Enhanced Performance of PCG Signal using Effective Feature Extraction Method

G. Venkata Hari Prasad, Lakshmi Narayana Thalluri

***Abstract:** Phonocardiography (PCG) is the realistic portrayal of sounds created in the heart auscultation. PCG is an improvement for ECG. Particularly in observing of patient and biomedical research, these signals need to do the diagnosis. This paper deals with the processing of heart sound signals i.e., Phonocardiography (PCG) Signals. The primary goal of analyzing these heart sound signals is to separate the signals from the noisy background and to extract some parameters which are used for patient monitoring and for other researches. Various momentum explore ventures are going on biomedical signal processing and its applications. The performance of the PCG signal will comprise of sectioning the signal into S1 and S2 and then compare, whether the PCG is normal or abnormal. In the previous framework the different change approaches are utilized to break down the PCG signal.In the primary stage, for include extraction; acquired heart sound signals were isolated to its sub-groups utilizing discrete wavelet change with Level-1 to Level-10. This upgraded strategy proposes a best component for Heart Signal Features, which are removed and changed in to other area to arrange signals. This enhanced method proposes a best feature for Heart Signal Features, which are extracted and transformed in to other domain to classify signals. In the proposed strategy the Wavelet is utilized for highlight extraction and different Statistical strategies are utilized. InformationGain (IG), Mutual Information (MI) and so on. Feature selection techniques are compared using classifiers like kNN(k-Nearest Neighbor), Naïve Bayes, C4.5 and Support Vector Machines (SVMs). MATLAB & WEKA Soft wares are used for analysis Purpose. In this paper, coiffelet technique is utilized to analyze the synthetic PCG and the classifier parameters are compared with one another.*

***Keywords:** Heart Sounds, Wavelets, Feature Extraction, Mutual Information, Information Gain (IG).*

I. INTRODUCTION

Biomedical signal processing is a potential research area and because of its flexibility it is adopted in many applications. The major applications are, design of monitoring systems to analyze the operation of human body using artificial limbs and aids. In Biomedical Applications heart sounds are significant component. Tracking sound is easier than other measures for diagnosis of human heart. Heart auscultation, i.e., the act of listening with a stethoscope to sounds delivered by the heart, lungs and blood, is an essential part of the physical examination and can provide information about the rhythm, rate, anatomical defects, valve functioning of the heart [1].Theheart functions by repeating cycles of systole and diastole. During systole the heart chambers eject theblood, while during diastole the heart chambers fill with blood[2].

Revised Manuscript Received on July 15, 2020.

Dr. G. Venkata Hari Prasad, Professor, Department of Electronics & Communication Engineering, CMR College of Engineering & Technology, JNT *University*, Hyderabad, India. E-mail: drgvenkatahariprasad@cmrcet.org

Dr. Lakshmi Narayana Thalluri, Assistant Professor, Department of Electronics & Communication Engineering, Andhra Loyola Institute of Engineering Technology (ALIET), Vijayawada, A.P, India.

A phonocardiogram or PCG is plot of recording of the sounds and murmurs made by the heart with the assistanceof phonocardiograph, which sounds resulting from vibrations created by closure of the heart valves namely Atrio Ventricular, Pulmonary & Aortic valves. PCG acquisition methods are plain, non-invasive, low-cost and accuratefor analyzing a wide range of heart disease like mitral value insufficiency aortic regurgitation etc.,. Advanced signal processing helps a lot to explicate the PCG into convenientformat.

A. Normal &irregular Heart sounds

By Phonocardiogram assessment ordinary heart sounds are identified into four according to the pattern in which they show up 1st, 2nd, 3rd and 4th heart sound with 1stand 2nd heart sounds heard usually**.** A few kids and youngsters carving 3rd heartsounds are sometimes heard [3]. 4th heart sound normally may be heard in 40 years of age or in above people.

B. The activity of normal heart sounds and their features:

i) First heart sound: The beginning of ventricular withdrawal isgenerally the mitral and tricuspid vibrations brought about by the abrupt end. First heart sound obtuse tone down, it sounds like "Le " sound much louder in the apex than the second heart sound high by 50 Percentage, the bottom of the second heart sound mind equal to or less than the second heart sound lasted over long, usually 0.10 - 0.16 s, at the apex and apical area and left sternal border between the loudestauscultation

ii) Second heart sound: The beginning mainly caused by ventricular diastolic pulmonary valve and aortic valve vibration caused. It sounds like "De" strong of its quality as the primary heart sound, shorter time than the main heart sound, normally 0.08 - 0.12 s in the aortic valve zone and aspiratory valve zone most intense straight prostrate auscultation.

