International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

ECG data Acquisition Errors Removal from Empirical Mode Decomposition Related Independent Component Values

Uma
Computer Science & Engg.
B.I.E.T., Lucknow, (U.P.)
uma.budhani.85@gmail.com

Mahesh Kumar Singh
Computer Science & Engg.
B.I.E.T., Lucknow, (U.P.)
mks.cse07@gmail.com

***Abstract*-- ECG is the important data of the human body that gives the information measures of the changes in the cardiovascular system. These clinical ECG signal is always corrupted by the electromagnetic field and power line interference which causes misleading results during the diagnosis of diseases by advanced software modules .So it is important to minimize these data acquisition recording errors in the ECG to make the accurate clinical analysis. We are developing an algorithm that can break three consecutive ECG signal data array into 3 equivalents Empirical Mode Decomposition these decomposed data matrix are passed through the Independent Component Generation algorithm from these components we will select and eliminate the noisy components and thereafter reverse ICA and EMD will be applied to get error free ECG data.**

Keywords-- ECG, EMD, Data Array, ICA, and Sapling.

1. Introduction:

A pair of surface electrodes placed directly on the heart will record a repeating pattern of changes in electrical "action potential." As action potentials spread from the atria to the ventricles, the voltage measured between these two electrodes will vary in a way that provides a "picture" of the electrical activity of the heart. The nature of this picture can be varied by changing the position of the recording electrodes; different positions provide different perspectives, enabling an observer to gain a more complete picture of the electrical events. The body is a good conductor of electricity because tissue fluids contain a high concentration of ions that move (creating a current) in response to potential differences. Potential differences generated by the heart are thus conducted to the body surface where they can be recorded by surface electrodes placed on the skin. The recording thus obtained is called an electrocardiogram (ECG or EKG). There are two types of ECG recording electrodes, or "leads."The bipolar limb leads record the voltage between electrodes placed on the wrists and legs. These bipolar leads include lead I (right arm to left arm), lead II (right arm to left leg),and lead III (left arm to left leg). In the unipolar leads, voltage is recorded between a single "exploratory electrode" placed on the body and an electrode that is built into the electrocardiograph and maintained at zero potential (ground). The unipolar limb leads are placed on the right arm, left arm, and left leg; these are abbreviated AVR, AVL, and AVF, respectively. The unipolar chest leads are labeled one through six, starting from the midline position (see below). There are thus a total of twelve standard ECG leads that "view" the changing pattern of the heart's electrical activity from different perspectives. This is important because certain abnormalities are best seen with particular leads and may not be visible at all with other leads. The unipolar limb leads are placed on the right arm, left arm, and left leg; these are abbreviated AVR, AVL, and AVF, respectively. The unipolar chest leads are labeled one through six, starting from the midline position (see below). There are thus a total of twelve standard ECG leads that "view" the changing pattern of the heart's electrical activity from different perspectives. This is important because certain abnormalities are best seen with particular leads and may not be visible at all with other leads.

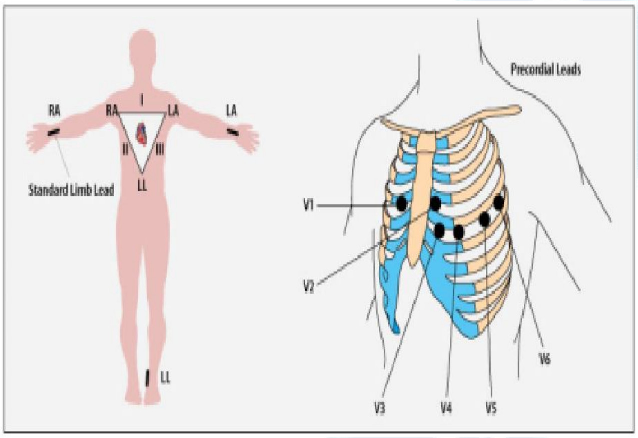

Fig. 1. The placement of the bipolar leads and the exploratory electrode for the unipolar chest leads in an electrocardiogram (ECG); (RA = right arm, LA = Left arm, LL = left leg.)

ECG is the important data of the human body that gives the information measures of the changes in the cardiovascular system. These clinical ECG signal is always corrupted by the electromagnetic field and power line interference which causes misleading results during the diagnosis of diseases by advanced software modules .So it is important to minimize these data acquisition recording errors in the ECG to make the accurate clinical analysis. We are developing an algorithm that can break three consecutive ECG signal data array into 3 equivalents Empirical Mode Decomposition these decomposed data matrix are passed through the Independent Component Generation algorithm from these components we will select and eliminate the noisy components and thereafter reverse ICA and EMD will be applied to get error free ECG data.

