IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 11, Issue 3, Ver. IV (May-Jun .2016), PP 47-52
www.iosrjournals.org

Non-Invasive Detection of Coronary Stenosis By Parametric And Harmonic Modeling of Heart Sounds

Poornima K[1], Roopashree[2]
[1](MTech student, Dept of ECE, SCEM, Mangalore, Karnataka, India)
[2](Assistant Professor, Dept of ECE, SCEM, Mangalore, Karnataka, India)

Abstract: *It has been found that the partial coronary stenosis generates sounds due to turbulent blood flow. The diastolic heart sounds gives the indications of this stenosis. The electronic stethoscopes can be used to record the heart sounds, which can then be analyzed using computers for detecting coronary occlusions. In this study, Heart sound segmentation is done using Duration dependant Hidden Markov Model which extracts diastolic duration from heart sound. Noise signals are reduced using sub segmentation of diastolic signal which is then modeled by AR method. Akaike's Information Criteria showed the AR pole orders 5-15 to describe diastolic period. The diastolic signal is analyzed over three overlapping frequency bands. The MUSIC method is also used to study the frequency spectrum of heart signal. The Principal Component Analysis is used the find the correlation between the features of AR model and peak frequencies obtained using MUSIC method.*
Keywords: *Auto Regressive Model, Coronary Artery Disease, Heart Sounds, Markov Model, Multiple Signal Classification*

I. Introduction

People all over the world are so much worried about Cardiovascular Diseases and Coronary artery Disease (CAD) is one among them. Timely detection of CAD is very important to save life of a person. Some commonly used techniques for finding heart related problems are Stress Echocardiography (non invasive imaging), Single-photon Emission Computed Tomography, Positron Emission Tomography, Cardiac MRI, Multislice Cardiac Computed Tomography, coronary angiography, auscultation etc. These methods prove beneficial to a person, if they are used for detection at an early stage. Hence there is a need to aid the physicians in detecting heart ailments when a person approaches him for first consultation itself. The heart sounds recorded through an electronic stethoscope can be used for this purpose. It provides a non-invasive, inexpensive and easy approach towards detection of coronary stenosis.

Previous studies have showed that, the blood flow through the stenosed coronary arteries generates sounds due to turbulent blood flow. These sounds are rarely audible, but different signal processing methods can be used to find the relationship between these murmurs and heart diseases. Earlier methods have also found the existence of both low frequency and high frequency components in case of occluded arteries. The FFT, STFT, wavelet Transforms and parametric methods have been used in different earlier studies.

Also the data used in earlier studies were collected using specially designed cardiac microphone or accelerometer. But these devices are not helpful in a small clinical environment. The present study aims at coronary artery disease detection at an early stage in a non-invasive, simple and inexpensive way using a commercially available electronic stethoscope. The identification of diastolic period in a heart signal is also an important issue. The study also includes acquiring features of a diastolic signal in various frequency bands. The ECG signal was used as a reference to find the beginning of second heart sound in many earlier studies. Hidden Markov Model was proposed later for segmentation. In this study, the diastolic period is extracted using Duration Dependant Hidden Markov Model (DDHMM). The poles of AR model are calculated using Yule-Walker method which is used in the detection of Coronary Artery Disease. The frequency spectrum is studied using Multiple Signal Classification (MUSIC) method. The numbers of correlated features are reduced using the Principal Component Analysis (PCA). The Discriminant Analysis is used to construct a score based on the selected features to detect the CAD.

II. Method

2.1 Preprocessing

In the previous studies, diastolic periods were studied only in high frequency region. In order to study the effect of CAD on lower frequencies, the frequency band is split into three overlapping components such as 250-1000Hz, 50-500Hz and 25-1000Hz in this study. The diastolic periods are identified using Duration Dependant Hidden Markov Model. The handheld electronic stethoscope generates friction spikes due to the friction between skin and stethoscope. This is removed by dividing diastole duration into overlapping sub segments of small duration (500ms). Since the sub segment duration is small, the CAD murmur can be assumed to be stationary during this period. The blood flow is highest during diastole period. Hence analysis is done during diastolic period.

2.2 Parametric Model

Spectral density estimation or spectrum analysis implies estimating spectral density of a random signal which is a function of time. In this process, the complex signal is divided into a number of simple components. Usually in spectral analysis, the signal parameters such as amplitude, power, phase etc. are analysed as a function of frequency. Different techniques exist for spectral estimation. Broadly it can be classified as parametric and non parametric models.