iii) Third heart sound: Sometimes after the start of the resulting heart sound 0 .12 - 0.20s can hear the sound of a short and feeble called third heart sound. This is because of ventricular quick filling sound when the vibration of the ventricular divider. Third heart sound muted tone sounds powerless, dull and sloppy, auscultation was "Le - De - He" sound appears to reverberate the subsequent heart sound. Span is short, commonly 0.03 - 0.8s in the wake of showing up in the subsequent heart sound tone 0.12 - 0.18s.

iv) Fourth heart sound: Occurs in late diastole before the beginning of the primary heart sound 0.07 - 0.10s, which is because of atrial ventricular end-diastolic pressure in overcoming the force Generated by the vibration of contraction.

This sound is very weak in normal circumstances, the general heard.

II. EXISTING METHOD

In the existing research paper the improvement and analysis of the PCG is done by identifying the PQR values, de-noising the PCG signal by DWT method [3]. The wavelet technique subject to three thresholding limits is used for heart sound signals de- noising. A complete heart sound investigation framework covering division of beat cycles to definite assurance of heart conditions was introduced by Kao and Wei. The distinction in wavelet methodology can give different results. The hard-thresholding and its shortcoming discontinuous function performance can be improved by removing the permanent bias in soft-thresholding. The authors proposed a function to reduce the noises in the PCG signal simply by incorporating distinctive verity of wavelets. Which results the selection of appropriatewavelet that is best for a complex signal. Using the methods the functionality as the normality, abnormality is verified in the existing methodologies.

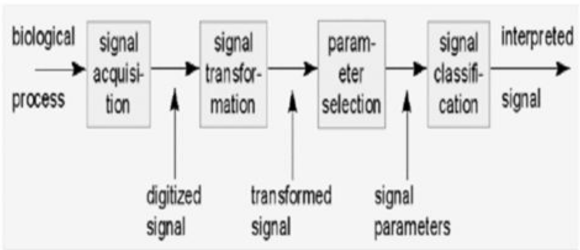

Fig. 1.The PCG generation from a human body

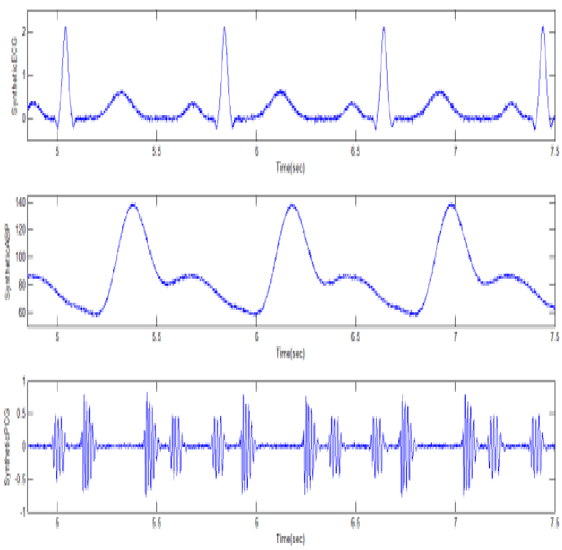

Fig. 2. Synthetic Electrocardiogram (ECG), Arterial blood pressure (ABP), PCG Waveforms

The basic processing stages of PCG signal is as shown in Fig. 1. The Fig. 2 and Fig. 3 shows Synthetic or System evaluated PCG signals. The complete information in the existing system uses the feature extraction method mainly. The feature extraction through various methods likeEmpirical Mode Decomposition (EMD), Singular Value Decomposition(SVD), and Independent Component Analysis (ICA) and finalize the normality and abnormality of the PCG signal [4].

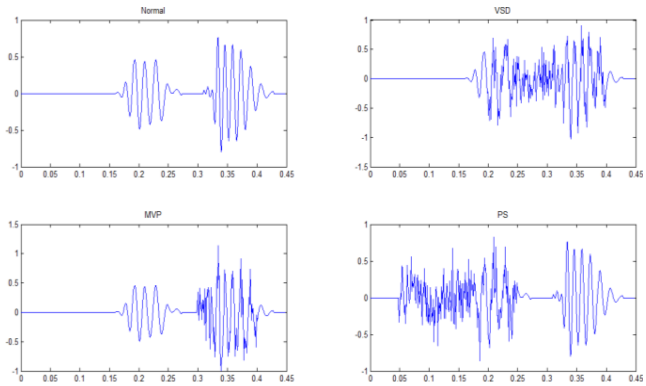

Fig. 3. PCG Normal, anomalous [(VSD) Ventricular Septal Defect, (MVP) Mitral Valve Prolapse and (PS) Pulmonary Stenosis] System Evaluated Signals

III. METHODOLOGY

A. The Pre-Processing

Filtering means essential to remove the unwanted noise. In general the noise can be expelled using analogue (or) digital signal processing methods. Especially the PCG signal is a sensitive and weak so easily interfere with noise. To avoid these unwanted interferences, it is needed to incorporate the differential amplifiers andfilters.