2. Related Work:

Sim Kuan Goh, Hussein A. Abbass1, Kay Chen Tan, and Abdullah Al Mamun (2014), [6] according to them Independent Component Analysis (ICA) has been widely used for separating artifacts from Electroencephalographic (EEG) signals. Still, a few challenging problems remain. First, in real-time applications, visual inspection of components should be replaced with an automatic identification method or a heuristic for artifacts detection. Second, as they will explain more in the work, they expect to have a clear order relationship between an electrode and a corresponding component. Third, they need to minimize the EEG information loss during artifact removal while also minimizing the residue of the artifact in the cleaned signal.
In this work, they propose a decomposition of the independent components. This decomposition separates each component into two vectors, one - they call local vector - maintains maximum information from the unique EEG information encoded by an electrode, while the other - they call shared vector - absorbs over- lapping artifact information. They present an explicit Pareto-based multi-objective optimization formulation that trade-off similarity between the local vector and the original vector on the one hand, and the uncorrelatedness of all local vectors from all components on the other hand. They demonstrate that the proposed method can automatically isolate artifacts from an EEG signal while preserving maximum EEG information.

M Murugappan, Reena Thirumani, Mohd Iqbal Omar, and Subbulakshmi Murugappan (2014) [7] according to them the main objective of this work is to develop a portable and cost effective data acquisition (DAQ) system for clinical applications. This DAQ consists of several modules such as power supply, analog to digital converter (ADC), amplifiers, isolators, filters and interfacing circuits. The complete data acquisition circuit has been developed using This system mainly aims to collect the ECG signals of frequency between 0.05 Hz and 113 Hz with a gain of 3113. This frequency information from the ECG signal is highly useful clinical applications such as SCA prediction, cardiovascular disease (CVD) detection, etc. ECG signals will be collected from the subjects using 3 leads system and given to DAQ for recording the ECG signal. The acquired signal through this DAQ will then be transferred to the Notebook through NI6008 data acquisition card. This DAQ interface is used to convert the input analog signal to digital signal output and to save the ECG data in the notebook using Labview software. This acquired signal from Labview software is used for further clinical investigation. We also developed a Graphical User Interface (GUI) in LabVIEW software to continuously monitor the ECG signal traces and to record the ECG data with higher precision. The morphology of the acquired ECG signal in the system is highly precise and useful for clinical diagnosis. Furthermore, this proposed system is used for developing sudden cardiac arrest (SCA) prediction in our university.

Presence of artifacts in electroencephalographs (EEG) is major hurdles for the precise analysis of spectral behavior. For suppression of ocular artifact in EEG this work proposed by **Vandana Roy and Shailja Shukla (2015),** [8] a component based Independent Component Analysis (ICA) model. It involves the generating a set of individual components of given signal followed by rejection of unwanted artifacts. Further this work presents a novel method with combination of ICA, information sharing and double density wavelet transform to reject the artifacts from the signal. The Independent Component Analysis (ICA) here is used to segment artifact peaks in the signal. Then the Discrete Wavelet Transform is applied for multi-level transfer of signal data until the reception of significant result. The Wavelet ICA suppression not only removes artifacts but also preserves the spectral and coherence properties of brain signals.

Electrocardiogram (ECG) is a graphic recording of the electrical activity produced by the heart. The accuracy of any electrocardiogram waveform extraction plays a vital role in helping a better diagnosis of any heart related illnesses. They presented **Akinlolu A. Ponnle et. al. (2015)** [9], a computer-aided application model for detection of cardiac arrhythmia in ECG signal, which consists of signal pre-processing and detection of the ECG signal components adapting Pan-Tompkins and Hamilton-Tompkins algorithms; feature extraction from the detected QRS complexes, and classification of the beats extracted from QRS complexes using Back Propagation Neural Network (BPNN). The application model was developed for ECG signal classification under 'Normal' or 'Abnormal' heartbeats to detect cardiac arrhythmia in the ECG signal. The model was trained with standard arrhythmia database of Massachusetts Institute of Technology Division of Health Science and Technology/Beth Israel Hospital (MIT-BIH), and taking into account the Association for the Advance of Medical Instrumentation (AAMI) standard. The performance of the developed application model for classification of ECG signals was investigated using the MIT-BIH database. The accuracy of detection and extraction of the signal components and features (based only on the MIT-BIH database used) shows that the developed application model can be employed for the detection of heart diseases in patients.