In parametric spectral estimation, it is assumed that the signal is modeled by a stationary stochastic process whose structure can be described using a small number of parameters. It is also assumed to have spectral density as a function of some frequency and model parameters. Hence parametric power spectrum estimation methods avoid the problem of leakage and provide better frequency resolution than FFT based non parametric methods. AR estimation, MA estimation, ARMA estimation and maximum entropy spectral estimation are some of the parametric estimation techniques. Of the above linear models, the AR model is most widely used because it is suitable for representing spectra with narrow peaks (resonances) and gives simple linear equations.

2.2.1 Auto Regressive (AR) model

It assumes that the present sample is correlated with the previous samples. An AR model can be viewed as an all pole Infinite Impulse Response (IIR) filter's output, whose input is white noise. The definition is as follows:

$x_t = \sum_{i=1}^{N} a_i x_{t-i} + \epsilon_t$

The current term of the series is the linear weighted sum of previous terms in the series. These weights are called as the auto regression coefficients. The aim in AR analysis is to derive the best values for a_i given a series x_t.

The auto regression model with order p is given as AR (p) and is expressed as

$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \ldots\ldots + a_p x_{t-p} + \epsilon_t$

The solution to this difference equation depends on the 'p' starting values of $\{x_t\}$ and on $\{\epsilon_t\}$. Generally the estimates of the model improve as the order increases (except in case of very noisy signal). There is no straight forward way to detect the correct order of the model. The root mean square (RMS) error generally decreases quickly up to some order as the order of the model is increased and then decreases slowly. The appropriate order is that point after which the RMS error decreases slowly. The Akaike's Information Criteria (AIC) is a method used to determine the AR model order. AIC is calculated for a set of models and the model with the least AIC is considered the best. Fig.1 shows variations of final prediction error (FPE) as a function of model order. FPE decreases with increasing model order up to some order, and then is almost constant.

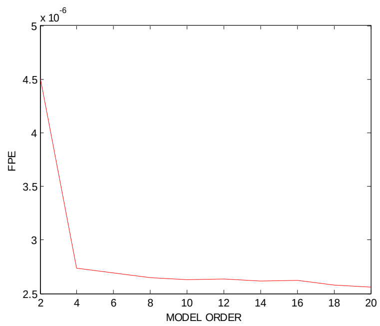

Fig.1. TheAkaike Final Prediction Criterion (FPE) v/s AR model order

A number of methods exist for computing AR coefficients (a_i). The main two categories are the least squares method and the Burg method. The most common least squares method is based on Yule Walker equations given below:

$$\gamma_m = \sum_{k=1}^{p} a_k \gamma_{m-k} + \sigma^2 \delta_{m,0}$$

The coefficients can be found by multiplying the definition above by x_{t-d}, taking the expectation values and normalizing gives a set of linear equations called the Yule-Walker equations that can be written in matrix form as

$$\begin{pmatrix} 1 & r_1 & r_2 & r_3 & r_4 & \cdots & r_{N-1} \\ r_1 & 1 & r_1 & r_2 & r_3 & \cdots & r_{N-2} \\ r_2 & r_1 & 1 & r_1 & r_2 & \cdots & r_{N-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ r_{N-1} & r_{N-2} & r_{N-3} & r_{N-4} & r_{N-5} & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \cdot \\ \cdot \\ \cdot \\ a_N \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \\ \\ \\ r_N \end{pmatrix}$$

$$Ra = r$$

Where autocorrelation coefficient at delay d is r_d . The diagonal is $r_0 = 1$.

2.3 Harmonic Model

Complex frequency components of a signal in the presence of noise can be estimated using frequency estimation. The number of components should be assumed.Let a signal X (n) be expressed as a sum of p complex sinusoids in the presence of white noise w_n.

$$X(n) = \sum_{i=1}^{p} A_i e^{jn\omega i} + w_n$$

Where A_i represents complex amplitude represented as $A_i=|A_i|e^{j\phi i}$

ϕi is uncorrelated and uniformly distributed over [$-\pi$ π].

The frequencies ω and magnitudes $|A_i|$ are not random, but they are unknown. The power spectrum of X(n) consists of a set of p impulses of area $2\pi|A_i|$ at frequencies ω_i for i=1,2,……,p plus the power spectrum of the white noise w_n.

The Eigen Vector method, Pisarenko's method, the Multiple Signal Classification (MUSIC) method and the Minimum norm methods are the most popular techniques used for noise subspace based frequency estimation.

2.3.1 MUSIC (MUltiple SIgnal Classification)

It is a technique used for frequency estimation. The MUSIC algorithm estimates the pseudo spectrum from the signal or a correlation matrix using Schmidt's Eigen space analysis method. If correlation matrix is not known, then estimation of the Eigen values and Eigen vectors is done.

