PHONOCARDIOGRAM BASED DIAGNOSTIC SYSTEM

Lubaib.P[1], Ahammed Muneer KV[2], Abdu Rahiman V[3]

[1]Department of Applied Electronics and Instrumentation, Govt. Engineering College, Kozhikode, India
lubaib.p@gmail.com
[2]Department of Applied Electronics and Instrumentation, Govt. Engineering College, Kozhikode, India
ahammedcet@gmail.com
[3]Department of Applied Electronics and Instrumentation, Govt. Engineering College, Kozhikode, India
vkarahim@gmail.com

ABSTRACT

A Phonocardiogram or PCG is a plot of high fidelity recording of the sounds and murmurs made by the heart with the help of the machine called phonocardiograph. It has developed continuously to perform an important role in the proper and accurate diagnosis of the defects of the heart. As usually with the stethoscope, it requires highly and experienced physicians to read the phonocardiogram. A diagnostic system based on Artificial Neural Networks (ANN) is implemented as a detector and classifier of heart diseases. The output of the system is the classification of the sound as either normal or abnormal, if it is abnormal what type of abnormality is present. In this paper, Based on the extracted time domain and frequency domain features such as energy, mean, variance and Mel Frequency Cepstral Coefficients (MFCC) various heart sound samples are classified using Support Vector Machine (SVM), K Nearest Neighbour (KNN), Bayesian and Gaussian Mixture Model (GMM) Classifiers. The data used in this paper was obtained from Michigan university website.

KEYWORDS

Phonocardiogram; K Nearest Neighbour; Support Vector Machine; Gaussian Mixture Model; Bayesian

1. INTRODUCTION

Cardiovascular diseases are among those seriously threatening human health. Heart sound analysis provides good information of these diseases, which is significant for diagnosing different kinds of cardiovascular diseases. In cardiac auscultation, an examiner uses a stethoscope to listen for these sounds, which give important information about the condition of the heart. Heart is one of the important organs in our body pumping blood continuously throughout the entire life-time. Heart is made up of a strong muscle called myocardium, and has four valves for regulating the blood circulation [5]. It beats almost regular intervals and is controlled by the electrical pulses generated from the sinus node near the heart. The rhythmic beating of heart produces a sound referred to as “Lub-Dub” due to the closing of the atrio-ventricular and the aortic-pulmonic valves. Moreover, there are other sounds, which are due to the structural and functional defects of the heart, called as murmurs. Stethoscope is still used as the primary auscultation device for heart and lung sounds. However, analysis based on heart sounds using a stethoscope are depends on the doctor’s experience and hearing ability.

The purpose of this paper is to develop an algorithm for Study and Design of Peak Spacing Analysis for Estimating Arrhythmic Heart-Beat, Based on the algorithm, and every cycle of the PCG signals is separated into four parts: the first heart sound, the systolic period, the second heart sound and the diastolic period. Based on this information, the intervals of the systolic and diastolic period are obtained consequently. Then both the systolic and diastolic periods are analysed separately and at the same time in a combined manner. After that the heart signal are analysed whether the sound is arrhythmic or not. If it is arrhythmic then of which type whether it is overall arrhythmic or the systolic and diastolic periods are arrhythmic.

2. BACKGROUND

Hippocrates (460-377 BC) provided the foundation for auscultation when he put his ear on the chest of a patient and described the sounds he could hear from it. The next step was made by Robert Hooke (1635-1703) who realized the diagnostic use of cardiac auscultation. Although Dr. Laennec's invention, the stethoscope, has been in clinical use for more than 180 years, and electronic stethoscopes with variable amplification gain have been available for over 80 years, it is still difficult to understand auscultation findings. The phonocardiogram, first developed in 1894, visualizes auscultatory signals. The spectral phonocardiogram has proven to be a reliable tool that gives information of whether or not the murmur is pathological. Phonocardiography and electronic stethoscopy attempt to improve the diagnostic accuracy of cardiac auscultation [7]. In the most recent studies, digital acoustic analysis has demonstrated the validity of these methods. Since the 1980's, phonocardiographic research activity had decreased due to the improvements of echocardiography, which yields more visual information. During the past few years, however, the improvements of personal computers have made it possible to design new low-cost, high quality phonocardiographic devices.The phono-spectrogram combines traditional phonocardiogram with time-frequency distribution presentation of the signal. The spectrogram was introduced for heart sound analysis as early as 1955 by McKusik et al, but was afterwards almost forgotten. The stethoscope (from the Greek word stethos , meaning “chest” and skopein, “to examine”)invented during the early twentieth century, was one of the most primitive devices designed to aid a doctor in listening to heart sounds. It was a simple tool used for transmitting the sound energy from the chest wall of the patient to the ear of the physician via a column of air. The 20th century witnessed extraordinary advances in the diagnosis of heart diseases, corrective surgeries, and other forms of treatment for heart problem [8].

