Advanced Diagnosis Approach using EMD and ICA for PQRS Detection.

Shobhana Awasthi, Ambuj Pandey
Computer Science and Engineering,
B. N. College of Engineering and Technology, Lucknow, India

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, EKG, ICA, CVD

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.

2. Related Work:

Aapo Hyvärinen and Erkki Oja (2000), [1] worked on a fundamental problem in neural network research, as well as in many other disciplines, is finding a suitable representation of multivariate data, i.e. random vectors. For reasons of computational and conceptual simplicity, the representation is often sought as a linear transformation of the original data. In other words, each component of the representation is a linear combination of the original variables. Well-known linear transformation methods include principal component analysis, factor analysis, and projection pursuit. Independent component analysis (ICA) is a recently developed method in which the goal is to find a linear representation of nongaussian data so that the components are statistically independent, or as independent as possible. Such a representation seems to capture the essential structure of the data in many applications, including feature extraction and signal separation. In this work, we present the basic theory and applications of ICA, and our recent work on the subject.
ICA is a very general-purpose statistical technique in which observed random data are linearly transformed into components that are maximally independent from each other, and simultaneously have "interesting" distributions. ICA can be formulated as the estimation of a latent variable model. The intuitive notion of maximum nongaussianity can be used to derive different objective functions whose optimization enables the estimation of the ICA model. Alternatively, one may use more classical notions like maximum likelihood estimation or minimization of mutual information to estimate ICA; somewhat surprisingly, these approaches are (approximatively) equivalent. A computationally very efficient method performing the actual estimation is given by the FastICA algorithm. Applications of ICA can be found in many different areas such as audio processing, biomedical signal processing, image processing, telecommunications, and econometrics.

Huang's data-driven technique of Empirical Mode Decomposition (EMD) is presented by **Gabriel Rilling et. al., (2003),** [2], and issues related to its effective implementation are discussed. A number of algorithmic variations, including new stopping criteria and an on-line version of the algorithm, are proposed. Numerical simulations are used for empirically assessing performance elements related to tone identification and separation. The obtained results support an interpretation of the method in terms of adaptive constant-Q filter banks.
EMD is a promising new addition to existing toolboxes for nonstationary and nonlinear signal processing, but it still needs

to be better understood. This work discussed algorithmic issues aimed at more effective implementations of the method, and it proposed some preliminary performance measures. The results reported here are believed to provide with new insights on EMD and its use, but they are merely of an experimental nature and they clearly call for further studies devoted to more theoretical approaches.

Arnaud Delorme and Scott Makeig, (2004), [3] have developed a toolbox and graphic user interface, EEGLAB, running under the cross-platform MATLAB environment (The Mathworks, Inc.) for processing collections of single-trial and/or averaged EEG data of any number of channels. Available functions include EEG data, channel and event information importing, data visualization (scrolling, scalp map and dipole model plotting, plus multi-trial ERP-image plots), preprocessing (including artifact rejection, filtering, epoch selection, and averaging), Independent Component Analysis (ICA) and time/frequency decompositions including channel and component cross-coherence supported by bootstrap statistical methods based on data resampling. EEGLAB functions are organized into three layers. Top-layer functions allow users to interact with the data through the graphic interface without needing to use MATLAB syntax. Menu options allow users to tune the behavior of EEGLAB to available memory. Middle-layer functions allow users to customize data processing using command history and interactive 'pop' functions. Experienced MATLAB users can use EEGLAB data structures and stand-alone signal processing functions to write custom and/or batch analysis scripts. Extensive function help and tutorial information are included. A 'plug-in' facility allows easy incorporation of new EEG modules into the main menu. EEGLAB is freely available (http://www.sccn.ucsd.edu/eeglab/) under the GNU public license for non commercial use and open source development, together with sample data, user tutorial and extensive documentation.
Another relative disadvantage of using Matlab to process high-density EEG data is that Matlab currently converts all floating-point numbers to 64-bit doubleprecision, thus requiring large amounts of main memory to process large data sets. Though hopefully some future Matlab versions may allow the option of processing data in 32-bit floating-point format, we
have taken care to address this issue in EEGLAB by including various options to minimize memory usage, such as constraining EEGLAB to work on a single dataset, or computing the 'activation' time courses of independent components only as needed. However, this issue remains a serious problem for large datasets: parts of the toolbox may have to be updated to allow very large (e.g., long 256-channel) datasets to be analyzed within the current Linux 2GB/process limit. One possibility is to use the Matlab MEX language, an interface between C and Matlab that allows a wider variety of data types including single precision. Another possibility is to have EEGLAB load into main memory only a part of the dataset at a time. However, as 64-bit processors become more available, the current data space limits of operating systems and Matlab should increase, in which case the remaining problem would only be the burden of purchasing the necessary RAM. Current development of EEGLAB focuses on processing of large datasets (>1 Gb), semiautomatically grouping independent component across subjects, and component source localization. EEGLAB will also be linked to our FMRLAB toolbox (http://www.sccn.ucsd.edu/fmrlab) to process simultaneously recording EEG and fMRI data (Duann et al., 2002a). We also have begun working with codevelopers to increase the range of EEGLAB functions using the 'plug-in' facility, whereby contributors may easily contribute optional EEGLAB code that is readily incorporated into the EEGLAB menu. The plug-in facility is designed so that plug-in functions can be used and distributed both within EEGLAB and independently. By this mechanism we hope to encourage the open source development of comprehensive EEG (and MEG) signal processing tools under EEGLAB.