This method assumes that p complex exponentials exist in a signal X (n), in the presence of white Gaussian noise. If R_x corresponds to autocorrelation matrix of the order M X M, with the Eigen values sorted in the descending order, the Eigen vectors corresponding to the p largest Eigen values will span the signal subspace. The remaining M-p Eigen vectors span the orthogonal space, where there is only noise. Mathematically this can be given as follows:

Let $\mathbf{v_i}$ and λ_i be the Eigen vectors and Eigen values of Rx. The Eigen values are arranged in descending order.

$$\lambda_1 \geq \lambda_2 \geq \ldots\ldots\ldots \geq \lambda_M$$

Let the variance of white noise be represented as σ_w^2 . The Eigen values of R_x are given as

$$\lambda_i = \lambda_i^s + \sigma_w^2$$

Where λ_i^s are the Eigen values of R_x

R_x is a matrix of rank p. Hence first p Eigen values of R_x will be greater than σ_w^2 and the last M-p Eigen values will be equal to σ_w^2. Hence, it can be noted that two groups of the Eigen values and the Eigen vectors of R_x exist: the signal Eigen vectors $v_1, v_2, \ldots, v_p$ that have Eigen values greater than σ_w^2 and the noise Eigen vectors $v_{p+1}, v_{p+2}, \ldots\ldots, v_M$ that have Eigen values equal to σ_w^2. The noise Eigen vectors are averaged for frequency estimation.

The MUSIC estimate is given by:

$$P_{MU}(e^{j\omega}) = \frac{1}{\sum_{i=p+1}^{M} |\mathbf{e}^H \mathbf{v_i}|^2}$$

2.4 Principal Component Analysis (PCA)

Information can be extracted from confusing data sets in a simple, non parametric way using PCA. It is a statistical procedure that uses an orthogonal transformation to convert a set of possibly correlated values into a set of values of variables which are not correlated, called Principal Components.

The linear combination of weighted observed variables can be called as Principal Component. Let us consider a data matrix of p variables with n samples. The first principal component Y_1 is given as a linear combination of variables X_1, X_2,….., X_p as follows:

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \ldots\ldots\ldots\ldots + a_{1p}X_p$$

Or in the matrix form as follows:

$$Y_1 = a_1^T X$$

Where a_{11}, a_{12}, ……., a_{1p} are called as the weights. The first principal component will have the highest variance in the data set. This can be done by choosing proper values for weights. Usually the weights are calculated with the following constraint:

$$a_{11}^2 + a_{12}^2 + \ldots\ldots + a_{1p}^2 = 1$$

Similarly second principal component is calculated, with the condition that it has no correlation with the first principal component, and accounts for the second highest variance.

$$Y_2 = a_{21}X_1 + a_{22}X_2 + \ldots\ldots\ldots\ldots + a_{2p}X_p$$

This continues until p numbers of principal components are calculated. In general, this can be represented in matrix form as:

$$Y = AX$$

The number of principal components obtained is equal to the number of variables being analysed. But in most of the analyses, only the first few components will have meaningful level of variance. Hence only first few components are used for future analysis. The Eigen vectors of covariance matrix are called the Principal Components, and hence orthogonal.

III. Results

The AR spectrum is calculated for the model orders 2, 4, 6, 8, 10 and 12. Fig 2(a) shows large variations in spectrum with model orders for a signal indicating CAD. Fig 2(b) shows spectrum plot of a normal heart signal.

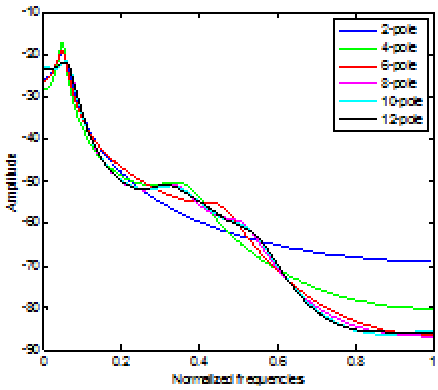

Fig. 2(a) AR spectrum for a CAD data

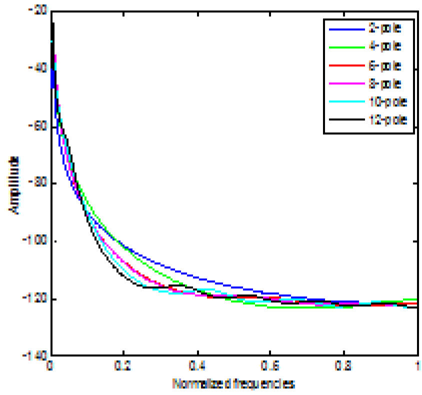

Fig. 2(b) AR spectrum for a non_CAD data

The location of AR poles differs for normal signals and CAD signals. The poles move closes towards the unit circle on the right half of the plane in case of CAD signals. The poles may sometimes be located outside the unit circle also in case of abnormal heart signals. This is shown in fig. 3(a) and 3(b).