B. Feature Extraction

Feature extraction methods are used to map the real signals into lower dimension feature vectors. Feature extraction is to compute and identifying parameters from each cardiac cycle. Feature extraction finds a transformation that converts the original heart signal into a feature space preserves the information required for the application and enables meaningful comparisons. The first step in the characterization of the PCG signals is to compute the multi-resolution decomposition of the signal by using of the wavelet transform [1].

i) Discrete Wavelet Transform

A wavelet allows one to do multi resolution analysis; it helps to achieve both time & Frequency localization [5].Atime-frequency approach, Discrete Wavelet Transform (DWT) was utilized in numerous applications. DWT effectively inspected non-stationary signals. But, DWT yields a high dimensional feature vector that increases classifier’s learning parameters. So, a method to lower feature vector dimensionality that preserves classification accuracy must beused.

High pass filter output is set of DWT wavelet coefficients associated with all discrete wavelets at smallest single scale available for specific digitized neuro-electric waveform that goes into the filter. The output captures all high frequency waveform energy. The low pass filter output is set of DWT coefficients associated with companion functions set called scaling factor which captures all waveform’s low frequency energy.

Full DWT for signal *x (t)* is represented regarding a shifted scaling function version $\varphi_{j,k}$ and a shifted/dilated version of a mother wavelet function $\psi_{j,k}$.

The DWT representation is can be written as:

$$x(t)=\sum_{k\in Z} u_{j0,k}\phi_{j0,k}(t)+\sum_{j=-\infty}\sum_{k\in Z} w_{j,k}\psi_{j,k}(t) \quad \text{.....(I)}$$

Where $w_{j,k}$ are wavelet coefficients and $u_{j,k}(j<j0)$ are scaling coefficients.

C. Feature Selection

Feature selection is important, with many statistical classification methods and machine learning techniques being applied for this recently.

1) Mutual Information

Mutual Information is two random variables in information theory and in this context it is measure to evaluate features. It should be termed point wise mutual information.MI measures a document feature value ie presence/absence of a word which gives about document class. Mutual information is a criterion in statistical language modeling of word associations and related applications. MI between random variables X and Y, I(X; Y), measures information in X that is predicted when Y is known. If X and Y arecontinuous,

$$I(X;Y) = H(X) - (H(X \mid Y)$$

$$= \int P_{XY}(x,y) \log [P_{XY}(x,y)/P_X(x)\, P_Y(y)]\, dxdy \quad \text{.....(II)}$$

Where H(X) is entropy of X, which is a measure of its uncertainty, H(X/Y) is conditional entropy, which represents uncertainty in X after knowing Y.

2) 2) Information Gain

Information Gain is information obtained for category prediction by knowing the presence/absence of a document term. Let $m\{C_i\}_{i=1}^{m}$ denote set of categories in target space. Information gain of term t is defined:

$$G(t) = -\sum_{i=1}^{m} p_r(c_i)\log p_r(c_i) + p_r(t)\sum_{i=1}^{m} p_r(c_i \mid t)\log p_r(c_i \mid t) + p_r(\bar{t})\sum_{i=1}^{m} p_r(c_i \mid \bar{t})\log p_r(c_i \mid \bar{t})$$

--- (III)

Information gain is computed for every word of training set and words whose information gain is less than a predetermined threshold are removed.

IV. CLASSIFICATIONALGORITHMS

A. Naïve Bayes

Naïve Bayes classifier is a Bayes Theorem based classification method. Let C_j denote that an output belongs to j-th class, *j=1,2,...,J* out of J possible classes. Let $P(C_j \mid X_1, X_2, ..., X_p)$ denote (posterior) probabilityof belonging in j-th class given individual characteristics $X_1, X_2, ..., X_p$. Further, let $P(X_1, X_2, ..., X_p \mid C_j)$ denote probability of a case with individual characteristics $X_1, X_2, ..., X_p$ belonging to j-th class and $P(C_j)$ denote unconditional (without individual characteristics) prior probability of belonging to j-th class. For a total of J classes, Bayes theorem gives the following probability rule to calculate case-specific probability of falling into j-th class [6].

$$P(C_j \mid X_1, X_2, ..., X_p) = \frac{P(X_1, X_2, ..., X_p \mid C_j)P(C_j)}{Denom} \quad \text{-------- (IV)}$$

Where *Denom* =

$$P(X_1, X_2, ..., X_p \mid C_1)P(C_1) + ... + P(X_1, X_2, ..., X_p \mid C_J)P(C_J)$$

Of course the conditional class probabilities of the above equation are exhaustive in that a case $X_1, X_2, ..., X_p$ has to fall in one of the J cases.