Empirical mode decomposition (EMD) has recently been pioneered by **Patrick Flandrin, Gabriel Rilling, and Paulo Gonçalvés, (2004),** [4] for adaptively representing nonstationary signals as sums of zero-mean amplitude modulation frequency modulation components. In order to better understand the way EMD behaves in stochastic situations involving broadband noise, we report here on numerical experiments based on fractional Gaussian noise. In such a case, it turns out that EMD acts essentially as a dyadic filter bank resembling those involved in wavelet decompositions. It is also pointed out that the hierarchy of the extracted modes may be similarly exploited for getting access to the Hurst exponent.
They reported here on first numerical experiments aimed at supporting the claim that, in the case of structured broadband stochastic processes such as fractional Gaussian noise, the built-in adaptivity of EMD makes it behave spontaneously as a "wavelet-like" filter bank. An interesting by-product of this interpretation is that EMD may offer a new way of analyzing self-similar processes. Thorough comparisons (which are beyond the scope of this letter) with other existing methods are in progress. Let us just mention that benefits very similar to those of wavelet-based methods are obtained when using EMD: in particular, the technique happens to naturally cope with superimposed smooth trends. From a more general perspective, the results presented here clearly call for theoretical elements which would explain the observed behaviors (e.g., the -dependence of the filter bank structure), a task which is made difficult by the fact that EMD does not admit an analytical definition. The purpose of the present experimental study was to be a contribution aimed at a better understanding of one specific aspect of EMD (the way it decomposes broadband noise), filling somehow the gap between a still nonexisting theory and the application of an appealing method to real-world situations.

An Extended Kalman Filter (EKF) has been proposed by **Reza Sameni, M.B. Shamsollahi, Christian Jutten, and Massoud Babaie-Zadeh (2007)**, [5] for the filtering of noisy ECG signals. The method is based on a modified nonlinear dynamic model, previously introduced for the generation of synthetic ECG signals. An automatic parameter selection method has also been suggested, to adapt the model with a vast variety of normal and abnormal ECG signals. The results show that the EKF output is able to track the original ECG signal shape even in the most noisiest epochs of the ECG signal. The proposed method may serve as an efficient filtering procedure for applications such as the noninvasive extraction of fetal cardiac signals from maternal abdominal signals.
In this work an Extended Kalman Filter (EKF) was designed for the filtering of ECG signals. The EKF's dynamic model was based on a modified three dimensional nonlinear dynamic model previously introduced for the generation of synthetic ECG signals. This nonlinear model was linearized in order to

be used in an EKF. The designed filter was later applied to noisy ECG signals, and the results show the filter's capability in tracking and filtering noisy ECG signals. The evaluation of the EKF implemented in this work was quite qualitative. In practical applications it is necessary to represent more quantitative measures, together with issues concerning the stability and convergence of the Kalman filter. The filtering performance is highly reliant on the underlying dynamics assumed for the ECG signal. It was
shown that by using a flexible nonlinear dynamical model, together with the EKF, it is possible to construct a filter which can remove environmental noises and artifacts. The proposed method can serve as a base for the design of a robust ECG filter, with vast applications for low SNR ECG signals such as the noninvasive fetal cardiac signal extraction.
Future works include the combination of the proposed EKF model with source separation techniques, for the extraction of maternal and fetal cardiac signals from multi-channel surface electrode recordings.

