Evaluation of different parameters of noisy ECG signal generated by jointly using MIT-BIH Noise Stress Test Database and CSE multi-lead clear ECG database

[1]Sucharita Mitra(Sarkar), [2]Priyanka Samanta
Department of Electronics [1] Netaji Nagar Day College ,Kolkata.[2] Vidyasagar college for women,,Kolkata.

Abstract- A realistic 12 lead of noisy ECG signal database is developed here using MIT- BIH Noise Stress Test Database (NSTDB) of three type of noise with clear ECG from CSE multi-lead database. Then used this generated noisy ECG database on a algorithm of detection and identification of ECG wave by histogram approach to analyze its performance over noisy ECG signal, as it does on clear ECG, for different value of signal to noise ratio (SNR) and compare the result by computing average percentage error for different parameters of each ECG lead.

Keyword: Noise, Lead, Database, signal to noise ratio, Average percentage error.

I. INTRODUCTION

Electrocardiogram is a most frequently used diagnostic tool for recording various features of human heart, which is very important for human healthcare purpose. When an ECG is recorded, a reading of voltage vs time is produced, which is normally displayed as millivolts (mV) and seconds. It shows a series of peaks and waves that corresponds to ventricular or atrial depolarization and repolarization, with each segment of the signal representing a different event associated with the cardiac cycle. An ECG signal contains P wave, QRS complex and a T wave. These units of electrical activity can be further broken down into the PR interval, the ST segment, and the QT interval. End of the QRS complex and beginning of the ST segment marked by the J point. **Figure 1** illustrates a general indication of the P-wave, QRS complex, T-wave and **Table 1** illustrates the standard value of frequency and duration of waves, segments and intervals of ECG signal. The deviation from these standard values of the normal ECG signal leads to various cardiac abnormalities.

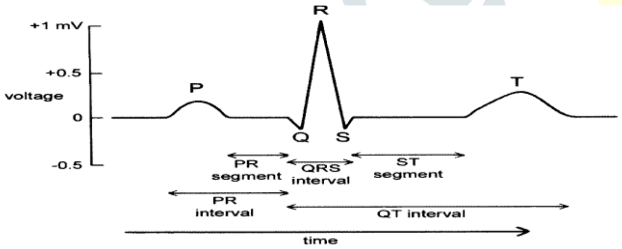

Fig1: A typical ECG wave form

At present Cardiac diseases are most common problem of human life, Early diagnosis and treatment are crucial to Ensure sustainable medical treatment which may improved survival rates. Electrocardiogram is most common diagnostic tool, used for recording various feature of heart, but this tool needs an expert intervention that's why continuous monitoring is not possible. So our aim is to developed an auto ECG diagnostic program which can be operated by any common man for continuous monitoring . But less expert intervention may introduce more noise, so noise handling is a major issue in continuous ECG monitoring . In this present work we generate noisy ECG database using a suitable algorithm ,developed on matlab and then use this generated database on another algorithm of dection and identification of ECG wave by histogram approch [2] to observe noisy ECG signals various feature and compare it with original ECG signal.

Table1: Amplitude and duration of various part of ECG signal.

SL No.	Feature	Frequency(HZ)	Duration(ms)
1	P wave	8-10	60-80
2	QRS complex	10-40	80-120
3	T wave	5-8	60-120
4	PR interval	-	120-200
5	QT interval	-	360-440
6	ST interval	-	100-120

II. METHODOLOGY

In this work clean ECG samples were taken from CSE multi-lead ECG database . Each record of the database includes 15 simultaneously measured signals: the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank lead ECGs (vx, vy, vz). Each signal is digitized at 500 samples per second with 16 bit resolution and measured in micro-volt range. In order to generate noisy database the MIT- BIH PhysioNet noise stress test database (NSTDB) [5],[6] has been used as noise samples. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. The database contains samples for three types of noise; They are denoted as bw, em and ma.

Since the sampling rates were different for noise and clean database a re-sampling has been done to make the number of samples equal for noisy data and clean data.For generation of the 12 lead noisy ECG data, noise samples (bw, em or ma) were taken from the MIT- BIH NSTDB, But as it has only two lead of noise data, so to generate third orthogonal lead of NSTDB Principal Component analysis(PCA) has been used, This method was described by [3].A Matlab function developed by him to generate principal components of the NSTDB two lead ECG data and the 1st column of the transposed orthogonal matrix gives the third orthogonal lead's data. Assuming the resultant leads form an orthogonal lead set with arbitrary orientation, the Dower transformation [8] was then used to create realistic correlated 12-lead sets of noise.In the present experiment we use 125 clear ECG data from CSE multi-lead database to generate noisy ECG database. Here in this work SNR (Signal to noise ratio) has great importance, SNR can be defined as the ratio of the power of a signal (original or meaningful information) to the power of noise (unwanted signal) ,expressed as : **[SNR= P $_{signal}$ / P $_{noise}$]**.