3. METHODOLOGY

The Phonocardiogram (PCG) could be used as a simple but effective source for analysing and detecting the arrhythmic nature of heart beat pattern. The heart rhythm is directly related to the electrical activity of the heart as well as its valve movements. The data used in this paper was obtained from Michigan university website [1]. The following methodology has been employed to study the effectiveness of PCG in cardiac diagnosis.

3.1. Pre- processing

The recorded heart sound signal is required to pre process [6] before envelope extraction. The heart sound signals were normalized according to the equation (1) as shown below.

$$x_{norm}(t) = \left(\frac{x(t)}{\max\ |x(t)|}\right)^2 \tag{1}$$

Where *x(t)* is the original signal and the square operation leads to make peak signal more prominent and weaken the noise.

3.2. Envelope detection

The envelope of the heart sound signal can be detected using different methods [6] like absolute value of the signal, squared energy, etc.

$$\begin{aligned} &\text{Absolute Value} \quad E = |x| \\ &\text{Square Energy} \quad E = x^2 \end{aligned} \tag{2}$$

The squared energy method based on exponential weighing factors to high intensity components. The absolute value technique associates same weighing factor to all components, which is difficult to separate low from high amplitude signals.

Figure. 1 shows the energy based envelop of a simple PCG signal which is convenient to find S1 and S2 locations.

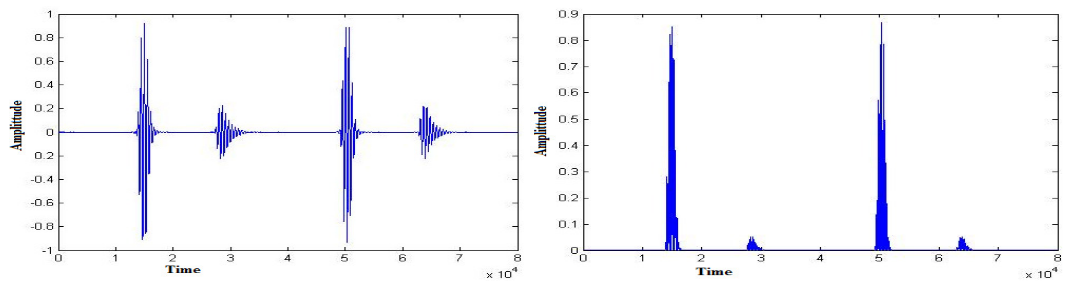

Figure 1: Energy based envelop of a simple PCG signal

3.3. Separation of Systole and Diastole

In a small uni-variate time-series, it is easy to identify peaks; there is a need to formalize an algorithm to automatically detect peaks in any given time-series. A data point in a time-series is a *local peak* if it is a large and locally maximum value with respect to all other data points in a window. By identifying peak-locations and peak-amplitude values in a phonocardiogram wave located the S1 and S2 peaks. It has been observed that S1-peak values are normally larger than S2-peak values. The time-gap between consecutive S1 and S2 peaks represents systole period whereas diastole period is measured as the time-gap between consecutive S2 and S1 peaks [6]. Every systole and diastole periods are recorded along the entire sample sequence for further analysis. The time gap between two S1 peaks called cardiac cycle, which includes one systole and diastole, which are clearly illustrated in Figure. 2

From the separation of systole and diastole it is easy to calculate systolic period, diastolic period, heart rate etc. Then total phonocardiogram wave of one minute duration is split into many cardiac cycles [9]. Every systole and diastole periods, cardiac cycles are recorded along the entire sample sequence for further analysis.

