abnormal (PVC and SVPB) beats.

6.3.2. Modeling and synthesis using artificial neural network (ANN)

Due to universal acceptance and advantage of ANN for modeling non-linear systems, we tried to establish a non-linear connectivity between the pressure signals (ABP & CVP) and the ECG signals and thereby synthesizing ECG.

The proposed technique utilizes one of the most popular ANN algorithm − Radial basis function (RBF) to map the synchronously sampled pressure signals (ABP & CVP) as input and the corresponding ECG as the output. We have done ANN modeling on 16 records each comprising 20,000 samples of ABP, CVP and ECG signals of MGH/MF waveform

database to model and synthesize the ECG from which R peaks are detected. We have also done ANN modeling on 16 records each comprising 1,00,000 (85,000 for training and 15,000 for testing) samples of MGH/MF waveform database to model and synthesize the ECG which was used for comparison with original ECG by 'cosine measure' and 'cross-correlation' techniques.

In our first approach of ANN modeling with 20,000 samples, the number of peaks obtained in the synthesized ECG are compared with expert annotations for ECG peaks available from the database. The overall accuracy, sensitivity and positive predictivity are 95.96%, 97.05% and 98.99% respectively. In the second approach on 1,00,000 samples for modeling of ECG, the suynthesized ECG signals are tested by cosine measure and cross-correlation for testing resemblance with original ECG signals. The highest cosine measure and cross-correlation achieved are 0.9493 and 0.9498 in mgh006.

6.4. Scope for future work

We have attempted to develop the frameworks for feature extraction of ECG, ABP and CVP signals and thereby modeling and synthesis of ECG, but due to limitations beyond the scope of this research, the research could not cover many very interesting and important issues as stated below. In view of this, we propose that future researchers can take up the extended parts of this work as stated below —

a) Although the use of window based amplitude thresholding algorithm reduce the accuracy due to of missing peaks or false peaks detection, however the detection accuracy can be further improved if the thresholding method can be further modified to add more adaptability within the predefined window.

b) The performance of PVC detection algorithm described in Chapter 2, can be further enhanced if energy feature for various data records can be used as input to the artificial neural network (ANN) to train it and classification of a record can be obtained. Energy feature can also be combined with correlation and the algorithm performance can be further enhanced.

c) The proposed algorithm for ABP feature extraction can be employed to detect myocardial ischemia, detection of premature ventricular contraction (PVC) and

---

premature supra-ventricular contraction (PSC) beats along with ECG signal to enhance the accuracy of disease identification.

d) Performance of CVP feature extraction algorithm can be further enhanced by constructing new wavelets. The algorithm after further modifications can be used to detect different diseases such as *atrial fibrillation, tricuspid stenosis, cardiac temponade* etc.

e) Modeling of ECG proposed in Chapter 4 and 5, can be extended to synthesize other biomedical signals such as CVP, ABP and PAP which are acquired invasively.

REFERENCES

[1] Lim, S.S, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet*, **380**(9859), 2224–2260, 2012.

[2] The global burden of disease: 2004 update, Geneva, World Health Organization, 2008.

[3] Bin, He. et al. Grand challenges in interfacing engineering with life sciences and medicine, *IEEE Transactions on Biomedical Engineering*, **60**(3) ,589-598, 2013.

[4] Coblentz, B. et al. The relationship between electrical and mechanical events in the cardiac cycle of man, *J. Br Heart*, **11**, 1-22, 1949.

[5] Headley, J.M. et al. *Invasive hemodynamic monitoring: physiological principles and clinical applications*, Edwards Life Sciences, 31-33, 2002.

[6] American Heart Association, The heart : how it works, *Heart and Stroke Facts*, 1-78, 1992-2003.
http://www.americanheart.org/downloadable/heart/1056719919740HSFacts2003text.pdf

[7] Dash, P.K. Electrocardiogram monitoring, *Indian Journal of Anaesth*, **46** (4), 251-260, 2002.

[8] Schamroth, L., *An introduction to electrocardiography*, Blackwell Scientific Publications, 1990.

[9] Daimiwal, N., Microcontroller based ECG and blood pressure simulator", *J. Instrum. Soc. India*, **37** (4), 243-248.

[10] Pressmant, G. L. & Newgardt, P. M. A transducer for the continuous external measurement of arterial blood pressure, *IEEE transactions on bio-medical electronics*, 73-81, 1963.

