Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 14 (14): 1628-1632, 2014
ISSN 1812-5654 / DOI: 10.3923/jas.2014.1628.1632
© 2014 Asian Network for Scientific Information

K-nearest Neighbour Method of Analysing the ECG Signal (To Find out the Different Disorders Related to Heart)

[1]S. Jayalalitha, [2]D. Susan, [1]Shalini Kumari and [1]B. Archana
[1]Department of Electronics and Instrumentation Engineering,
[2]Department of Electronics and Communication Engineering, SASTRA University, Thanjavur, India

Abstract: Heart attack is very common thing now-a-days. This results because of many disorders occurring in the heart. These disorders can be found out by analysing the ECG signal. The conventional method of analysing the ECG signal is done by the doctors who are expert in that. So far there is no systematic procedure for analysing the ECG signal. The main goal of this study is to obtain a systematic method of analysing the Electrocardiogram (ECG) from the patient and display the types of disorder. The ECG from the patient is taken and by loop analysis every point in the data is checked and analysed using two new methods namely the direct comparison method and kNN method. The ECG parameters such as shape of the waveform, heart rate, time interval, amplitude (voltage) etc., are taken for analysis. kNN (K-Nearest Neighbour) algorithm is a search algorithm which finds the nearest distance from the training data and this method is used to compare the corresponding predefined values of different disorders of the heart with the sampled data and the inference from the comparison results are displayed. In case of multiple disorders, the KNN finds the most prominent one. Simulation is done using MATLAB.

Key words: Electro cardio gram, k nearest neighbour, sinus bradycardia, heart rate, heart disorder

INTRODUCTION

Heart disease has become the most common disease that affects humans worldwide. Now a day's large number of people are experiencing heart attack due to stress and some of them are taking up coronary artery bypass surgical treatment and balloon angioplasty for progressive heart problem. Early detection and timely treatment can prevent such events. This would slow down the progression of heart failure and prevent the loss of life.

Electro Cardio Gram is a graphical representation of the electrical potentials produced during each cardiac cycle by the heart. It replicates the recurring electrical depolarisation and repolarisation of the myocardium (heart muscle) which accompanies with the narrowing of atria and ventricles. Normally body fluids consists of principle ions such as Sodium (Na^+), Potassium (K^+), Chlorine (Cl^-) responsible for producing cell potentials. The membrane of the cell readily permits K+ and Cl^- but not Na+. Hence, concentration of Na+ outside the cell is quite high, making it highly positive and highly negative (-90 mV) (Arumugam, 2006; Wagner, 2001) on the inside of the cell under the resting state. Potential measured as a result of the unequal charge distribution is called resting potential. The cell is polarized under this condition. When a voltage is applied to the cell, the cell membranes outside area temporarily turns out to be negative with respect to the inner area. A decrease in this resting membrane potential difference is called depolarization (+20 mV) (Arumugam, 2006). When the normal state of the cell is resumed the membrane is negative inside with respect to outside. This process is called repolarization (Khandpur, 2003).

These voltages are sensed on the body surface using probes fastened to the extremities and chest wall. By fastening the probes to the extremities and to the chest wall, these voltages are sensed on the body surface. Signal conditioning is done and then this signal is amplified by the electrocardiograph machine and displayed on special graph.

A true ECG reading is judged by the three waves P, qRs and T which occurs in a sequential way. The P wave being small in magnitude represents the impulse of the two regions of the heart's atria that gets the blood from the veins of the body. It is characterised by depolarization of atrial muscles. The qRs complex that has high magnitude and spikes due to the rise in the conduction velocity, measures the depolarization of the ventricles. Lastly, the depolarization and revitalization of the ventricles is made known by the T wave.

Corresponding Author: S. Jayalalitha, Department of Electronics and Instrumentation Engineering, SASTRA University, Thanjavur, India

In the diagnosis procedure, first the ECG of the patient is taken. It is a non-invasive diagnostic tool used for the assessment of a patient's heart condition. The data's obtained from ECG is distinguished by careful remarks and when shared with heart rate gives an exact and quick diagnosis.

There are different heart conditions such as sinus tachycardia, sinus bradycardia, premature atrial contraction, atrial fibrillation, idioventricular rhythm, premature junctional contraction, apnea, bundle block, myocardial infraction, coronary insufficiency that is diagnosed by this method (Webster, 2009; Proakis and Manolakis, 2007).