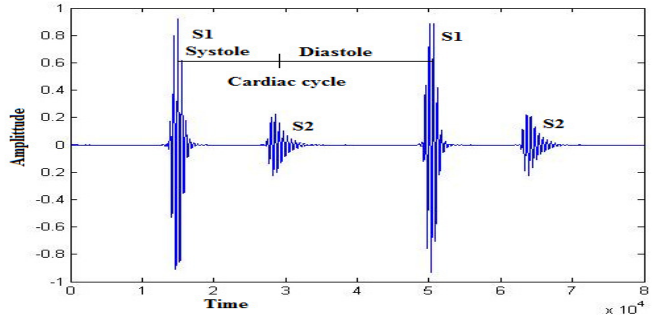

Figure 2: Systole and Diastole Period with reference to S1 and S2 peaks.

3.4. Feature extraction

Automatic extraction of features [2], [3], [4], [9], [10] depends on accurate knowledge about the timing of the heart cycles. Segmentation into the first heart sound (S1), systole, the second heart sound (S2) and diastole is thus needed. In this paper, both time domain and frequency domain features are extracted.

3.4.1 Time domain features

The envelope of heart sound was extracted with the normalized average energy, the heart sound signal was divided into each cardiac cycles. Each cardiac cycle was divided into short overlapping segments of fixed and equal duration. Energy was calculated in each segment and which is recorded and plotted as shown in Figure.3. Energy of corresponding segments in each cardiac cycle is stored in an array, and calculated their mean and variance.

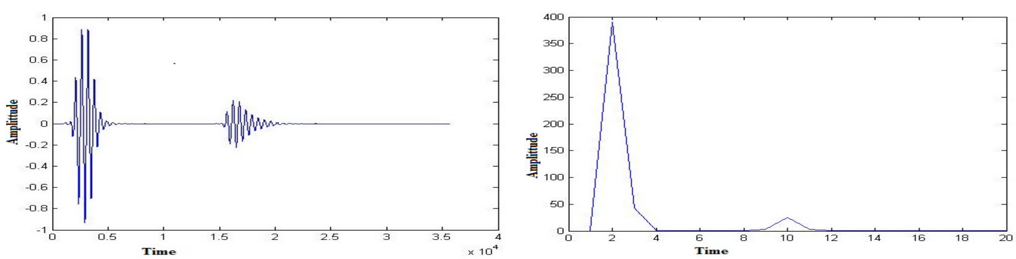

Figure.3 One cardiac cycle and its energy

3.4.2 Frequency domain features

There are so many transforms available to obtain frequency domain features [3], [4], [9]. In this paper Mel Frequency Cepstral Coefficients (MFCC) [2], [4] are used as frequency domain features. MFCCs are a way of representing the spectral information in a sound. Each coefficient has a value for each frame of the sound. The changes within each coefficient across the range of the sound are examined here. Obtaining the MFCCs involves analyzing and processing the sound according to the following steps

- Divide the signal into frames
- Obtain the amplitude spectrum of each frame
- Take the log of these spectrums
- Convert these to the Mel scale
- Apply the Discrete Cosine Transform (DCT)

The computation of the MFCC includes Mel-Scale filter-banks [4], as shown in Figure .4 .The Mel-Scale filter-banks are computed as follows[2], [4]:

$$m = 1127 log_e\left(\frac{f}{700} + 1\right) \tag{3}$$

Where f is the frequency in the linear scale and m is the resulting frequency in Mel-Scale.

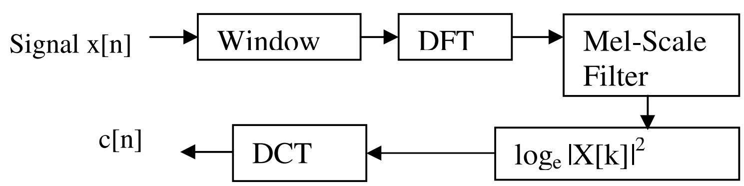

Figure. 4. MFCC feature extraction process.