3. Methodology:
The extraction of high resolution ECG signals from noisy measurements is among the most tempting open problems of biomedical signal processing. Specifically, the extraction of ECG signals from low SNR measurements is the state of the art in applications such as the noninvasive extraction of fetal ECG signals, recorded from an array of electrodes placed on the maternal abdomen [27].
On the other hand, in recent years some research has been conducted towards the generation of synthetic ECG signals. Regarding the physiological bases of ECG signals, a true ECG model should consider the morphology of the PQRST complex, together with the RR-wave timing. In a previous work [12], a synthetic model has been proposed which has unified the morphology and pulse timing of the ECG signal in a single nonlinear dynamic model. Concerning the simplicity and flexibility of this model it is believed that it can be easily adapted to a broad class of normal and abnormal ECG signals. This model may be further used in dynamic adaptive filters, such as the Kalman Filter, for ECG filtering applications. Meanwhile, the dynamic model of [12] is nonlinear and requires the nonlinear counterparts of the conventional Kalman Filter.
In a recent work [13], the authors have developed an Extended Kalman Filter (EKF) based on the mentioned dynamic model for noisy ECG filtering. In this paper, the synthetic ECG model has been further modified to fulfill the requirements of the EKF filter. The EKF model
parameter selection has also been automated in order to adapt the method to different normal and abnormal ECG signals. The results show that the proposed method can fully track the ECG signal even in the noisy epochs, where the observed ECG signal is almost lost in noise.
Electrocardiogram (ECG) signals plays a vital role in clinical diagnosis especially for diagnosing heart related diseases and disorders such as, cardiovascular disease (CVD), pulmonary disease, sudden cardiac arrest (SCA), etc [7]. ECG signal is generated by a nerve impulse stimulus to a heart. The current is diffused around the surface of the body and build on the voltage drop, which is a normally 0.0001 to 0.003volt and the signals are within the frequency range of 0.05 to 100 Hz [27] [22]. ECG signals are usually recorded at the surface of the body and processed to give important information about the electrical activity of heart. A typical ECG tracing of a normal heartbeat consists of a P wave, a QRS complex and a T wave (Figure 3.1). Usually, the signal which is acquired from the human body is of very low potential and difficult to analyze the signal variance. Hence, necessary amplification is required before processing the ECG signal to derive any give useful information about the cardiac abnormalities.

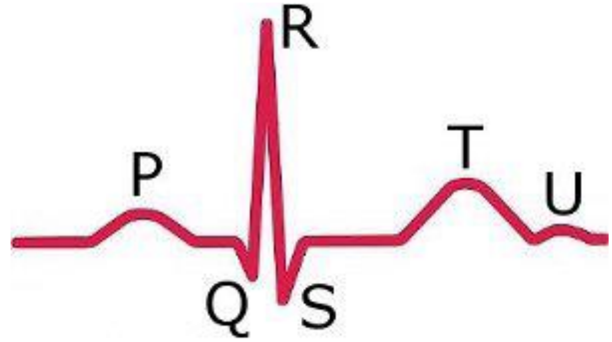

Fig 1: The elements of ECG complex

Biomedical signals are observations of physiological activities of organisms, ranging from protein sequences, tissue and organ images, to neural and cardiac rhythms. Biomedical signals are obtained by electrodes that record the variations in electrical potential generated by physiological processes. Each physiological process is associated with certain types of signals that reflect their nature and activities. Observing these signals and comparing them to their known norms, diseases or disorders can often be detected. When such measurements are observed over a period of time, a one dimensional time-series is obtained which is called a physiological signal. Arrhythmia is a generalized term used to denote any disturbances in the heart's rhythm. Cardiac Arrhythmia is an abnormal rate of muscle contractions in the heart. These abnormalities of heart may cause sudden cardiac arrest or cause damage to heart. Proper diagnosis of arrhythmia requires an electrocardiogram.
The development of the model for the application can be divided into the following stages: ECG Signal Pre-processing, Feature Detection, Feature Extraction, and Feature Classification using BPNN. For the signal pre-processing and feature detection, we made use of Pan-Tompkins and Hamilton-Tompkins algorithms [2], and adapted them to suit our application. The algorithms are more popular in QRS detection methods. For the feature classification by BPNN, we adopted MATLAB in-built Levenberg-Marquardt (LM) algorithm. The model accepts and works on already digitally acquired ECG signal, and MATLAB software was used to both implement and evaluate the application model, using MIT-BIH database. Figure 2 shows the block diagram representation of the developed ECG beat classifier.