The most significant technological development is the computer which plays an important role in cardiac investigations. It has been said that the application of computers in cardiology is perhaps the one of the most important uses of computer in medicine. Hence, the conventional method of diagnosis by the doctor can be replaced by computer based diagnosis and the result can be shown immediately to the patient. This reduces the risk and helps them to understand their situation better and go for the relevant treatment.

Computerized interpretation is possible because of availability of digitized ECG to which mathematical techniques of signal conditioning and processing can be applied. Many software and classification methods are available for analysis (Swarnalatha and Prasad, 2010; Vijila *et al*., 2006).

In this study, MATLAB based stimulation of ECG signal analysis using k Nearest Neighbour is carried out, because MATLAB is user friendly and gives output in an understandable form with high accuracy. Next section details about the proposed study followed by kNN classifier explanation and results. The final section summarizes the conclusion of this study.

MATERIALS AND METHODS

The abnormalities of the heart are found by doctors by just observing the deviation of P, qRs, T signal from the normal signal, by observing the various parameters (such as time, amp, shape etc.,) in the ECG graph of the patient obtained from the ECG machine. The accuracy of finding the abnormalities depends on the experience of the experts. The main aim of this study is to provide a systematic and simple method of analyzing the condition of the heart (Kohama *et al*., 1999). It is an accurate, cost effective and easy handling procedure which gives the response within a few seconds for which MATLAB is used for this analysis.

The diagnosis is done for 10 disorder types. It can be further extended for more disorders by following the same methodology and algorithm. By extending it further, the values can be given to a learning algorithm using Neural Network techniques (Hendel *et al*., 2010). It can then be followed by Genetic Algorithm in case of working with large number of data's. A miniature hardware can be designed with the analysis code interfaced. This will help the patient to find their heart condition at any moment. Hence, it will decrease the risk of heart failures since the diagnosis is done at the earlier stages.

PARAMETER ESTIMATION USING MATLAB

The parameters such as P-R interval, R-R interval and peak values of P, qRs complex, T wave in an ECG signal are of prime importance in finding out the condition of the patient's heart. For this analysis purpose MATLAB is used.

ECG from the patient is obtained in discrete form and then loaded into MATLAB for signal processing. The various MATLAB functions such as get file, load, length, square root sort etc., are used. Heart rate is calculated by considering the number of cycles in 60 sec since one ECG cycle represents one heart beat.

$$\text{ECG rate} \quad \frac{\text{int}32(\text{fix}(\text{cyl}))}{\max(\text{time})} \times 60$$

Then, using the loop analysis every point in each cycle in the waveform is checked and a factor is calculated that relates the total number of samples in one ECG cycle and the number of samples in a particular wave in that cycle of ECG waveform. It helps in finding out each peak of the waves such as P, Q, R, S and T waves in each cycle using the functions max and min. This calculation of the factor is based on the observation of different real time data. Duration of the waves and the intervals between them are essential for ECG analysis. The analysis is done as follows:

- **R-R interval:** R-R interval is calculated by finding the difference in corresponding time of the two consecutive R-R peaks
- **p-R interval:** p-R interval is calculated by finding the difference in the corresponding time of p and R wave peak in one cycle
- **qRs complex duration:** qRs complex duration is calculated by finding the difference in the corresponding time of q and s wave peak in one cycle

- **s-T interval:** s-T interval is calculated by finding the difference in the corresponding time of s and T wave peak in one cycle

By calculating the slope of the wave from the baseline it is found, if the T wave is upright or depressed. Using the same method it is found if the s-T segment is elevated or normal. The amplitude of each wave is calculated from the baseline. In case of baseline shift, the value of shift is considered. The mean of all the parameters are calculated and used for further analysis. Next, the analysis code is split into two parts. The first part is, finding the disorder by comparison and the second part is the use of algorithm which follows the concept of minimum distance.

COMPARISON TO IDENTIFY THE DISORDERS USING ECG PARAMETERS

ECG characteristics of input are compared with different heart arrhythmias of predefined values. Average of R-R interval, P-R interval, qRs duration and s-T interval for the signal is compared with the standard values of various disorders. Thereby the condition of the heart along with the disorders (if any) of the patient is displayed (Babikier *et al*., 2011). In certain cases, a patient may suffer from more than one heart disorder. Then, kNN classification (Liu *et al*., 2010), is applied for finding the most prominent one. Table 1 shows the various conditions of the heart for various abnormality conditions (ECG Review, 2001).

KNN CLASSIFIER

kNN algorithm: k- Nearest Neighbour might be the most simplest of all the classification algorithms. This method is usually used to carry out classifications. The classification is performed by finding the minimum distance from a data set which contains the input or training data and data set which contains the reference values.