Analysis of EEG activity usually raises the problem of differentiating between genuine EEG activity which is introduced by **Arjon Turnip, and Dwi Esti Kusumandari (2015)**, [10] through a variety of external influence. These artifacts may affect the outcome of the EEG recording. In this work, wavelet denoising and band pass filter for preprocessing and an adaptive principal component analysis based recursive least squares algorithm for extraction are proposed to remove the artifacts. The algorithm is designed to adaptively derive a relatively small number of decorrelated linear combinations of a set of random zero-mean variables while retaining as much of the information from the original variables as possible. The proposed method was tested in real EEG records acquired from eight subjects. The experimental result show that the proposed method can effectively remove the artifacts from all subjects.

3. Methodology:

Usually while ECG data is being taken, different types of noise are added to the ECG signal such as electrode motion,

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

power-line interferences, baseline wander, muscles noise etc., and corrupt the original signal. In order to get rid of the noise, a proper filter must be designed [31]. Since very fine features present in an ECG signal may convey important information, it is important to have the signal as clean as possible. A digital ECG signal is read by MATLAB, and is then normalized. The ECG data is then sampled (or re-sampled) at a frequency of 360 Hz (frequency used in the MIT/BIH records). The proposed ECG signal preprocessing model is shown in Fig 2. The preprocessing stages consist of low pass filtering, high pass filtering, differentiation, Hilbert transform, squaring and moving average. The low pass and high pass filters are cascaded to form a band pass filter.

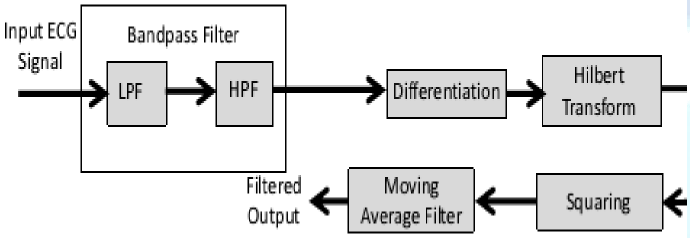

Fig. 2. Block diagram of the proposed ECG signal pre-processing stage

A Butterworth low pass filter (LPF) is used to eliminate noise such as the electromagnetic interference and 50Hz power line noise. The adopted designed LPF is of order 6, a cutoff frequency of 11 Hz and a sampling frequency of 360 Hz. The difference equation representing the LPF [29] is

$$y(n) = 2y(n-1) - y(n-2) + x(n) - 2x(n-6) + x(n-12)$$

The IIR LPF was designed in MATLAB and the designed filter object was then used to filter the input ECG signal. The output of the LPF is passed into the high pass filter (HPF) to eliminate motion artifacts. The adopted designed HPF has a cutoff frequency of 5Hz, and the difference equation is given [29] as

$$y(n) = y(n-1) + x(n) + x(n-1) + x(n-5) - x(n-10)$$

The filtered ECG signal is then differentiated to give the slope information by accentuating QRS complexes relative to P and T wave. The differentiator also helps to overcome baseline
wandering in the signal. The adopted difference equation used to design the differentiator [29] is

$$8y(n) = 2x(n) + x(n-1) - x(n-4) - 2x(n-5)$$

After the differentiation of the ECG signal, Hilbert Transform is applied to the signal to find the location of R-peak in the ECG signal. The output of the Hilbert transform is squared in order to emphasize the higher frequency component and attenuates the lower frequency component. This helps to suppress the P and T waves.
The squared signal is then passed into the moving average filter (MAF) to produce a waveform with smoothed features by performing moving window integration. The difference equation designed for the MAF is given as

$$y(n) = \frac{x[n-(N-1)] + x[n-(N-2)] + \ldots + x[n]}{N}$$

where N is the length of the MAF, i.e. N-point MAF. In our developed application, we used 3 for the value of N.

A. Detection of the ECG Signal Components:

The next step after preprocessing of the ECG signal is the detection of the R, Q, S, P and T points on the signal waveform. These points were first detected before extracting the features needed for training the neural network object. The flow chart of Fig 4 represents the steps of obtaining the points. This was achieved by taking a beat out of the ECG waveform. The beats are extracted using 128 samples centered on R points. Pan-Tompkins and Hamilton-Tompkins algorithms were adapted for this operation to suit our application [10].

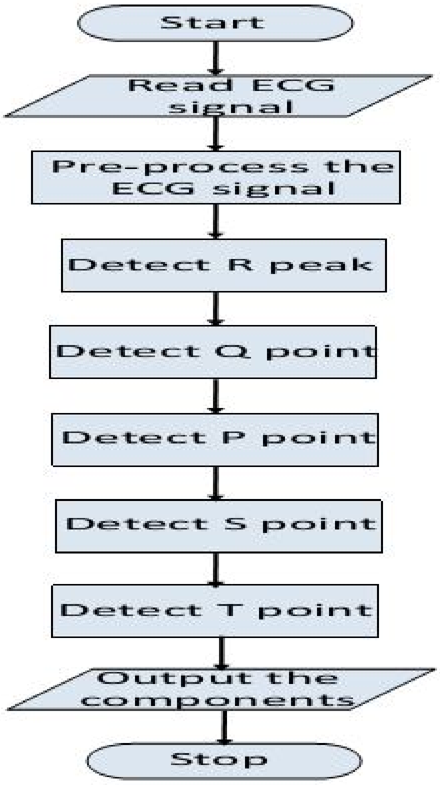

Fig. 3. Flow Chart of Detection of the ECG signal components.