B. C4.5

C4.5 is a standard algorithm to induce classification rules as decision tree. A popular decision tree algorithm is C4.5 which is an extension of basic ID3 algorithm. Improvements to C4.5 include: (1) using information gain ratio instead of information gain as a measurement to choose attributes splitting; (2) both discrete and continuous attributes are handled; (3) handles incomplete training data with missing values; (4) prunes during trees construction to avoidover-fitting.

As an ID3 extension, default criteria to choose splitting attributes in C4.5 are *information gain ratio*. C4.5 handles numeric attributes. It incites from a training set incorporating missing values through use of corrected gain ratio criteria.

C. K-nearest neighbor(kNN)

K-nearest neighbors (kNN) rule is an old and basic example characterization technique which delivers good outcomes. Most kNN classifiers by default use simple Euclidean distances to measure dissimilarities. Euclidean distance metrics do not capitalize on statistical data regularity that can be estimated from a large labeled examples training sets. KNN classification is improved by learning a distance metric from labeled examples. A simple (global)input features linear transformation produced much better kNN classifiers

D. Support Vector Machine(SVM)

SVMs introduced by Boser, Guyon & Vapnik in COLT-92 and became popular. Theoretically well motivated algorithm developed from Statistical Learning Theory (Vapnik & Chervonenkis) since the 60s.The support vector machine usually deals with pattern classification that means this algorithm is used mostly for classifying the different types of patterns. Now, there is different type of patterns i.e. Linear and non-linear.

Basically, the main idea behind SVM is the construction of an optimal hyper plane, which can be used for classification, for linearly separable patterns. The optimal hyper plane is a hyper plane selected from the set of hyper planes for classifying patterns that maximizes the margin of the hyper plane i.e. the distance from the hyper plane to thenearest point of each patterns. The main objective of SVM is to maximize the margin so that it can correctly classify the given patterns i.e. larger the margin size more correctly it classifies the patterns. The equation shown below is the hyperplane:

Hyper plane, aX + bY=C ------------- (5)
The given pattern can be mapped into higher dimension space using kernel function, Φ(x)[7].

V. PRECISION ANDACCURACY

Precision refers to the fidelity of the measurement; if the measurement is repeated on the same subject, the same result will be obtained. Accuracy refers to the tendency of measured values to be symmetrically grouped around the variable's true value. Variability of medical data can arise from intra- and inter- instrumental and observer variations (analytical or metrological variability) or intra- and inter-individual variations (biological variability); the total is the combination ofthese.

VI. RESULT AND DISCUSSION

To extract feature vectors related to heart sound signal DWT is useful. The accuracy of the system depends on the identifying the prior level of Wavelet with decomposition at best level. Wave menu is used for analysis purpose. Select the input which is shown in Fig.4 and taken the coiffelet i.e. in Fig.5.In this work PSNR, MSE, RMS, Variance, Energy and Entropy parameters are calculated as shown inFig.6.
In the process of system validation, overall three verities of heart states like VSD, MVP, and PS with heart sounds are analyzed. 150 signals comprising of 80 normal heart sounds, 30 MVP, 20 VSD, 20 PS and Featuresof heart sound signals are taken from Discrete Wavelet Transform.
Different feature selection like Information Gain, Mutual Information is applied to get a feature subset which is classified using Naïve Bayes, C4.5, SVM and kNN. Table I presents the summary the results of Classifiers for syntheticsignals.