K-nearest-neighbour classification was developed from the need to perform discriminant analysis. In order for the classification to occur:

- Assume the availability of training data
- Size of the training data set (k)

For each training data item, distance is calculated using distance equation and then arranged in ascending order. x and y are the true and measured value of ECG parameters and N is the number of values taken:

$$\text{Distance } d(x,y) = \sum_{i=1}^{N} \sqrt{(x_i - y_i)^2}$$

As this method is intuitive, simple to implement and understandable, it is used for biomedical research purposes. Table 2 shows the training data for various heart conditions.

Training data:

- A = [knn_hr,pr_int_avg,qrs_dur_avg, st_int_avg, knn_sig_regular,knn_st,knn_t]

RESULTS AND DISCUSSION

Figure 1a and b shows the ECG signal obtained from two patients having the heart disorder.

Table 1: Variation in the parameters of ECG signal showing different heart disorders

Heart conditions	Rhythm	Heart rate (beats min^{-1})	P-R interval (sec)	QRS complex duration (sec)	ST segment
Normal sinus rhythm	Regular	60 to 100	0.12 to 0.20	0.04 to 0.10	Normal in baseline
Sinus tachycardia	Regular	>100	0.12 to 0.20	0.04 to 0.10	Normal in baseline
Sinus bradycardia	Regular	<60	0.12 to 0.20	0.04 to 0.10	Normal in baseline
Apnea	Regular	60 to 67	0.12 to 0.20	0.04 to 0.10	Normal in baseline
Bundle block	Regular	---	---	>0.10	Normal in baseline
Coronary insufficiency	Regular	---	---	---	Depressed
Myocardial infraction	Regular	---	---	---	Elevated
Idioventricular rhythm	Regular	20 to 40	---	>0.12	Normal in baseline
Premature junction contraction	Irregular	<100	<0.12	0.04 to 0.10	Normal in baseline
Premature atrial contractions	Irregular	<100	0.12 to 0.20	0.04 to 0.10	Normal in baseline
Atrial fibrillations	Irregular	400-700	Not visible	---	Normal in baseline

Table 2: Training data using kNN classifier for various heart disorders

Heart condition	Reference data	Heart condition	Reference data
Normal sinus rhythm	A1 = [0.80, 0.16, 0.07, 0.10, 0.1, 0, 0]	Sinus bradycardia	A2 = [0.60, 0.16, 0.07, 0.10, 0.1, 0, 0]
Sinus tachycardia	A3 = [1.10, 0.16, 0.07, 0.10, 0.1, 0, 0]	Apnea	A4 = [0.65, 0.16, 0.07, 0.10, 0.1, 0, 0]
Premature atrial contractions	A5 = [0.9, 0.16, 0.07, 0, 0, 0]	Atrial fibrillation	A6 = [5, 0, 0, 0]
Idioventricular rhythm	A7 = [0.75, 0.15, 0.1, 0, 0]	Bundle block	A8 = [0.75, 0.15, 0.1, 0, 0.1]
Coronary insufficiency	A9 = [0.8, 0.1, 0, 0.1]	Myocardial infraction	A10 = [0.9, 0.16, 0.1, 0.1, 0]
Premature junction contraction	A11 = [0.9, 0.06, 0, 0, 0]		

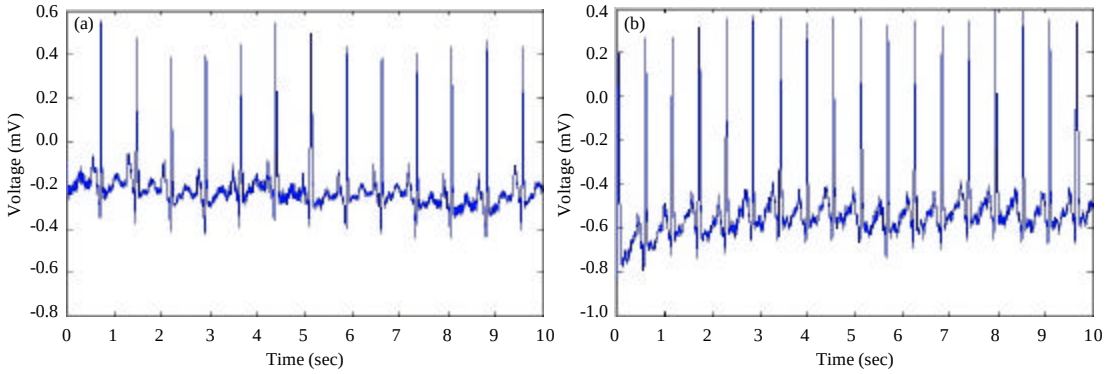

Fig. 1(a-b): (a) Sample of ECG taken from a patient (sample-1) and (b) Sample of ECG taken from a patient (sample-2)