Basically in our application, the detection of the points consists of the following steps, which is illustrated in Fig 4.

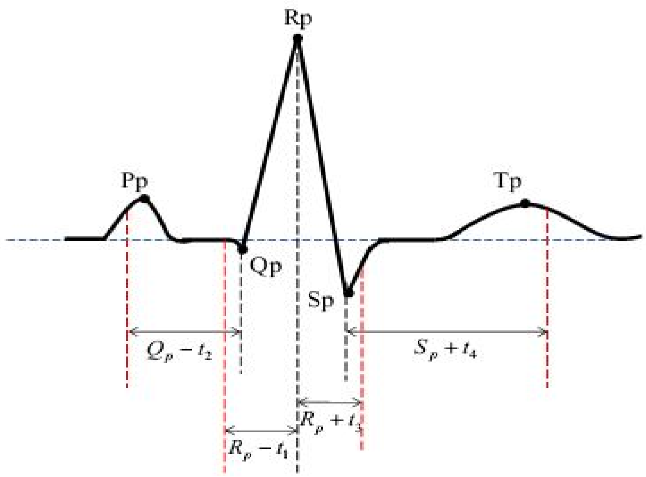

Fig. 4. Detection of the ECG signal components.

ISSN: 2454-6844

a. Locate the maximum amplitude in the signal beat; this is the R peak. The point at which the R peak is detected is the R point Rp.
b. Shift some steps to the left of R peak, that is Rp - t1, and locate the minimum amplitude; this is the Q point Qp. t1 is the number of steps taken until the signal to the left of R begins to rise i.e. changes direction.
c. Shift some steps to the left of Qp, that is Qp - t2, and locate the maximum amplitude; this is the P point Pp. t2 is the number of steps taken until the signal to
the left of Qp begins to fall i.e. changes direction.
d. Shift some steps to the right of Rp, that is Rp + t3, and locate the minimum amplitude; this is the S point Sp. t3 is the number of steps taken until the signal to the right of Rp begins to rise i.e. changes direction.
e. Shift some steps to the right of Sp, that is Sp + t4, and locate the maximum amplitude; this is the T point Tp. t4 is the number of steps taken until the signal to the right of Sp begins to fall i.e. changes direction.

In adapted algorithms, the determination of t1, t2, t3 and t4 to be adaptive within some predefined time intervals with respect to some pre-determined values based on knowledge of standard normal ECG waveform; as for example, there could be possibility of false peaks to occur due to noise or otherwise between Sp and Tp. The step size is made equal to the sampling interval of the signal.

4. Result and Discussion:

We have considered an ECG data having length of 50000X1 samples. It is an ECG data of healthy person having 10 ECG cycles. The data plot with respect to time axis is shown in figure 1 and magnified view of 1st cycle is shown in figure 5.

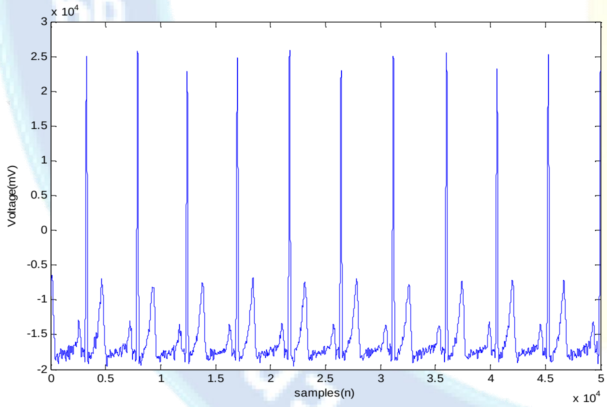

Fig. 5. ECG waveform of data prior to adding noise.

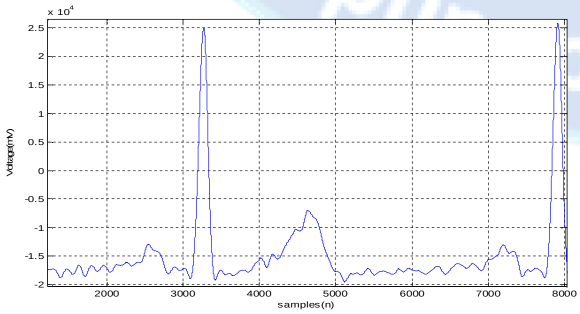

Fig. 6. Magnified View of 1st ECG waveform of data prior to adding noise.