The power spectral density (PSD) of the spectrum is mapped onto the Mel-Scale by multiplying it with the filter-banks constructed earlier and the log of the energy output of each filter is calculated as follows:

$$s[m] = log\left(\sum_{k=0}^{N-1} |X[k]|^2 H_m\ [k]\right) \tag{4}$$

Where $H_m[k]$ is the filter-banks and m is the number of the filter-bank. Finally, to obtain the MFCC the discrete cosine transform (DCT) of the spectrum is computed:

$$c[n] = \sum_{m=0}^{N-1} S[m] \cos\left(\frac{\pi n}{M}\left(m - \frac{1}{2}\right)\right), n = 0,1,2 \ldots\ldots M \tag{5}$$

Where M is the total number of filter banks.

3.5 Classification

3.5.1 support vector machine (SVM)

After the feature extraction process, support vector machine [2], [13] was used to classify the data. Support vector machine is in fact a Two-class classifier in which a linear boundary is used to divide the classes. In this method, samples closest to the decision boundary are called the Support vectors. These vectors define the decision boundary equation. Each sample is shown as a vector. To find the optimal decision boundary, the maximum margin method is used. So, the decision boundary in addition to all instances of both classes should not only properly divide all the samples into two categories, in this method it is assumed that the samples hold $y_i = \{-1, +1\}$ label. The mathematical expression of the decision boundary in vector space can be expressed as the following equation:

$$f(\bar{x}) = sgn(\bar{w}.\bar{x} + b) \tag{6}$$

Where $\bar{w}$ is the normal vector of the hyper plane and b is the intercept. The decision making boundary should accurately classify the samples as the following equation:

$$y_i(\bar{w}.\bar{x} + b) \geq 1 \tag{7}$$

The boundary of decision making should have the most distance from the samples of each class as maximizing $\frac{2}{\bar{w}}$ according to Figure.5 shown below

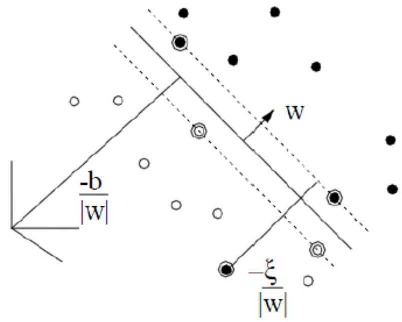

Figure.5 Hyper plane separation

We can define an optimization problem as follows:

$$min\frac{1}{2}\ ||\bar{w}||^2 \quad (8)$$

$$s.t.\ y_i(\bar{w}.\bar{x}+b) \geq 1 \quad (9)$$

The method of Lagrange multipliers [2], [13] is used to solve the problem of optimizing, which is

$$\min max\left\{\frac{1}{2}||w||^2 - \sum_{i=1}^{n} \propto_i [y_i(w.x_i+b)-1]\right\} \quad (10)$$

Karosh-Kuhen- (KKT) state [2],[13] that the optimal value is shown as

$$w = \sum_{i=1}^{n} \propto_i x_i y_i \quad (11)$$

The value of b is calculated through equation (12)

$$b = \frac{1}{N_{sv}} \sum_{i=1}^{n} y_i - \bar{w}.\bar{x}_i \quad (12)$$

Where N_{sv} is the number of support vectors.

3.5.2 K- Nearest Neighbour (KNN)

The K Nearest Neighbour (KNN) [13] method is a widely used technique in clustering and classification applications. In this method given that a set of *N* points (training set), whose class labels are known, classify a set of *n* points (testing set) into the same set of classes by examining the *k* closest points around each point of the testing set and by applying the majority vote scheme. There are different classification approaches based on the KNN such as the Density Based KNN Classifier (DBKNN), Variable KNN Classifier (VKNN), Weighted KNN Classifier (WKNN) and Class Based KNN Classifier (CBKNN).

Structural Density (SD) is defined as the number of points in the neighbourhood of an element over the volume of this neighbourhood. Radius (*r*) is the parameter involved in defining the neighbourhood, in this method. First, we calculate the density of all the elements as a function of *r* and the average density of the whole set. We then search for a value of *r* such that the mean of the individual densities is equal to the average density calculated earlier.