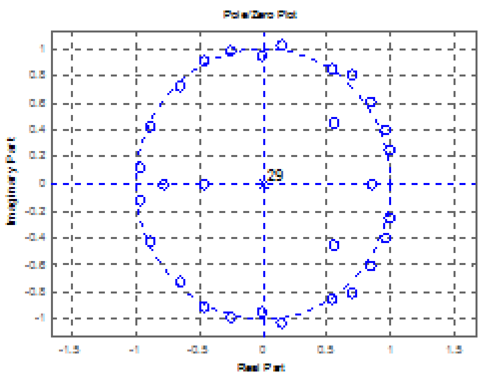

Fig. 3(a) AR poles for a CAD signal

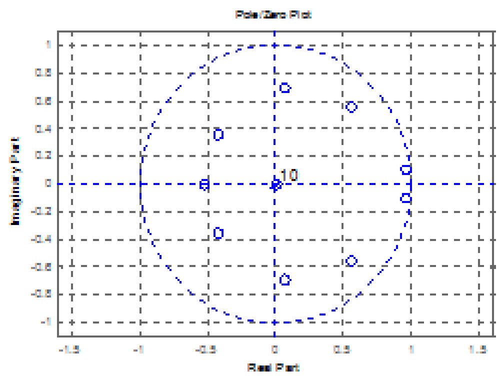

Fig. 3(b) AR poles for normal signal

The MUSIC method requires the number of complex exponentials to be specified. Hence analysis is done here for p=2, 4, 6, 8, 10, 12 signal components. Fig. 4 shows the power spectrum using MUSIC method assuming p = 4.

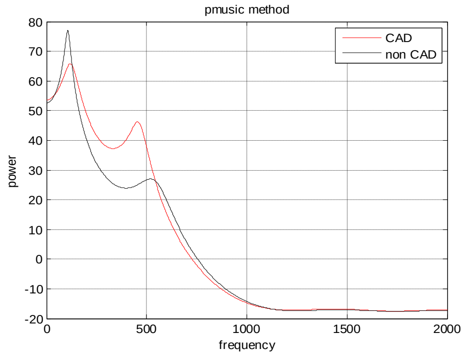

Fig. 4 Spectrum obtained for CAD and Non-CAD signals using MUSIC method

CAD signals shows extra peaks compared to normal signals. The power levels of the first and the second peaks are also different for normal heart signals and abnormal heart signals. Table 1 shows the ratio of second peak power (P2) to the first peak power (P1) in both the cases for some signals used in this study.

Table I: Ratio Of Power Levels P2 And P1 For Some Normal Signals And Cad Signals

P2/P1 for CAD signals	P2/P1 for normal signals
1.5749×10^{-6}	0.5095
2.59×10^{-5}	1.6654
2.84×10^{-6}	1.9282
3.977×10^{-5}	2.1830

In Fig. 5, the plot shows the variance obtained for four Principal Components. The Principal Component Analysis is used to reduce the amount of variables obtained using AR model poles and MUSIC system. Using one test signal having coronary artery disease gave four Principal Components with first component having a variance of about 52% as shown in Fig. 5

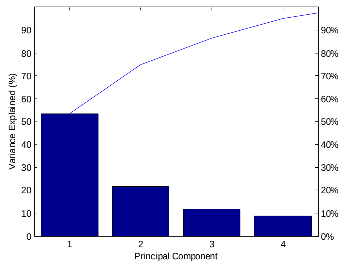

Fig. 5 Variances obtained by four Principal Components of a CAD signal

IV. Conclusions

In this study, the heart signal recorded using electronic stethoscope is used for finding the presence of coronary artery disease. The signal is split into three overlapping frequency bands and then segmented using duration dependant Hidden Markov Model. Signal is modeled using Yule-Walker AR model. The AR model is used with different model orders. The same signal is represented in harmonic model and then analysed using MUSIC method. PCA is used in order to reduce the total number of variables obtained using both of these methods. The results show that both AR method and MUSIC method are capable of distinguishing normal and abnormal heart signals. The presence of excessive noise during recording of heart sounds reduces the efficiency of the detecting system. It is necessary to eliminate noise completely before analysis which becomes impossible. Hence an alternative is to increase the number of parameters considered for analysis. The AR poles and MUSIC frequencies should be used with additional parameters for accurate and exact detection of heart disease using electronic stethoscope signals.