Since above signal shown in figure 1 consist of 50000 samples for only 10 cycles it will take large memory space hence processing time. So the number of samples are reduced by down sampling by 8 times thus total samples we obtained are 50000/8=6250 data points. This down sampled signal are shown in figure 7.

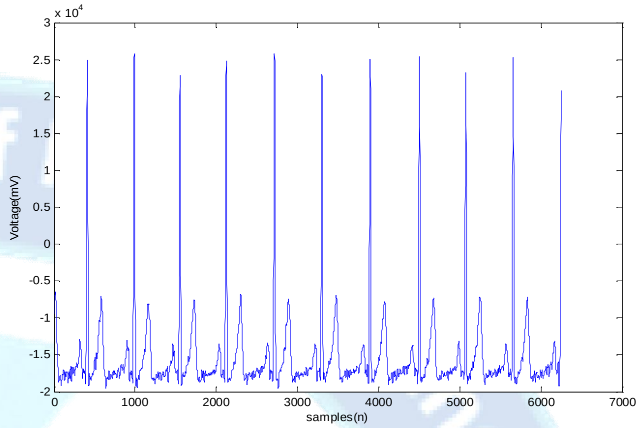

Fig. 7. ECG waveform of data prior to adding noise after down sampling by 8 times.

We can see that there are no significant differences in figure 1 and figure 3 even eliminating the data information by 8 times but along with this denoising processing time will become fast.The above ECG data is to be passed through Empirical mode decomposition (EMD) prior to this we will have to clip multiple portions of above signal to make a multiple dimension data. For clipping out the ECG data we have pointed the peaks location foe the given ECG records as shown in figure 4.From the figure 4 we have defined the approximated position of peaks as an array x.

x= [409 998 1552 2127 2716 3300 3896 4520 5075 5659 6250].

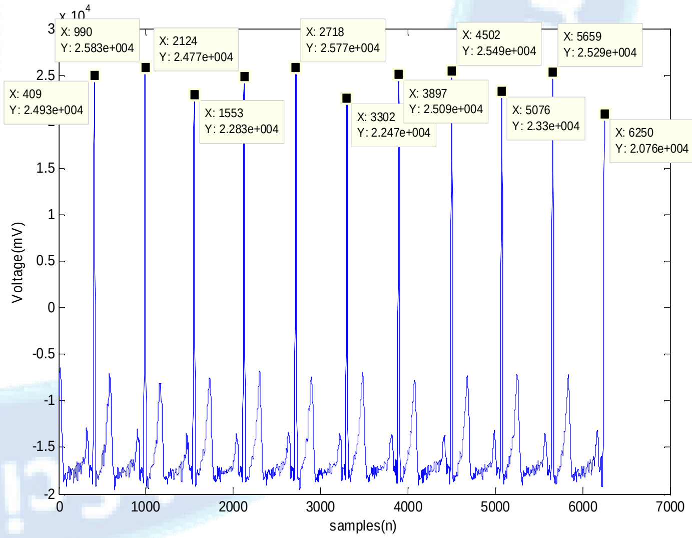

Fig. 8. ECG data peak location for all 10 cycles. (x:location of peak value, y:value of peak ECG voltage).

Since available database of any biomedical signals are contributed by standard research lab hence they are extremely high quality instrument based measurements values and taken under several precise environment. Due to this these data do not consist of any noise. For testing our denoising algorithm we require to add noise in these data such that they exhibit distortions. We have added noise shows distortions in there waveforms. We have added Gaussian noise in the signal having noise power 10% of the signal power. The generated noisy signal and its magnified view are shown in figure 5 and 6.

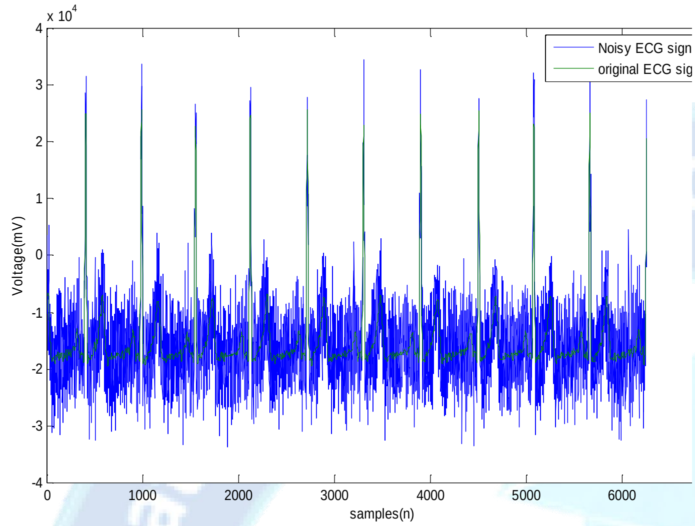

Fig. 9. Noisy ECG data (blue) and original ECG data (red).