[11] Kain, H.K., et al. Arterial blood pressure measurements with a portable recorder in hypertensive patients : I. variability and correlation with casual pressures, *Circulation*, **30**, 882-892,1964.

[12] Biomedical signal processing laboratory, CSL database (http://bsp.pdx.edu), 2011.

[13] Christov, I., et al. Automatic detection of atrial fibrillation and flutter by wave rectification method", *J. Medical Engineering and Technology*, **25** (5), 217-221, 2001.

[14] Burke, M.J. & Nasor, M. Wavelet based analysis and characterization of ECG signal, *J. Medical Engineering and Technology*, **28** (2), 47-55, 2004.

[15] Chuang-Chien-Chiu et al. Using correlation coefficient in ECG waveform for arrhythmia detection, *Biomedical Engineering Applications, Basis and Communications,* **17** (3), 147-152, 2005.

[16] Inan, O.T., et al. Robust neural-network based classification of premature ventricular contractions using wavelet transform and timing interval features, *IEEE transactions on Biomedical Engineering* , **53** (12), 2006.

[17] Silipo, R., & Marchesi, C. Artificial neural networks for automatic ECG analysis, *IEEE Transactions on Signal Processing*, **46**(5), 1417-1424, 1998.

[18] Ferguson C., et al. Atrial fibrillation : stroke prevention in focus, *Aust Crit Care* 2013, Epub 2013 in press, http://dx.doi.org/10.1016/j.aucc.2013.08.002.

[19] Izacovic, M. Central venous pressure–evaluation, interpretation, monitoring, clinical implications, *Bratisl Lek Listy*, **109**(4), 185-187, 2008.

[20] Cuiwei, Li., et al. Detection of ECG characteristic points using wavelet transforms, *IEEE Trans. Biomed. Eng*, **42**(1), 1995.

[21] Sahambi, J.S., et al. Using wavelet transforms for ECG characterization, an online digital signal processing system, *IEEE Engineering in Medicine & Biology*, 77-83, 1997.

[22] Saxena, S.C., et al. QRS detection using new wavelets, *J. of Medical Engineering & Technology*, **26**, 7-15, 2002.

[23] Legarreta, I.R., et al. R-wave detection using continuous wavelet modulus maxima, *IEEE Computers in Cardiology*, **30**, 565-568, 2003.

[24] Mahmoodabadi, S.Z., Ahamadian, A., Abolhasani, M.D., Eslami M. & Bidgoli, J.H. ECG feature extraction based on multiresolution wavelet transform, Proceedings of IEEE, Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 2005.

[25] Rudnicki, M. & Strumillo, P., A real-time adaptive wavelet transform - based QRS complex detector, ICANNGA 2007, Part II, LNCS 4432, Springer-Verlag Berlin Heidelberg, 281–289, 2007.

[26] Tu, C., et al., A new approach to detect QRS complexes based on a histogram and genetic algorithm, *J. Medical Engineering & Technology*, **29** (4), 176-180, 2005.

[27] Ghongade, R., & Ghatol, A.A. A brief performance evaluation of ECG feature extraction techniques for artificial neural network based classification, TENCON 2007, 1-4, doi:10.1109/TENCON.2007.4429096.

[28] Chen, H.C. & Chen, S.W. A moving average based filtering system with its application to real-time QRS detection, *Computers in Cardiology*, **30**: 585-588, 2003, doi:10.1109/CIC.2003.1291223.

[29] Wrublewski, T.A., Ying, S. & Beyer, J.A. Real-time early detection of R-waves of ECG Signals, ECG signal processing III, IEEE Engineering in Medicine & Biology Society, 11th Annual International Conference 1989.

[30] Lee, R.G., et al. A novel QRS detection algorithm applied to the analysis for heart rate variability of patients with sleep apnea, *Biomedical Engineering Applications, Basis & Communications*, **17**(5)258 – 262, 2005.

[31] Miyakawa, K., et al. *Mechanism of blood pressure wave*, Japan Scientific Press, Springer- Verlag, Tokyo, 1984.

[32] Sun, J.X., et al. Estimating cardiac output from arterial blood pressure waveforms: a critical evaluation using the MIMIC II database, www.physionet.org/physiotools/cardiac-output/doc/s54-5.pdf.

[33] Aboy, M., et al. A novel algorithm to estimate the pulse pressure variation index, *IEEE Transactions on Biomedical Engineering*, **51** (12), 2004.