References

[1]. M. Akay, J. L. Semmlow, W. Welkowitz, M. D. Bauer and J. B. Kostis, Detection of coronary occlusions using autoregressive modeling of diastolic heart sounds, *IEEE Trans. Biomed. Eng.*, Apr, 1990,vol. 37, pp. 366-373.
[2]. Y. M. Akay, M. Akay, W. Welkowitz, J. L. Semmlow and J. B. Kostis, Noninvasive acoustical detection of coronary artery disease: A comparative study of signal processing methods, *IEEE Transactions on Biomedical Engineering*, 1993 vol. 40, pp. 571-578.
[3]. M. Akay, Y. M. Akay, D. Gauthier, R. G. Paden, W. Pavlicek, F. D. Fortuin, J. P. Sweeney and R. W. Lee, Dynamics of diastolic sounds caused by partially occluded coronary arteries, *IEEE Trans. Biomed. Eng.*, Feb, 2009, vol. 56, pp. 513-517.

[4]. S. E. Schmidt, C. Holst-Hansen, C. Graff, E. Toft and J. J. Struijk, Detection of coronary artery disease with an electronic stethoscope, *Computers in Cardiology*, 2007, pp. 757-760.

[5]. Schmidt SE, Hansen J, Zimmermann H, Hammershøi D, Toft E,Struijk JJ,Coronary artery disease and low frequency heart sound signatures, *Computers in Cardiology* , 2011,38:481–484

[6]. M. Akay, J. L. Semmlow, W. Welkowitz, M. D. Bauer and J. B. Kostis, Noninvasive detection of coronary stenoses before and after angioplasty using eigenvector methods, *IEEE Trans. Biomed. Eng.*,1990, vol. 37, pp. 1095-1104, 11.

[7]. Semmlow J, Rahalkar K,Acoustic detection of coronary artery disease, *Annu Review Biomedical Eng* ,2007, 9:449–469

[8]. Jin-Zhao, W., Bing, T., Welkowitz, W., Semmlow, J.L., Kostis,J.B., "Modeling sound generation in stenosed coronary arteries," *IEEE Transactions on Biomedical Engineering*, , 1990, vol. 37, pp. 1087-1094.

[9]. J. Semmlow, W. Welkowitz, J. Kostis and J. W. Mackenzie, "oronary artery disease--correlates between diastolic auditory characteristics and coronary artery stenoses, *IEEE Trans. Biomed. Eng.*,1983, vol. 30, pp. 136-139, Feb

[10]. Akay, Y.M., Akay, M., Welkowitz, W., Kostis,J., Noninvasive detection of coronary artery disease, *IEEE Engineering in Medicine and Biology Magazine*, 1994, vol. 13, pp. 761-764

[11]. M. Akay, Y. M. Akay, W.Welkowitz, J. L. Semmlow and J. Kostis, *Application of adaptive FTF/FAEST zero tracking filters to noninvasive characterization of the sound pattern caused by coronary artery stenosis before and after angioplasty*, *Ann. Biomed. Eng.* 1993, vol. 21, pp. 9-17

[12]. Ricke AD, Povinelli RJ, Johnson MT. *Automatic segmentation of heart sound signals using hidden Markov models*. Computers in Cardiology 2005;32:953-6,.

[13]. Wang P, Lim CS, Chauhan S, Foo JY, Anantharaman V. *Phonocardiographic signal analysis method using a modified hidden Markov model*. Ann.Biomed.Eng 2007; 35:367-74.

[14]. Lima CS, Cardoso MJ. *Phonocardiogram segmentation by using hidden Markov models*. BioMed 2007;5:415-8.

[15]. D. Gauthier, Y. M. Akay, R. G. Paden, W. Pavlicek, F. D. Fortuin, J. K. Sweeney, R. W. Lee and M. Akay, *Spectral analysis of heart sounds associated with coronary occlusions,* in Information Technology Applications in Biomedicine, 2007. ITAB 2007. 6th International Special Topic Conference on, 2007, pp. 49-52

[16]. D. Gauthier, Detection of coronary artery disease using an electronic stethoscope, M.S. thesis, Arizona State Univ., 2007

[17]. *3M™ Littmann® Electronic Stethoscope Model 4000*, Manual

[18]. C. M. Hurvich and C. Tsai, *Regression and Time Series Model Selection in Small Samples*, *Biometrika*, 1989, Vol. 76, pp. 297–307

[19]. Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd edition.Springer

[20]. Steven M Holland, *Principal Component Analysis*,University of Georgia, Athens, 2008