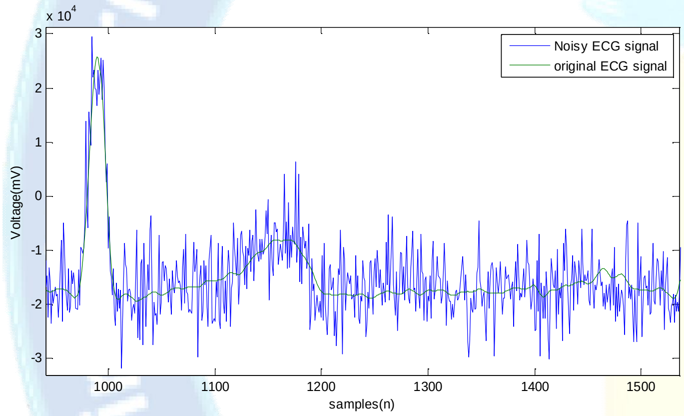

Fig. 10. Magnified View of Noisy ECG data (blue) and original ECG data (red).

As per the location ECG signal cycles are one by one for performing the EMD of the noisy ECG data cycles. For example first two cycles is taken by assigning samples from first peak to third peak location given in array x. All the clipped ECG cycles are shown in figure 7 below.

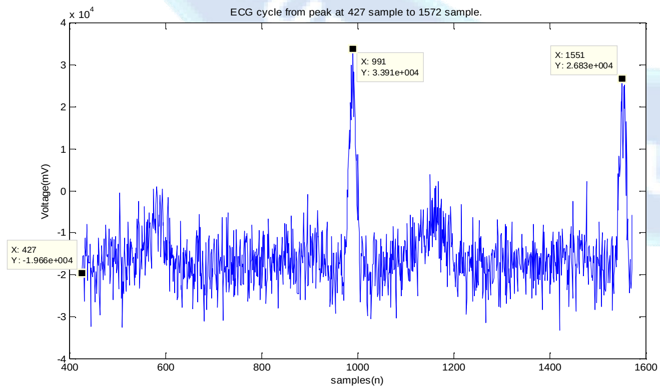

Fig. 11. Clipped Noisy ECG signal from first two cycles.

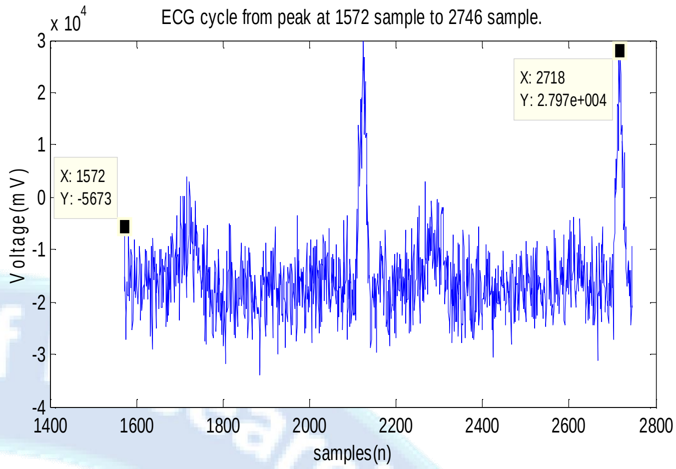

Fig. 12. Clipped Noisy ECG signal from third to fifth cycles.

Similar to cases of figure 7 and we have taken the discrete data blocks one by one for the processing under EMD algorithm to obtain these cycles 12 decompositions. All the 12 decomposed component of EMD for the cycles of figure 7 are shown in figure 8 below. The top left is the ECG signal and remaining 12 are EMD components.