In Variable KNN classifier, a classification of it is performed based on various neighbourhoods for each one of the training set elements the *k* value that maximises the *DC* of each classification is found. Therefore, for each training set there corresponds a particular *k* value which is considered the best available. Afterwards, for each unknown element, the nearest neighbour is found and its *k* value is assumed. Then, the KNN classifier is applied on that test element, using that *k* value.

In Class Based KNN, the *k* nearest elements of each class is taken for every test element. The value of *k* is automatically selected by the classifier, so as to maximise the *DC* of the classification. Afterwards, the harmonic mean of the distances of these neighbours is calculated. Finally, these means are compared and the class yielding the lowest value is chosen for the classification.

3.5.3 Bayesian Decision Theory

Bayesian decision theory [13] is a fundamental statistical approach to the problem of quantifying pattern classification. This approach is based on decisions using probability Bayes formula [13] can be expressed informally.

*Posterior = (likelihood * prior)/evidence*

There are many different ways to represent pattern classifiers. One of the most useful is in terms of a set of discriminant functions $g_i(x)$, $i = 1,\ldots c$. The classifier is said to assign a feature vector x to class w_i if $g_i(x) > g_j(x)$ for all $j \neq i$. The minimum error rate classification can be achieved by use of the discriminant functions

$$g_i(x) = ln\, P\left(\frac{x}{w_i}\right) + lnP(w_i) \tag{13}$$

This expression can be readily evaluated if the densities $p\left(\frac{x}{w_i}\right)$ are multivariate normal, In this case, then,

$$g_i(x) = \left(\frac{1}{2}\right)(x-\mu_i)^T \sum\nolimits_i^{-1} (x-\mu_i) - \left(\frac{d}{2}\right) ln2\pi - ln|\Sigma| + lnP(w_i) \tag{14}$$

3.5.4 Gaussian Mixture Model (GMM)

Gaussian Mixture Model (GMM) [11] belongs to the stochastic modelling and it is based on the modelling of statistical variations of the features. The GMM model is trained for each person and a model is generated for each user which contains information based on statistical processing of data. During testing, user is identified based on maximum probability criteria. In the first step we choose the number of component densities required to specify a user. GMM algorithm is named based on number of component densities. For example, if the number of component densities is four, the GMM is called as GMM-4. Similarly, if the number of components is eight, the GMM is called as GMM-8.

Mixture density in GMM is the weighted sum of *M* component densities and is given by the equation,

$$p(\bar{x}|\lambda) = \sum_{i=1}^{M} p_i b_i(\bar{x}) \tag{15}$$

Where, x is a D-dimensional feature vector, $b_i(\bar{x})$, i=1,2....M, are the component densities and p_i ,1,2....M are the mixture weights. Each component density is a Gaussian function given by [11],

$$b_i(\bar{x}) = \frac{1}{2\pi^{\frac{D}{2}}|\Sigma i|^{\frac{1}{2}}} exp\left\{-\frac{1}{2}(\bar{x}-\bar{\mu_i})' \sum_i^{-1} (\bar{x}-\bar{\mu_i})\right\} \quad (16)$$

Where, μ_i is the mean vector and Σ_i is the covariance matrix.

The goal of GMM model is to estimate the parameters λ of the GMM from the training heart signals feature vectors. The most popular and well-established method for estimating the parameters of GMM is maximum likelihood (ML) estimation [12]. ML estimation finds the model parameters for the given training data and maximizes the likelihood of the GMM [11] For a given T training vectors $X = \{x_1 , x_2 ,\ldots x_T\}$, GMM likelihood is given by

$$p(x|\lambda) = \prod_{t=1}^{T} p(\bar{x}_t|\lambda) \quad (17)$$

This is a nonlinear function of the parameter λ and hence, direct maximization is not possible. Hence we estimate ML parameters iteratively using an algorithm known as Expectation-Maximization (EM) algorithm [12].