[34] Palaniappan, R., Gupta, C.N., Luk, C.K. & Krishna, S.M., Multi-parameter detection of ectopic heart, International Workshop on Biomedical Circuits & Systems (BioCAS' 2004), S2.4-1-44, IEEE

[35] Palaniappan, R., & Krishnan, S.M., Detection of ectopic heart beats using ECG and blood pressure signals, International Conference on Signal Processing & Communications (SPCOM' 2004).

[36] Miall-Allen, V.M., et al, Mean arterial blood pressure and neonatal cerebral lesions, *Archives of Disease in Childhood*, 1068-69, 1987.

[37] Davies, J.I. & Struthers, A.D. Beyond blood pressure: pulse wave analysis – a better way of assessing cardiovascular risks, *Future Cardiology* 1, 69–78, 2004.

[38] Ferreira, A., et al. Three-section transmission line arterial model for noninvasive assessment of vascular remodeling in primary hypertension, *Biomedical Signal Processing and Control* 4 (1) 2–6, 2009.

[39] Aboukhalil, A., et al. Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform, *J. Biomed Informatics*, 41(3), 442–451, 2008.

[40] Navakatikyan, M.A., et al. A real-time algorithm for the quantification of blood pressure waveforms, *IEEE Transactions on Biomedical Engineering*, 49 (7), 2002.

[41] Hocksel, S.A.A.P. et al. Detection of dicrotic notch in arterial pressure signals, *J. Clinical Monitoring*, 13, 309-316, 1997.

[42] Aboy, M., Crespo, C., McNames, J., & Goldstein, B. Automatic detection algorithm for physiologic pressure signal components, Proceedings of Second Joint EMBS/BMES Conference (Houseton, TX USA, 2002), 196-197.

[43] Zong, W., et al. An open source algorithm to detect onset of arterial waveform pulses, *Computers in Cardiology*, 30, 259-262, 2003.

[44] Aboy, M., et al. An automatic beat detection algorithm for pressure signals, *IEEE Transactions on Biomedical Engineering*, 52 (10), 2005.

[45] Li, B.N. et al. On an automatic delineator for arterial blood pressure waveforms, *Biomedical Signal Processing and Control*, 5, 76–81, 2010.

[46] Zong, W., et al. Effects of vasoactive drugs on the relationship between ECG-pulse wave delay time and arterial blood pressure in ICU patients, *Computers in Cardiology,* **25**, 673-676, 1998.

[47] Daubechies, I. & Sweldens, W., Factoring Wavelet Transforms into Lifting Steps, Preprint, Bell Laboratories, Lucent Technologies (1996).

[48] Kaushik, G. et al. Biomedical signal analysis through wavelets: A review, *International Journal of Advanced Research in Computer Science and Software Engineering* 2(9), 22-28, 2012. www.ijarcsse.com

[49] Rai, H.M. et al. De-noising of ECG waveforms based on multi-resolution wavelet transform, *International Journal of Computer Applications,* 45(18), 25-30, 2012.

[50] Patil, P.B. et al. A wavelet based method for denoising of biomedical signal, in Proceedings of the IEEE International Conference on Pattern Recognition, Informatics and Medical Engineering, 2012, 278-283.

[51] Ergen, B. Signal and image denoising using wavelet transform, In Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology In Tech, Europe, 2012, 495-514.

[52] Brechet., I. Compression of biomedical signals with mother wavelet optimization and basis wavelet packet selection, *IEEE Transactions on Biomedical Engineering,* **54**(12), 2186-2192, 2007.

[53] Rajoub, B.A. An efficient coding algorithm for the compression of ECG signals using the wavelet transform, *IEEE Transactions on Biomedical Engineering,* Vol. **49** (4) 355-362, 2002.

[54] Nielsen, M. et al. Optimal wavelets for biomedical signal compression, *Med Bio Eng Comput,* 44, 561–568, 2006. DOI 10.1007/s11517-006-0062-0.

[55] Wu, T.C. et al. Wavelet-based ECG data compression optimization with genetic algorithm, *J. Biomedical Science and Engineering,* 6, 746-753, 2013.

[56] Mallat S., Zero-crossings of a wavelet transform, *IEEE Transactions on Information Theory,* **37**(4), 1019-1033, 1991.

[57] Venkatraman, M., and Govindaraju, V. Zero crossings of a non-orthogonal wavelet transform for object location, 57-60, 1995 IEEE.