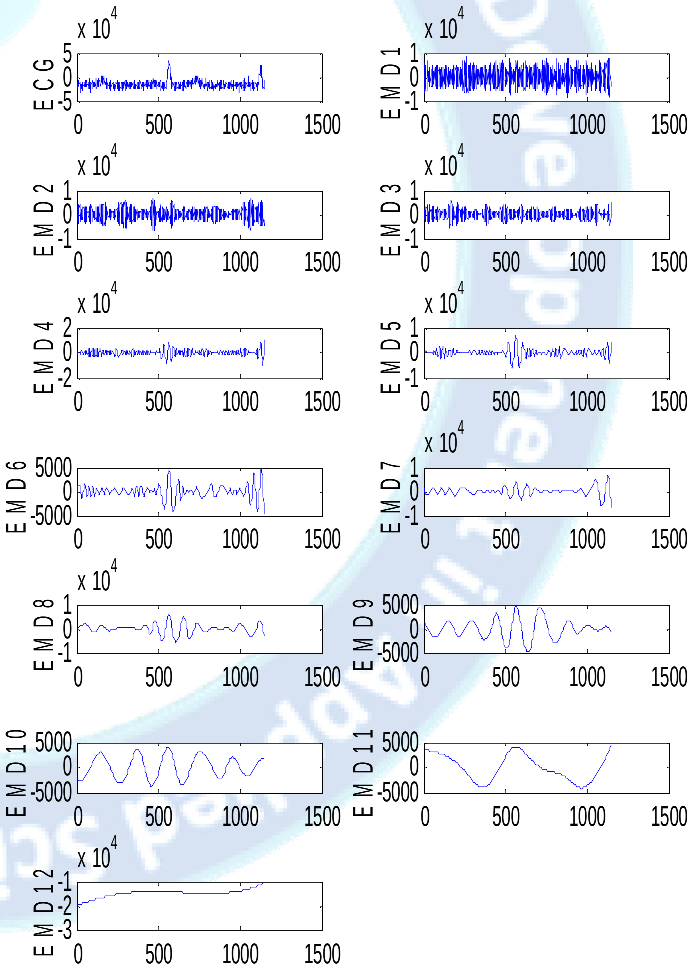

Fig. 13. Initial two cycles of noisy ECG and its EMD components from EMD1 to EMD12.

The algorithm after performing the ICA decomposition shows all the ICA components in one window along with this it also displays the kurtosis(degree of independence) of all the generated ICA components as given below.

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

Kurtosis of_ICA comoponents = (1) 2.4905 (2) 3.4545 (3) 2.4358 (4) 2.4409 (5) 2.7921 (6) 2.6080 (7) 1.9007 (8) 2.9257 (9) 2.7919 (10) 2.7599 (11) 2.7069 (12) 3.2093.