A SNR or signal to noise ratio compares a level of signal power to a level of noise power, It is most often expressed as a measurement of decibels(dB).Higher value of SNR generally mean a better specification, The SNR have been maintained for these data are -6db, 0db, 6db, 12db and 24db. A coefficient η has been computed to control the signal to noise ratio (SNR). The SNR, *S*, was controlled by computing the coefficient η for each lead as follows:

$$N_{ecg} = C_{ecg} + \eta * v \quad , \quad \eta = \sqrt{\exp\left(\frac{-\ln(10)*s}{10}\right)\frac{P_C}{P_v}}$$

where N_{ecg} being the output signal (i.e. the noisy signal generated from the clean sample), C_{ecg} is the initial 10 s clean ECG signal, *v is* a 10 s noisy sample from *NSTDB* selected at random, η is the amount of the noisy sample that we add to the clean signal, P_C is the power of the clean signal and P_v is the power of noisy signal.After generation of noisy ECG database , those 12 lead noisy ECG data are used, to analyze the performance of previously approached algorithm for detection and identification of ECG waves by histogram approach [2] .At first this Algorithm help in filtering and base line correction of generated signals. Then using this Algorithm P wave, QRS complex ,T wave, Duration of QT,PR,QRS and ST intervals of noisy ECG signals are try to be detected and verified. The detection of the QRS complex is the fundamental and most important component for feature extraction of any ECG signal, for which we have to concentrate on identifying the R point for each beat of the signal. All other characteristic points of ECG signals are detected with respect to that R point. Thus to detect an accurate QRS complex is an important task in ECG analysis.

III. RESULT

In this present work 125 clear ECG data from CSE multi-lead database has been used, to generate noisy ECG database using three type of noises bw, ma, and em for different values of SNR. A sample of generated noisy ECG signals of different leads with above mentioned SNR is given below :

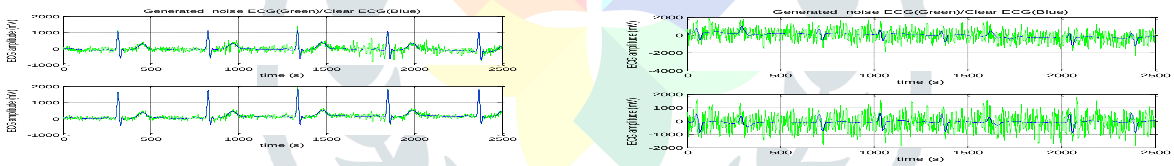

Fig.2. Clear ECG sample in blue and noisy sample, After adding *ma* in green for Left : MO1_008 of CSEDB for lead V3 and V4, SNR=6db, Right :MO1_122 of CSEDB for lead V5 and V6, SNR= -6db.

After Generating noisy ECG signals each lead are verified using algorithm for detection and identification of ECG waves by histogram approach [2] . sample of verified signal are given in Fig.3.While verifying , P wave, QRS complex and T wave of noisy ECG signals are detected and duration values of PR,QRS,QT,ST intervals of noisy ECG signals is given by the algorithm. But as shown in Fig .3 by the decreasing value of SNR the signals are distorted very badly and the position of P wave, QRS complex and T wave can not identified correctly and difference between duration of PR,QRS,QT,S intervals of noisy signal with original is very large. Percentage error of this durations are calculated using the formula:

[| Approximate value – Exact value | / |Exact value|* 100%]

Using above formula in Excel there are many tables are generated for different parameters of about 125 different noisy ECG signals. Average Calculated errors of different intervals are shown below in Table- 2.