[58] Robi Polikar, R. The wavelet tutorial part IV, multiresolution analysis: The discrete wavelet transform,www.cs.ucf.edu/courses/cap5015/WTpart4.pdf, (2008).

[59] Paul S Addition, P.S. Wavelet transforms and the ECG: a review, *Institute of Physics Publishing, Physiol. Meas.*, 26, R155-R199. 2005.

[60] Gargour, C.S. & Ramachandran, V. A scheme for teaching wavelets at the introductory Level, in Frontiers in Education Conference, 1997, 27th Annual Conference. Teaching and Learning in an Era of Change, Proceedings. Pittsburgh, PA, USA, 1997. http://fie.engrng.pitt.edu/fie97/papers/1030.pdf

[61] Mallat S. et al. Characterization of signals from multiscale edges, *IEEE Trans. Pattern Anal. Mach.Intell.*, 14, 710–732, 1992.

[62] Minhas, F.A. et al. Robust electrocardiogram (ECG) beat classification using discrete wavelet transform, *Physiol. Meas.* 29, 555–570, 2008.

[63] Sharma, S. *Signals and Systems*, S. K. Kataria & Sons, New Delhi, 2008.

[64] Hohlfeld, R.G. et al. Wavelet signal processing of physiologic waveforms, *Wavelet Technologies, Inc.*, Copyright 2004.

[65] Keenan, D.B., Detection and correction of ectopic beats for HRV analysis applying discrete wavelet transforms, *Int. J. Information Technology*, 2 (1), 54-60.

[66] Narasimhan, S.V. & Veena, S. *Signal Processing, Principles and Implementation*, Narosa Publishihng House, New Delhi, 2008.

[67] Gelsema, E.S. et al. Application of the method of multiple thresholding to white blood cell classification, *Comput, Biol, Med*, 18(2), 65-74, 1988.

[68] Donoho, D.L. et al. Minimax estimation via wavelet shrinkage, *The Annals of Statistics*, 26(3), 879-921, 1998.

[69] Sezgin, M. et al. Survey over image thresholding techniques and quantitative performance evaluation, *Journal of Electronic Imaging*, 13(1), 146–165, 2004.

[70] Chaabane, S.B. et al. Color edges extraction using statistical features and automatic threshold technique: application to the breast cancer cells, *Bio Medical Engineering*, 13(4), 1-18, 2014.

[71] Ballesteros, D.M. et al. Discrete wavelet transform in compression and filtering of biomedical signals, Discrete wavelet transforms - Biomedical applications, In Tech, Croatia, China, 17-32, http://cdn.intechopen.com/pdfs-wm/19507.pdf

[72] Söderström, T. Identification of linear lystems in time domain, *Control Systems, Robotics and Automation*, V, Encyclopedia of Life Support Systems (EOLSS).

[73] Chang, H. et al. Analysis of the dynamic characteristics of pressure sensors using ARX system identification, *Sensors and Actuators*, A 141, 367–375, 2008.

[74] Verhaegen, M. & Verdult, V. *Filtering and System Identification, A Least Square Approach*, Cambridge University Press, 2007.

[75] Juricek, B.C., Seborg, D.E., & Larimore, W.E. Process control applications of subspace and regression-based identification and monitoring methods, in IEEE American Control Conference, 2005, Portland.

[76] Smith, J.O. *Physical Audio Signal Processing*, W3K Publishing, 2010.

[77] Chon, K.H., et al. Linear and non-linear system identification of autonomic heart rate modulation, *IEEE Engineering in Medicine and Biology*, 96 – 105, 1997.

[78] Cohen, I. Relative transfer function identification using speech signals, *IEEE Transactions on Speech and Audio Processing* 12 (5), 451–459, 2004.

[79] Wu, M., & Wang, D. A two-stage algorithm for one-microphone reverberant speech enhancement, *IEEE Transactions on Audio, Speech and Language Processing* 14 (3), 774–784, 2006.

[80] Westwick, D.T. Nonlinear system identification in biomedical engineering: Techniques, applications and challenges, 4th IFAC Symposium on System Identification, 2006, 14(1), 128-130, 10.3182/20060329-3-AU-2901.00012.

[81] Abbas, A.K. & Basaam, R. Adaptive system identification and modeling of respiratory acoustics, 13th International Conference on Biomedical Engineering – ICBME' 2008), 3-6.