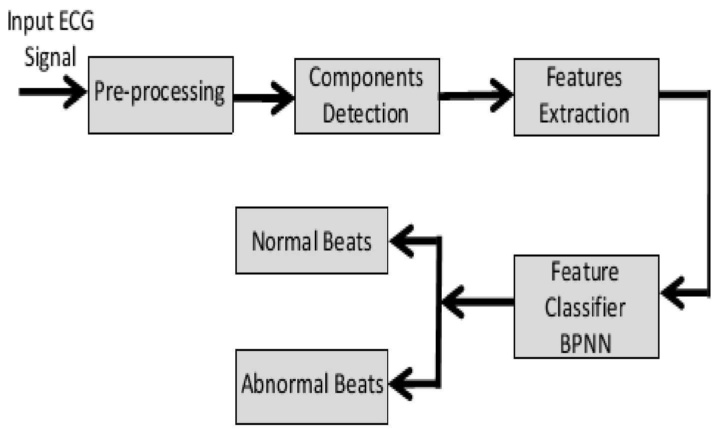

Figure 2: Block diagram of the ECG beat classifier.

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 3 and magnified view of 1st cycle is shown in figure 4.

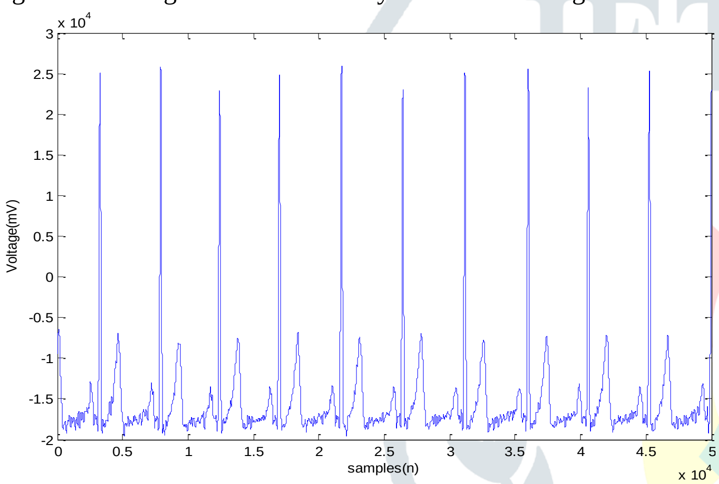

Fig 3: ECG waveform of data prior to adding noise.

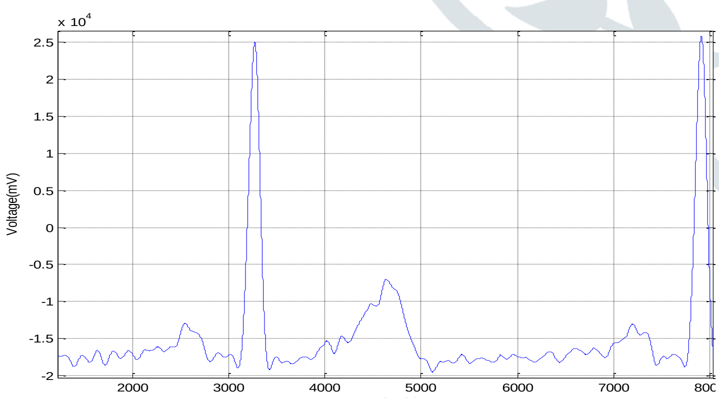

Fig 4: Magnified View of 1st ECG waveform of data prior to adding noise.

Since above signal shown in figure 3 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 5.