Fig.4. Selecting Input

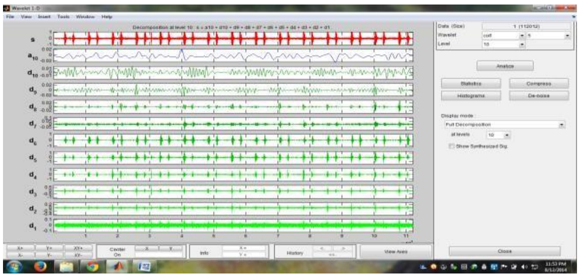

Fig.5: Analysis of One Method i.e., Coiffelet

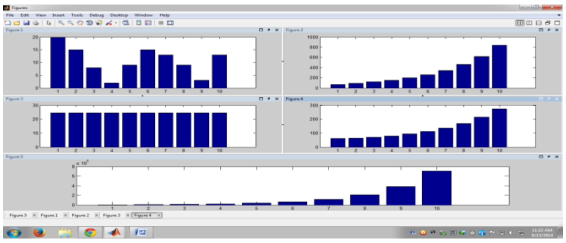

Fig. 6. Calculating Parameters

Table I: Summary Results of Classifiers for Synthetic Signals

Technique	Accuracy %	Precision	Recall	F Measure
IG_Naive_Bayes	84.88	0.8214	0.8283	0.8248
IG_C4.5	86.23	0.834	0.8494	0.8416
IG_KNN	83.54	0.8133	0.8136	0.8134
IG_SVM	90.12	0.8770	0.8956	0.8881
MI_Naive Bayes	80.41	0.7811	0.7936	0.7873
MI_C4.5	83.25	0.8053	0.8304	0.8176
MI_KNN	79.96	0.7754	0.7933	0.7842
MI_SVM	88.54	0.8627	0.8836	0.9260

VII. CONCLUSION & RECOMMENDATION

Especially, the Heart sound signal is quasi-periodic signal, complex and non-stationary. Wavelet methods are adopted to analyze the heart sounds. In this paper, Various features calculated and analyzed the PSNR, Variance, Mean Square, Entropy, Energy and denoising the values along level 10 decomposition. The decomposition and Denoising are calculated and plotted under the basis of wave menu from MATLAB. Feature extraction is analyzed by using DWT. Synthetic signalsoperation conducted with four different classifiers. Information Gain& MI is used for PCG signals feature selection. Classifier Parameters are compared with one another. Results revealed that SVM classifier gave the enhanced results compared to the Naïve Bayes, C4.5, KNN classifiers.

ACKNOWLEDGMENT

The author would like to thank Shri CH. Gopal Reddy, Secretary of CMR College of Engineering & Technology, JNT *University*, Hyderabad, giving support for this work. The author also grateful to Major Dr. V. A. Narayana, Principal of CMR College of Engineering & Technology, JNT *University*, Hyderabad, for his valuable support, especially during the final stage of the work.

REFERENCES

1. G. Venkata Hari Prasad, P. Rajesh Kumar. "Performance analysis of for feature selection methods for feature extracted PCG signals", 2015, 13th International Conference on Electromagnetic Interference and Compatibility (INCEMIC), 2015
2. Anita Devi, Abhishek Misaland Dr. G.R.Sinha, "Performance Analysis of DWT at different levels for Feature Extraction of PCG Signals," International Conference on Microelectronics, Communication and Renewable Energy (ICMiCR-2013), pp. 1-5.
3. "Denoising of PCG Signal by using Wavelet Transforms" Abhishek missal, Advances in Computational Research ISSN: 0975-3273 & E-ISSN: 0975-9085, Volume 4, Issue 1, 2012.
4. A Single channel Phonocardiograph Processing using EMD, SVD, and EFICA, Anil Dada Warbhe, 2010 IEEE.
5. Bai Fang-Fang, Miao Chang-yun, Zang Cheng,Gan Jing-Meng (2010) ICAP,1797-1800
6. Detection of Heart Diseases by Mathematical Artificial Intelligence Algorithm Using Phonocardiogram Signals, Prakash D, Uma Mageshwari T, Prabakaran K, and Suguna A, International Journal of Innovation and Applied Studies ISSN 2028-9324 Vol. 3 No. 1 May 2013.
7. Wang, Q., Guan, Y., Wang, X., & Xu, Z. (2006). A novel feature selection method based on category information analysis for class prejudging in text classification. Proceedings of the International Journal of Computer Science and Network Security, 6(1), 113-119.

AUTHORS PROFILE

Dr. G. Venkata Hari Prasad is currently working as a professor in CMR College of Engineering & Technology, JNT *University,* Hyderabad. He obtained his degree in Doctor of Philosophy in Andhra University and Master of Technology in VTU, Belgaum.

Dr. Lakshmi Narayana Thalluri was born in A.P, India. He received Bachelor's degree in Electronics and Communication Engineering, from JNTUK in the year 2009 & Master's degree in VLSI, from K L University in the year 2012 & Ph.D. in RF MEMS Switches, from K L E F (Deemed to be University) & working as an Assistant Professor, Department of Electronics &Communication Engineering in Andhra Loyola Institute of Engineering Technology (ALIET), Vijayawada, A.P, India. His current research area is RF MEMS Switches, Internet of Things. He is member of IE,IEEE, IACSIT.