[82] Feng, K.P.W.D. & Siu, W.C. Fast system identification algorithm for non-uniformly sampled noisy biomedical signal, IEEE ENCON - Digital Signal Processing Applications' 1996, 559-564.

[83] Feng, D. et al, An unbiased parametric imaging algorithm for non-uniformly sampled biomedical system parameter estimation, *IEEE Transactions on Medical Imaging*, **15** (4), 1996.

[84] Zhao, Y. & Kearney, R.E. System identification of biomedical systems from short transients using space methods, 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, 2008.

[85] Muto, M.M. *Application of stochastic simulation methods to system identification*, PhD Thesis, California Institute of Technology Pasadena, California, 2007.

[86] Ljung, L. *System Identification – Theory for the User*, Prentice-Hall, Englewood, 1999.

[87] Sara, A. & Geer, D.D. Least squares estimation, encyclopedia of statistics in behavioral science, **2**, 1041–1045, John Wiley & Sons, Ltd, Chichester, 2005.

[88] McKelvey, T. Subspace methods for frequency domain data, in Proceeding of the 2004 American Control Conference, Boston, Massachusetts, 2004, 673-678.

[89] Zhao, Y. & Kearney, R.E., System identification of biomedical systems from short transients using space methods, 30th Annual International IEEE EMBS Conference, 2008, Vancouver, British Columbia, Canada.

[90] Poshtan, J. et al, Subspace system identification, *Iranian Journal of Electrical & Electronic Engineering*, **1** (1), 11-17, 2005.

[91] Favoreel, W. et al. Subspace state space system identification for industrial processes, *Journal of Process Control*, **10**, 149-155, 2000.

[92] Overschee, P.V. & Moor, B.D. *Subspace Identification For Linear Systems: Theory, Implementation, Applications,* Kluwer Academic Publishers, Boston/ London/ Dordrecht, The Netherlands, 1996.

[93] Moor B.D. *Mathematical concepts and techniques for modeling of static and dynamic systems.* PhD Thesis, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium, 1988.

[94] Ogata, K., Modern control engineering, Prentice Hall of India, 1991.

[95] Bakshi, U.A. & Bakshi, V.U. *Control Systems,* Technical Publications Pune, chapter 8, 4-8, 2010.

[96] Silva, L., et al. Medical multivariate signal reconstruction using recurrent neural network, *Computing in Cardiology,* 37, 445–447, 2010.

[97] Sullivan, A.M., et al. Reconstruction of missing physiological signals using artificial neural networks, *Computing in Cardiology,* 37, 317–320, 2010.

[98] Rodrigues, R., Filling in the gap: a general method using neural networks, *Computing in Cardiology,* 37, 453-456, 2010.

[99] McBride, J., et al. Reconstruction of physiological signals using iterative retraining and accumulated averaging of neural network models, *IOP Publishing Physiol. Meas.,* 32, 661-675, 2011.

[100] Skarlas, L.V. et al. Intelligent modeling of the electrical activity of the human heart, *International J. Signal Processing* 1(1), 38-41, 2005, www.waset.org.

[101] Al-Nashash, H.A., Zalzala, A,S., & Thakar, N.V, A Neural Networks Approach to EEG signals Modeling, Proceedings of the 25th Annual International Conference of the IEEE EMBS' 2003.

[102] Sordo, M., Buxton, H., & Watson, D. A hybrid approach to breast cancer diagnosis, in Practical Applications of Computational Intelligence Techniques. Jain, L. and DeWilde, P. Eds.). Kluwer Academic Publishers. 2001.

[103] Burke, H., Rosen, D., & Goodman, P. Comparing the prediction accuracy of artificial neural networks and other statistical models for breast cancer survival, in Tesauro,

G., Touretzky, D., & Leen, T. (Eds.) 1995, Advances in Neural Information Processing Systems, 7, 1063-1067, The MIT Press.

[104] Lapuerta, P., Azen, S. P., & LaBree, L. Use of neural networks in predicting the risk of coronary artery disease, *Computers and Biomedical Research*, 28, 38-52. 1995.

[105] Sordo, M. *Neural networks for detection of Down's Syndrome*. Master's thesis, Department of Artificial Intelligence, University of Edinburgh, 1995.

[106] Costa, A., Cabestany, J., Moreno, J., & Calvet, M. An artificial neural net based diagnostic aid tool for serum electrophoresis, in Ifeachor, E., Sperduti, A., & Starita, A. (Eds.), Third International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, 1998, 34-43, World Scientific.