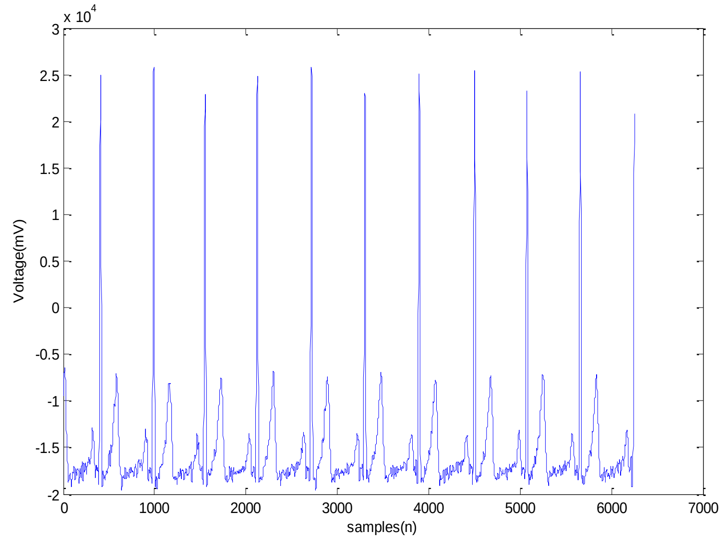

Fig 5: ECG waveform of data prior to adding noise after down sampling by 8 times.

We can see that there are no significant differences in figure 3 and figure 5 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 6 we have defined the approximated position of peaks as an array x.

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

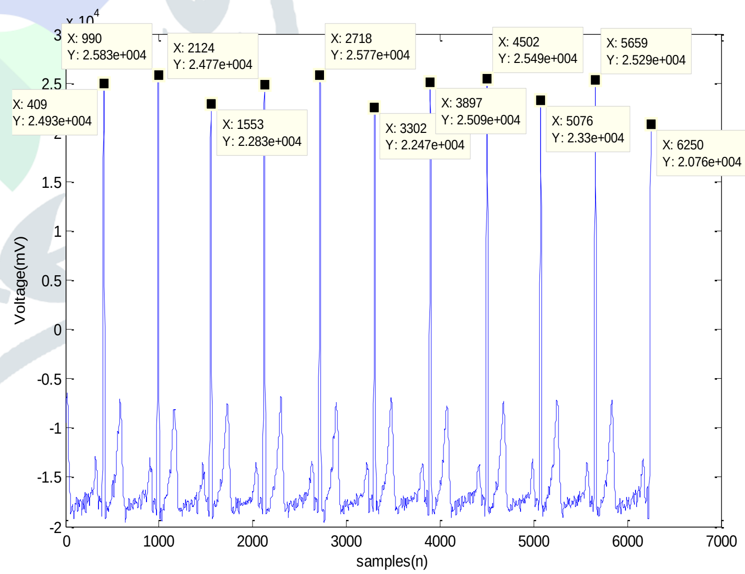

Fig 6: ECG data peak location for all 10 cycles. (x:location of peak value, y:value of peak ECG voltage).

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 7 and 8.

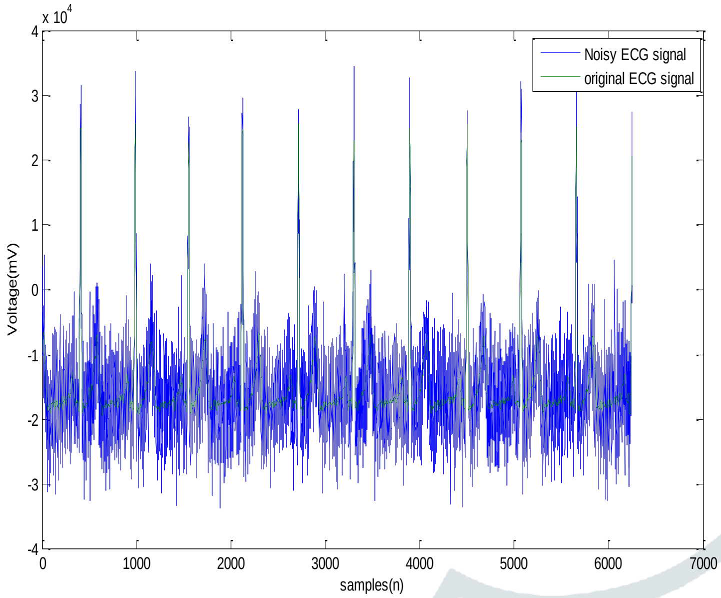

Figure 7: Noisy ECG data (blue) and original ECG data (red).