4. RESULT AND DISCUSSION

By the methodology mentioned above heart sounds are analyzed. From the energy envelop and S1, S2 peaks, heart beat per minute is calculated and also the peak-locations and peak-amplitude values are analyzed to identify S1 and S2 peaks in the test signal. Experimentally it has been observed that S1-peak values are normally larger than S2-peak values. The time-gap between consecutive S1 and S2 peaks represents systole period whereas diastole period is measured as the time-gap between consecutive S2 and S1 peaks, and also it has been observed that the period of Systole gap is always less than the diastole gaps, and before the time domain and frequency domain feature extraction it has been divided into each cardiac cycle containing one systole and one diastole.

In this paper twenty three heart sounds from Michigan university website are studied, which consisting normal and abnormal heart sounds of different types shown in Figure.6.

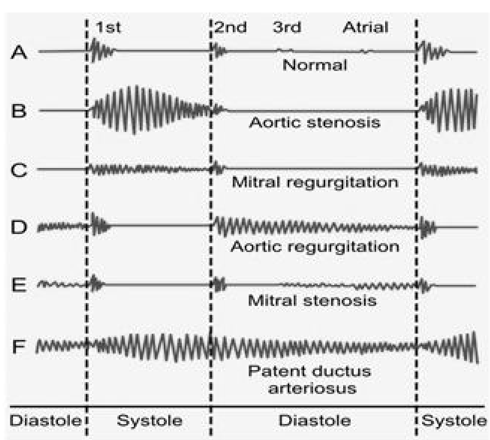

Figure.6.Phonocardiograms from normal and abnormal heart sounds

The result of these experiments will be based on classification accuracy,

$$Accuracy = \frac{C_n}{T_n} * 100 \quad (18)$$

Where C_n is the number of correct classifications and T_n is the total number of testing samples.

MFCC are extracted frequency domain features, based on these features KNN and SVM classifier is used to classify whether these signals are normal and abnormal. Mean and variance of average energy are extracted time domain features, based on these features Bayesian and GMM classifier is used to classify whether these signals are normal and abnormal, and more over which is used to check what type of abnormality is present. Classification accuracy of each method is shown in Table.1 and Table.2

Table.1Acceuracy obtained using MFCC features

Classifier	Case	Accuracy
SVM	Linear kernel	98.97%
	Polynomial kernel(d=2)	100%
	Polynomial kernel(d=3)	100%
KNN		99.80%

Table.1Acceuracy obtained using time domain features(mean and variance of energy)

Classifier	Case	Accuracy
Bayesian	Case-1 ($\Sigma_i=\sigma^2 I$)	98.98%
	Case-2($\Sigma_i= \Sigma$)	100%
	Case-3(Σ_i=arbitrary)	99.01%
GMM		99.98%

5. CONCLUSIONS

By the processing steps mentioned above heart sounds are processed and analyzed. From the energy envelop and S1, S2 peaks, heart beat per minute is calculated**.** Experimentally it has been observed that S1-peak values are normally larger than S2 peak values and also it has been observed that the time period of Systole gap is always less than the diastole gaps. It is concluded that the abnormality of heart sound is not only depend on the overall cardiac cycle whereas it depends upon the individual analysis of systole and diastole period. By feature extraction of each cardiac cycle and classification mentioned above, it is concluded that whether the signal is normal or abnormal, and what type of abnormality present. And SVM with polynomial kernel (d=2 and d=3) and second case of Bayesian classifier obtained 100% accuracy.

ACKNOWLEDGEMENTS

This paper will be incomplete without mentioning all the people who helped me to make this possible, whose encouragement was invaluable. The authors would like to thank Mr.Abdul Rahiman (Asst. Professor, AE&I), Prof. Shajee Mohan B.S, Associate Professor, Department of Applied Electronics & Engineering for their committed guidance, valuable suggestions and encouragement.

REFERENCES

[1] http://www.med.umich.edu/lrc/psb/heartsounds/

[2] Behnam Farzam, Jalil Shirazi " The Diagnosis of Heart Diseases Based on PCG Signals using MFCC Coefficients and SVM Classifier" IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 10, December 2014.

[3] X. J. Hu, J. W. Zhang, G. T. Cao, H. H. Zhu, and H. Li, "Feature extraction and choice in PCG based on hilbert transfer," in Proc. 4th International Congress on Image and Signal Processing (IEEE), 2011.