[107] Chazal, P.D., et al. Automatic classification of heartbeats using ECG morphology and heartbeat interval features, *IEEE Transactions on Biomedical Engineering*, 51(7), 1196-1206, 2004.

[108] Özkan, M., Sprenkels, H., & Dawant, B. Multi spectral magnetic resonance image segmentation using neural networks, in International Joint Conference on Neural Networks, 1990, 1,429-434.

[109] Fraser, H., Pugh, R., Kennedy, R., Ross, P., & Harrison, R. A comparison of backpropagation and radial basis functions, in the diagnosis of myocardial infarction, in Ifeachor, E., & Rosen, K. (Eds.), in International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, 1994, 76—84.

[110] Awad, M. et al. Prediction of time series using RBF neural networks: A new approach of clustering, *The International Arab Journal of Information Technology*, 6(2), 138-144, 2009.

[111] Dua, H. et al. Time series prediction using evolving radial basis function networks with new encoding scheme, *Neurocomputing*, 71, 1388-1400, 2008.

[112] Chen, S., Cowan, C.F.N. and Grant, P.M. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on Neural Networks*, 2(2), 302–309. 1991.

[113] McSharry, P.E., et al. Method for generating an artificial RR tachogram of a typical healthy human over 24-hours, *Computers in Cardiology*, **29**, 225– 228, 2002, doi: 10.1109/CIC.2002.1166748.

[114] McSharry, P.E., et al. A dynamical model for generating synthetic electrocardiogram signals, *IEEE Trans Biomed Eng* **50** (3), 289–294, 2003.

[115] Clifford G.D., & McSharry PE. A realistic coupled nonlinear artificial ECG, BP, and respiratory signal generator for assessing noise performance of biomedical signal processing algorithms, Proc of SPIE International Symposium on Fluctuations and Noise, 2004, **34**, 290–301.

[116] Clifford, G.D., et al. Model-based filtering, compression and classification of the ECG, *Int. J. Bioelectromagn*, **7**, 158–161, 2005.

[117] Sameni, R., et al. Multichannel ECG and noise modeling: application to maternal and fetal ECG signals, *EURASIP J. Adv. Signal Process.* 1-14, 2007, doi:10.1155/2007/43407.

[118] Clifford G.D., et al. An artificial multi-channel model for generating abnormal electrocardiographic rhythms, *Computers in Cardiology*, **35**, 773-776, 2008.

[119] Clifford G.D., et al. An artificial vector model for generating abnormal electrocardiographic rhythms, *Physiol. Meas.*, **31**, 595–609, 2010.

[120] Sayadi, O., et al. Synthetic ECG generation and Bayesian filtering using a Gaussian wave-based dynamical model, *Physiological Measurement*, **31**, 1309–1329, 2010.

[121] Sayadi, O., et al. Life-threatening arrhythmia verification in ICU patients using the joint cardiovascular dynamical model and a Bayesian filter, *IEEE Transactions On Biomedical Engineering*, **58** (10), 2748-2757, 2011.

[122] Plonsey, R. & Barr, R.C. Mathematical modeling of electrical activity of the heart, *J. Electrocardiology*, **20**, 219—226, 1987.

[123] Wang, J.T. et al. A knowledge-based system for qualitative ECG simulation and ECG analysis, 773-776, 1992, IEEE.

[124] Rasiah, A.I, Togneri, R. & Attikiouzel, Y. Modeling 1-D signals using hermite basis functions, IEE Proceedings on Vision, Image, and Signal Processing, **144** (6), 345-354, 1997.

[125] Dixit, P., Segmented modeling of ECG signal by using hermite basis function, *International J. Engineering and Mathematical Sciences (IJEMS)*, **1** (1), 40-46, 2012, www.ijems.org.

[126] Ahmadian, A., Karimifard, S., Sadoughi, H., & Abdoli, M. An efficient piecewise modeling of ECG signals based on Hermitian basis functions, Proceedings of the 29th Annual International Conference of the IEEE EMBS (2007), Lyon, France, 3180-3183.

[127] Ayatollahi, A., et al. A comprehensive model using modified Zeeman model for generating ECG signals, *Iranian J. Electrical & Electronic Engineering*, **1**(2), 88-93, 2005.

[128] Parvaneh, S., & Pashna, M. Electrocardiogram synthesis using a Gaussian combination model (GCM), *Computers in Cardiology*, **34**, 621-624, 2007.