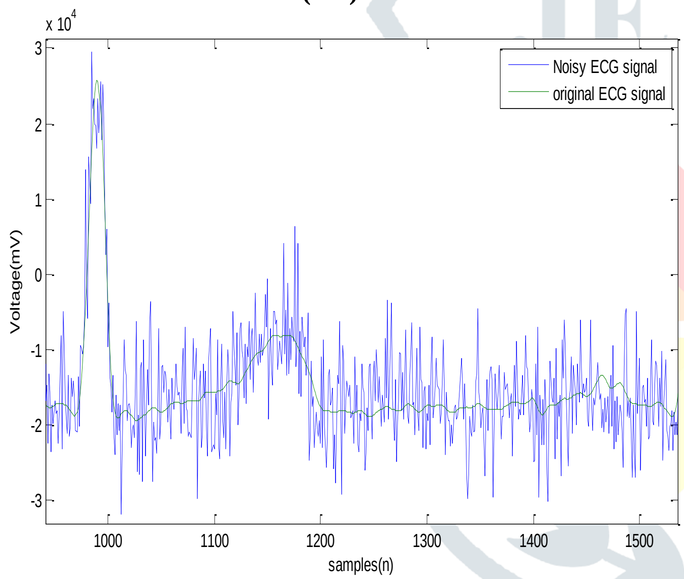

Figure 8: Magnified View of Noisy ECG data (blue) and original ECG data (red).

5. Conclusion:

In this work 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.

References:

[1] Aapo Hyvärinen and Erkki Oja, "Independent Component Analysis: Algorithms and Applications", Neural Networks, 13(4-5):411-430, 2000.

[2] Gabriel Rilling et. al., "On Empirical Mode Decomposition And Its Algorithms" IEEE 2003.

[3] Arnaud Delorme and Scott Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis" Journal of Neuroscience Methods 134 (2004), 9-21.

[4] Patrick Flandrin, Gabriel Rilling, and Paulo Gonçalvés, "Empirical Mode Decomposition as a Filter Bank" IEEE Signal Processing Letters, Vol. 11, NO. 2, February 2004.

[5] Reza Sameni, M.B. Shamsollahi, Christian Jutten, and Massoud Babaie-Zadeh, "Filtering Noisy ECG Signals Using the Extended Kalman Filter Based on a Dynamic ECG Model", Computers in Cardiology 2007;32:1017−1020.

[6] Sim Kuan Goh, Hussein A. Abbass1, Kay Chen Tan, and Abdullah Al Mamun, "Artifact Removal from EEG using a Multi-Objective Independent Component Analysis Model" Volume 8834 of the series Lecture Notes in Computer Science pp 570-577 2014.

[7] M Murugappan, Reena Thirumani, Mohd Iqbal Omar, and Subbulakshmi Murugappan, "Development of Cost Effective ECG Data Acquisition System for Clinical Applications using LabVIEW" 2014 IEEE 10th International Colloquium on Signal Processing & its Applications (CSPA2014), 7 - 9 Mac. 2014, Kuala Lumpur, Malaysia.

[8] Vandana Roy and Shailja Shukla, "A Two Stage Approach with ICA and Double Density Wavelet Transform for Artifacts Removal in Multichannel EEG Signals" International Journal of Bio-Science and Bio-Technology Vol.7, No.4 (2015), pp.291-304.

[9] Akinlolu A. Ponnle et. al., "Development of a Computer-Aided Application for Analyzing ECG Signals and Detection of Cardiac Arrhythmia Using Back Propagation Neural Network - Part I: Model Development" International Journal of Applied Information Systems (IJAIS)Volume 9 – No.3, June 2015.

[10] Arjon Turnip, and Dwi Esti Kusumandari, "Artifacts Removal of EEG Signals using Adaptive Principal Component Analysis" 3rd International Conference on Computation for Science and Technology (ICCST-3) 2015.

[11] Upasani D.E and Kharadkar R.D. 2012, Automated ECG Diagnosis, IOSR Journal of Engineering, Vol. 2 (5), pp. 1265-1269.

[12] Karpagachelvi S., Arthanari M., and Sivakumar M. 2010, ECG Feature Extraction Techniques: A Survey Approach, International Journal of Computer Science and Information Security, Vol. 8, No. 1, pp. 76-80.

[13] C. Li, C. Zheng (1993) Proc. Annual Int. Conf. IEE Eng. in Med. & Biol. Soc., San Diego, California, 330331.