[4] Jithendra Vepa," Classification of heart murmurs using cepstral features and support vector machines" IEEE, 2009

[5] P. de Vos, and M. M. Blanckenberg, "Automated Pediatric CardiacAuscultation", IEEE Trans. Biomed. Eng., Vol. 54, No. 2, (Febr. 2007), pp. 244-252

[6] Study and Design of a Shannon-Energy-Envelope based Phonocardiogram Peak Spacing Analysis for Estimating Arrhythmic Heart-Beat,Praveen Kumar Sharma,Sourav Saha,Saraswati Kumari , International Journal of Scientific and Research Publications, Volume 4, Issue 9, September 2014 1 ISSN 2250-3153G.

[7] Heart Sound Analysis for Diagnosis of Heart Diseases in Newborns ,Amir Mohammad Amiri*, Giuliano Armano ,University of Cagliari, Department of Electrical and Electronic Engineering(DIEE), 09123 Cagliari, Italy ,ICBET 2013: May 19-20, 2013, Copenhagen, Denmark.

[8] S. Cerutti, A. L. Goldberger, and Y. Yamamoto, "Recent Advances in Heart Rate Variability Signal Processing and Interpretation", IEEE Trans. Biomed Eng., Vol. 53, No.1 (Jan. 2006), pp. 1-3

[9] C. N. Gupta, et al., "Segmentation and Classification of Heart Sounds," CCECE/CCGEI, Saskatoon, IEEE, 1674-77, May2005.

[10] Jian-bo Wu, Su Zhou, "Research on the Method of Characteristic Extraction and Classification of Phonocardiogram", International Conference on Systems and Informatics, 2012.

[11] D.A. Reynolds, R.C. Rose, Robust text-independent speaker identification usingGaussian mixture speaker models, IEEE Trans. Speech Audio Process, 3 (1995) 72–83

[12] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Royal Srar. Soc., vol. 39,pp. 1-38, 1977. A text book of biological signal analysis, Dr.Ramaswami Palaniappan,university of Essex,United Kingdom

[13] Symon O Haykin "Neural Networks and Learning Machines", Third Edition, PHI, 2010

Authors

Lubaib P has received B. Tech degree in Electronics and Communication from KMCT College of Engineering, Kozhikode, Kerala, India, affiliated to University of Calicut (2012). Currently he is pursuing his M.Tech in Signal Processing at Government Engineering College, Kozhikode, Kerala, India, affiliated to University of Calicut. He is worked as Guest Lecturer at Government Women's Polytechnic College, Kottakkal, Kerala, India. His research interest includes Digital Signal Processing, Communication System.

Ahammed Muneer K V received the B.Tech degree in Electronics and Communication Engineering from College of Engineering Trivandrum, Kerala, India in 2002 and completed his post graduation-M.Tech in Control and Instrumentation Systems from Indian Institute of Technology, Madras, India in 2012. He has been working as Assistant Professor in Electronics Engineering in various universities of Kerala, India and has more than ten years of teaching experience. He has got the best M.Tech project award from Indian Institute of Technology Madras in 2012 for his thesis work. He recently published a paper in IEEE-EMBS international conference titled "Non-contact ECG recording Instrument for Continuous Cardiovascular Monitoring". Presently he is working as Assistant Professor in Govt. Engineering College Kozhikode, Kerala, India. His research area includes Electronic instrumentation and sensors, signal processing etc. He is pursuing his PhD degree in National Institute of Technology, Calicut, India and the current research focus is on biomedical image processing.

Abdu Rahiman V received the B.Tech degree in Electronics and Communication Engineering from Goverment College of Engineering, Kannur, Kerala, India in 1998 and completed his post graduation-M.Tech in Signal Processing from College of Engineering Trivandrum, Kerala, India in 2007. He has been working as Assistant Professor in Electronics Engineering in various universities of Kerala, India since 2009. Presently he is working as Assistant Professor in Govt. Engineering College Kozhikode, Kerala, India.He is pursuing his PhD degree in National Institute of Technology, Calicut, India and the current research focus is on Signal processing.