[129] Nunes, J.C. & Amine, N.A. Hilbert transform based ECG modeling, Biomedical Engineering, Springer, **39** (3), 133-137, 2005.

[130] Andreao, R.V. et al. ECG signal analysis through Hidden Markov Models, IEEE *Transactions on Biomedical Engineering*, **53** (8), 1541-1549, 2006.

[131] Zheying, Li., & Minjie, M., ECG modeling with DFG, 27th Annual Conference of IEEE Engineering in Medicine and Biology, 2691-2694, 2005.

[132] Moody, G.B., The physionet/computing in cardiology challenge 2010: mind the gap, *Computing in Cardiology*, **37**, 305-309, 2010.

[133] Thomas, C., et al. An approach to reconstruct lost cardiac signals using pattern matching and neural networks via related cardiac information, *Computing in Cardiology*, **37**, 441−444, 2010.

[134] Wei, W., The multi-parameter physiologic signal reconstruction by means of wavelet singularity detection and signal correlation, *Computing in Cardiology*, **37**, 457–459, 2010.

[135] Goldberger,A.L. et al. PhysioBank, PhysioToolkit, and PhysioNet components of a new research resource for complex physiologic signals, *Circulation*, **101**, e215-e220, 2000. doi: 10.1161/01.CIR.101.23.e215. (http://www.physionet.org/physiobank)

[136] MIT-BIH database (http://www.physionet.org/mitdb), 2009

[137] Fantasia database (http://www.physionet.org/physiobank/database/fantasia), 2011

[138] MIT-BIH polysmographic database

(http://www.physionet.org/physiobank/database/slpdb), 2011

[139] MGH/MF waveform database (http://www.physionet.org/mghdb), 2011

[140] Levkov Ch., et al. Removal of power-line interference from the ECG: a review of the subtraction procedure, *Biomedical Engineering Online*, **4** (50), 1-18, 2005.

[141] Barick S., & Bagha S., Removal of 50Hz powerline interference for quality diagnosis of ECG signal, *International Journal of Engineering Science and Technology (IJEST)*, **5** (5), 1149-155, 2005.

[142] Pallavi S., Kulkarni & Kakatkar, M., ECG Signal De-nosing using Wavelet Transform from MIT-BIH Database, *International Journal of Electronics Communication and Computer Engineering*, **5**(4), 149-152, 2014.

[143] Borries, R.F., J. H. Pierluissi & Nazeran, H., Wavelet transform-based ECG baseline drift removal for body surface potential mapping, in 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (EMB' 2005), Shanghai, China, 3891-3894.

[144] Thakor N.V., et al. Estimation of QRS complex power spectra for design of a QRS filter, *IEEE Trans. Biomedical Engg.*, **31**, (11), 702-705, 1984.

[145] Constantin Voiniciuc' Vasile Apope, ECG data acquisition system and processing in presence of noise artifact", *Optimization of Electrical and Electronic Equipments* – Brasov, 693-696, 1998.

[146] Rangayyan, R.M. *Biomedical signal analysis, A case-study approach*, Wiley Interscience, John Wiley & Sons INC, 2002.

[147] Harting, L.P., Fedotov, N.M. & Slump, C.H. On base line drift suppressing in ECG recordings, in Proceedings of IEEE Signal Processing Symposium (SPS' 2004), Benelux, Russia, 133-136.

[148] Paul, J.S., Reddy, M.R.S. & Kumar, V.J., Automatic detection of PVC's using autoregressive models, Proceedings - 19th International Conference - IEEE/EMBS (Chicago, IL USA.1997), 68-71.

[149] Zhao, L. Wiggins, M. & Vachtsevanos, G. Premature ventricular contraction beat detection based on symbolic dynamics analysis, Proceedings of the IASTED International Conference Circuits, Signals and Systems (Cancun, Mexico, 2003), 48-50.

[150] Lorena, S. C. de Oliveira, R V. Andreão & Sarcinelli-Filho, M., Detection of premature ventricular beats in ECG records using Bayesian Networks involving the P-wave and fusion of results, 32nd Annual International Conference of the IEEE EMBS, Buenos Aires, Argentina 2010.

[151] Joon S. L., Finding features for real-time premature ventricular contraction detection using a Fuzzy Neural Network system, *IEEE Transactions on Neural Networks*, 20 (3), 522-527, 2009.