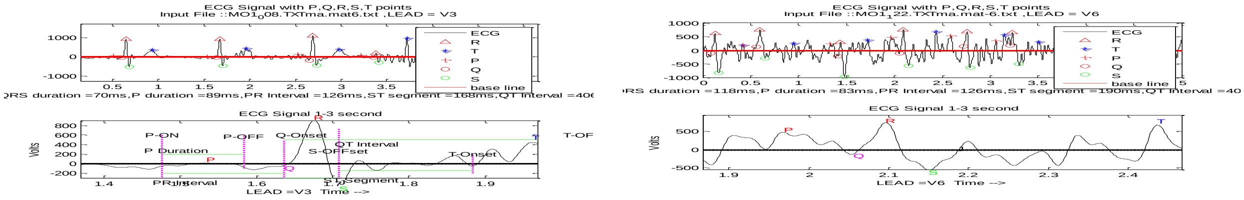

Fig.3. Using ma: Left : Characteristic feature of lead V3 of MO1_008 of CSEDB, SNR=6db . Right : Characteristic feature of lead V6 of MO1_122, SNR=-6db.

Table 2: Average percentage error calculation of PR,QRS,QT and ST intervals with different SNR.

Average errors in %													
No.	**SNR**	**PR interval**			**QRS interval**			**QT interval**			**ST interval**		
		ma	**em**	**bw**	**ma**	**Em**	**Bw**	**ma**	**em**	**bw**	**ma**	**Em**	**bw**
1.	24	18	16	17	9	9	7	4	3	3	8.53	11.39	11.80
2.	12	18.16	19	19	8	9	8	5	5	5	17	17	18
3.	6	20	20	20	14	14	15	6	6	6	32	32	33
4.	0	21	21	20	18	19	19	7	7	9	39	38	39
5.	-6	22	22	22	23	23	24	10	11	12	50	50	51

IV. CONCLUTION

A realistic noisy ECG database has been generated in this present work and used this database on a algorithm of detection and identification of ECG waves by histogram approach [2]. By running this algorithm using noisy ECG database we see that the performance of this software decreases gradually with introduction of noise though it works good for clear ECG. By Fig3. One can Clearly observe the above mentioned situation after that percentage error of each ECG lead has been calculated and form Table. 2 one can observed that Maximum error is obtained in ST interval. So in future we try to developed such filtering effect using software which can minimize the effect of noise, and can be operated by any common person very easily.

References :

[1] Anthony Dupre,Ms,Sarah Vicent,Ms, Paul A ,Iaizzo,Phd.Basic ECG Theory ,recording and interpretation.

[2] B.Halder, S.Mitra, M.Mitra, Detection and Identification of ECG waves byHistogram approach,2016.

[3] G D Clifford, J Behar1, Q Li and I Rezek, Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms, Physiol. Meas. Vol. **33,** 1419–1433, (2012)

[4] S P Nelwan Evaluation of 12-Lead Electrocardiogram ReconstructionMethods for PatientMonitoring, Ph.D. Thesis, Erasmus MC, November (2005)

[5] Moody G B, Muldrow W E and Mark R G, A noise stress test for arrhythmia detectors *Comput. Cardiol.*11 381–4, (1984).

[6] Goldberger A L, Amaral L A N, Glass L, Hausdorff J M, Ivanov P C, Mark R G, Mietus J E, Moody G B, Peng C-Kand Stanley H E , PhysioBank, Physiotoolkit, and PhysioNet: components of a new research resource for complex physiologic signals *Circ.* 101 e215–20 PMID: 1085218, (2000)

[7] Fye, W. B., A history of the origin, evolution, and impact of electrocardiography. *Am J Cardiol*,73, 1214–1219, (1994)

[8] Dower, G. E., Machado, H. B., and Osborne, J. A. ,On deriving the electrocardiogram from vectorcardiographic leads. *Clin Cardiol*, 3(2), 87–95, (1980).

[9] Drew, B. J., Califf, R.M., Funk, M., Kaufman, E. S., Krucoff, M.W., Laks, M. M.,Macfarlane, P.W.,Sommargren, C., Swiryn, S., and Van Hare, G. F., Practice standards for electrocardiographic monitoring in hospital settings: an American Heart Association scientificstatement from the Councils on Cardiovascular Nursing, Clinical Cardiology, and Cardiovascular Disease in the Young. *Circulation*, 110(17), 2721–2746, (2004)

[10] Krucoff, M. W., Parente, A. R., Bottner, R. K., Renzi, R. H., Stark, K. S., Shugoll, R. A., Ahmed, S. W.,DeMichele, J., Stroming, S. L., and Green, C. E., Stability of multilead ST-segment "fingerprints"over time after percutaneous transluminal coronary angioplasty and its usefulness in detecting reocclusion. *AmJ Cardiol*, 61(15), 1232–1237, (1988)