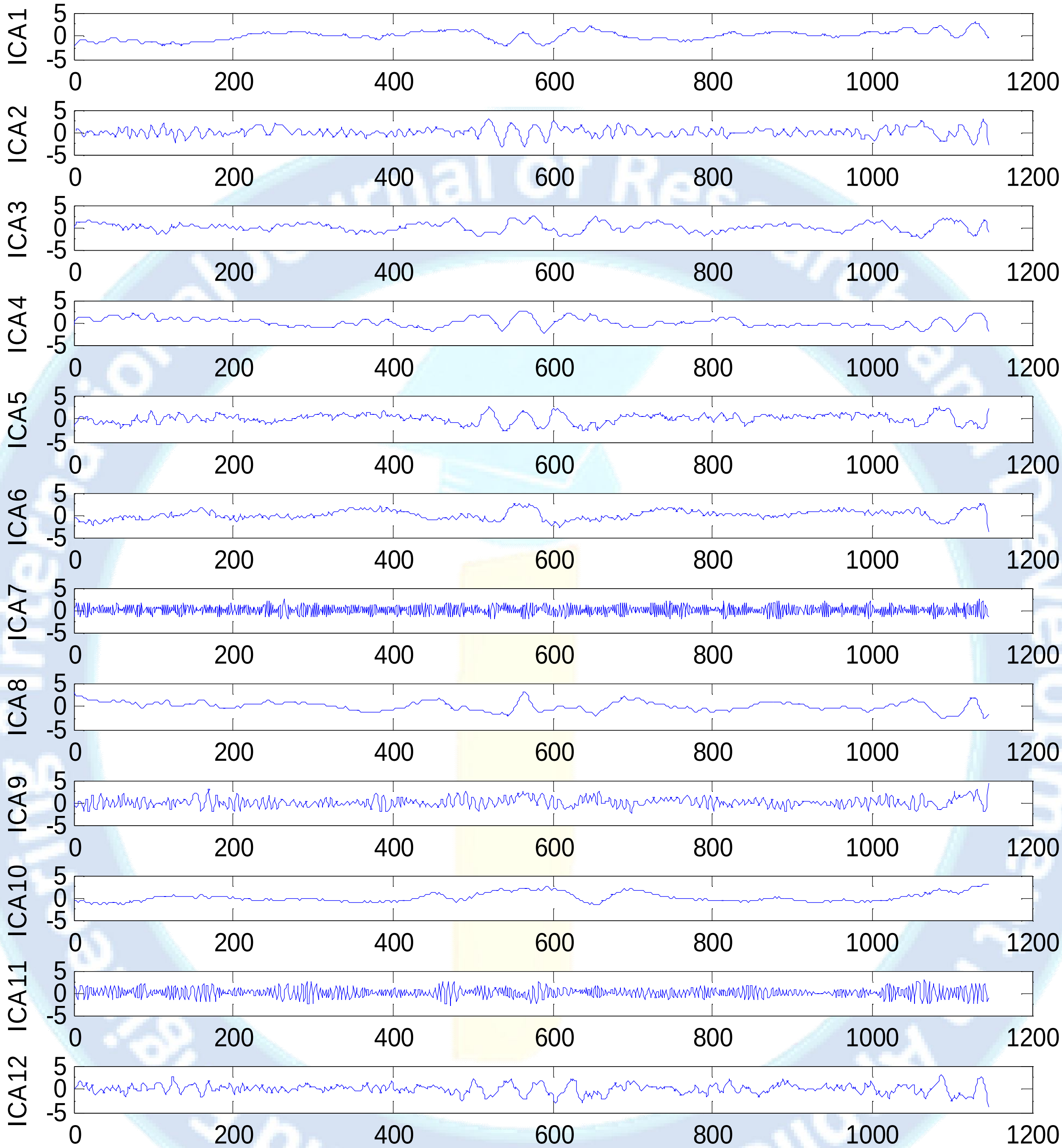

Fig. 14. Collective view of ICA components ICA1 to ICA12 for EMD of ECG cycles in one window.

5. Conclusion

In this paper we consider distortion related problem in diagnosis based on amplitude of ECG applied in common practice. By ECG amplitude analysis we can construct a new set of signals from the signal amplitudes at some defined points of the ECG, such as R peak or ST amplitudes or from time averages of delineated ECG segments. We have developed an algorithm by combining ICA and EMD decomposition techniques for error minimization in ECG database for improving diagnosis quality. ICA has found several applications in signal processing systems aimed at aiding in diagnostics. ECG based diagnostics applications in which ICA has been utilized in the applications of classification of ECG beats, analysis of parameterized ECG signals, heart rate variability analysis, arrhythmia estimation and atrial fibrillation extraction and analysis. In our literature survey we have observed that 12-lead ECG may sometimes be insufficient for efficient ICA based analysis of the phenomenon of interest Zhu et al. (2008) analyzed 72-lead and 98-lead ECG measurements using ICA and were

ISSN: 2454-6844

able to separate the P wave, QRS complex, and T wave. Thus, with high-density ECG measurements and ICA based analysis more detailed diagnostics applications might be realizable. In this works we have shown that even using single lead ECG records we can consider ICA decomposition upto 12 level if we apply EMD prior to ICA.

References:

[1] Jamal N. Al-Karaki, Ahmed E. Kamal," Routing Techniques In Wireless Sensor Networks: A Survey", IEEE Wireless Communications, Volume: 11, Issue: 6, 26- 28, December 2004.

[2]. Georgios Smaragdakis, Ibrahim Matta , Azer Bestavros, "SEP: A Stable Election Protocol for clustered wireless sensor networks", Computer Science Department, Boston University, Boston, MA 02215, USA

[3]. Li Qing *, Qingxin Zhu, Mingwen Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks", School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, PR China

[4]. Brahim Elbhiri, Saadane Rachid , Sanaa El fkihi, "Developed Distributed Energy-Efficient Clustering (DDEEC) for heterogeneous wireless sensor networks"

[5]. Ajay.K.Sharma,Department of Computer Science & Engineering, Parul Saini, Department of Computer Science & Engineering, "E-DEEC- Enhanced Distributed Energy Efficient Clustering Scheme for Heterogeneous WSN"

[6]. T. N. Qureshi, N. Javaid, M. Malik, U. Qasim, Z. A. Khan, "On Performance Evaluation of Variants of DEEC in WSNs".

[7] E. J. Duarte-Melo and M. Liu. Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks. In Proceedings of Global Telecommunications Conference (GLOBECOM 2002), pages 21–25. IEEE, November 2002.

[8]. Parul Saini, Ajay K Sharma, "Energy Efficient Scheme for Clustering Protocol Prolonging the Lifetime of Heterogeneous Wireless Sensor Networks"

[9] G. Smaragdakis, I. Matta, A. Bestavros, SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks, in: Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004), 2004.

[10] W.R. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application- specific protocol architecture for wireless microsensor networks, IEEE Transactions on Wireless Communications 1 (4) (2002) 660–670.