[152] Krasteva, V., & Jekova, I., QRS template matching for recognition of ventricular ectopic beats, *Annals of Biomedical Engineering*, 35 (12), 2065-2076, 2007.

[153] Bortolan, G., Jekova, I., & Christov, I., Comparison of four methods for premature ventricular contraction and normal beat clustering, *Computers in Cardiology*, 32, 921-924, 2005.

[154] Christov, I., Jekova, I., & Bortolan, G., Premature ventricular contraction classification by the Kth nearest-neighbours rule, *Physiological Measurement*, **26**, 123-130, 2005.

[155] Jekova, I., Bortolan, G., & Christov, I., Assessment and comparison of different methods for heart beat classification, *Medical Engineering & Physics*, **30**, 248-257, 2008.

[156] Herrero, G. et al. Relative estimation of the Karhunen-Loeve Transform basis functions for detection of ventricular ectopic beats, *Computers in cardiology*, **33**, 569-572, 2006.

[157] Jekova, I., Bortolan, G., & Christov, I., Pattern recognition and optimal parameter selection in premature ventricular contraction classification, *Computers in cardiology*, **31**, 357-360, 2004.

[158] Jekova, I., & Krasteva, V., Fast algorithm for vectorcardiogram and interbeat intervals analysis: Application for premature ventricular contractions classification, *Bioautomation*, **3**, 1-12, 2005.

[159] ANSI/AAMI CE57, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. (AAMI), recommended practice/American national standard. http://www.aami.org,1998.

[160] Hamilton P.S. & Tompkins, W.J., Quantitative investigation of QRS detection rules using MIT BIH arrhythmia database, *IEE Transactions on Biomedical Engg.* **33**(12), Dec 1986.

[161] Christov, I., et al. Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification, *Medical Engineering & Physics*, **28**, 876-887, 2005.

[162] Bernardo, D. D., et al. A model of two nonlinear coupled oscillators for the study of heartbeat dynamics, *Int. Journal of Bifurcation and Chaos*, **8** (10), 1998.

[163] Gopal, M. *Modern control system theory*, New Age International (P) Limited Publishers, 266, second edition, 2010.

[164] Wei-Ho Tsai & Hao-Ping Lin. Background music removal based on cepstrum transformation for popular singer identification, *IEEE Transactions on Audio, Speech, and Language Processing*, **19** (5), 1196-1205, 2011.

LIST OF PUBLICATIONS

[1] Pachauri, A., & Bhuyan, M. Robust detection of R-wave using wavelet technique, in International Conference on Signal Processing, Communication and Networking (ICSPCN' 2009), Singapore, 335-339, World Academy of Science, Engineering and Technology **56**, 2009, 901-905.

[2] Pachauri, A., & Bhuyan, M. Wavelet and energy based approach for PVC detection, in IEEE International Conference on Emerging Trends in Electronic and Photonic Devices & Systems (ELECTRO' 2009), Varanasi, India, 258-261.

[3] Pachauri, A., & Bhuyan, M. PVC detection by energy analysis", in 2nd International Conference on RF and Signal Processing Systems (RSPS' 2010), Guntur, India, 380-384.

[4] Pachauri, A., & Bhuyan, M. A new approach to ECG peak detection, in International Conference on Biomedical Engineering and Assistive Technologies (BEATS' 2010), Jalandhar, Punjab, India.

[5] Pachauri, A., & Bhuyan, M. ABP peak detection using energy analysis technique, IEEE International Conference on Multi-media, Signal Processing and Communication Technologies (IMPACT' 2011), AMU, Aligarh, India.

[6] Pachauri, A., & Bhuyan, M. Wavelet transform based arterial blood pressure waveform delineator, *International Journal of Biology and Biomedical Engineering*, **1** (6), 15-25, 2012.

[7] Pachauri, A., & Bhuyan, M. Modeling of ECG using ABP and CVP signals: A system identification based approach, *International Journal of Engineering, Science and Innovative Technology*, 2(6), 321 -330, 2013.

[8] Pachauri, A., & Bhuyan, M. Modeling of ECG from arterial blood pressure and central venous pressure using artificial neural network, in IEEE International conference on recent advances and innovations in engineering (ICRAIE' 2014), Jaipur, India.

Publication under review:

[1] Pachauri, A., & Bhuyan, M. System identification based modeling and synthesis of electrocardiograms, *International Journal of Modeling and Simulation.*