Table 3: Result obtained from the analysis of disorders using the comparison method and kNN method

ECG parameters of patient 1		ECG parameters of patient 2	
Sample_int	1.0000e-003	Sample_int	1.0000e-003
Enter number of cycles/peaks from the time graph	13	Enter number of cycles/peaks from the time graph	18
heart_rate	78	heart_rate	108
rr_val_avg	0.4502	rr_val_avg	0.3340
rr_int_avg	0.7386	rr_int_avg	0.5668
pr_int_avg	0.1549	pr_int_avg	0.1526
qrs_dur_avg	0.0570	qrs_dur_avg	0.0916
st_int_avg	0.2298	st_int_avg	0.1086
signal is regular		signal is regular	
sT segment normal		sT segment elevated	
t wave depressed		t wave depressed	
Comparison output	Coronary insufficiency	Comparison output	Sinus tachycardia Myocardial infraction Coronary insufficiency
Knn output	Coronary insufficiency	Knn output	Sinus tachycardia

Analysis is made using kNN technique and the inference from the analysis is given Table 3.

It is seen that using the comparison method a person may have more than one disorder. Hence he or she may not be able to realise as to which disorder requires the most immediate treatment. After this, the kNN algorithm is used which classifies the disorder based on the nearest neighbour of the input thereby giving a clear idea of the most prominent one in this condition. The analysis is done using real time data of patients aged between 20 to 40 years. Lead II readings are taken since it has large peak amplitude. Ferretti *et al*. (1992) explains the measurement of QT interval using linear regression method which has a slight effect on the sampling rate. Suppappola and Sun (1994), on the other hand explains about the MOBD algorithms which suggest a tradeoff between accurateness and response time. In this study, kNN method used for the identification of the disorder is precise.

CONCLUSION

As the technologies developed there was a great deal of advancement in the field of medical electronics. The diagnosis and analysis by expertise (doctors) was not sufficient to find out the heart disorders and for the increase in the cardiac problem cases that are seen by the hospitals every day. Hence, a more advanced, systematic and procedural fast analysis technique is required and this study is done for fulfilling the above requirement. To find the disorders of the heart many methods have been adopted so far in research but this study concentrates on developing two new methods namely direct comparison method and kNN method. Samples of ECG have been taken from two patients and the parameters are evaluated. Based on the evaluation the disorders of the heart are identified. Out of these two methods of study, kNN method gives better analysis of heart disorders.

REFERENCES

Arumugam, M., 2006. Biomedical Instrumentation. 2nd Edn., Anuradha Agencies, Kumbakonam, India.

Babikier, M.A., M. Izzeldin, I.M. Ishag, D.G. Lee and K.H. Ryu, 2011. Classification of cardiac arrhythmias using machine learning techniques based on ECG signal matching. http://sustech.edu/staff _publications /20111215103300603.pdf

ECG Review, 2001. ACLS program Ohio State University medical center. Department of Educational Development and Resources, OSU Medical Center.

Ferretti, G.F., L. Re, M. Zayat, D. Mazzara and C. Rimatori *et al*., 1992. A new method for the simultaneous automatic measurement of the RR and QT intervals in ambulatory ECG recordings. Proceedings of the Computers in Cardiology, October 11-14, 1992, pp: 171 174-10.1109/CIC.1992.269419.

Hendel, M., A. Benyettou, F. Hendel and H. Khelil, 2010. Automatic heartbeats classification based on discrete wavelet transform and on a fusion of probabilistic neural networks. J. Applied Sci., 10: 1554-1562.

Khandpur, R.S., 2003. Handbook in Biomedical Instrumentation. 2nd Ed., Tata McGraw-Hill Publishing Co., New Delhi, Pages: 994.

Kohama, T., S. Nakamura and H. Hoshino, 1999. Preprocessing MAMD and RSEF for R-R interval detection of ECG. Proceedings of the IEEE Region 10 Conference, Volume 2, December 1999, Cheju Island, pp: 1162-1165.

Liu, G.P., G.Z. Li, Y.L. Wang and Y.Q. Wang, 2010. Modelling of inquiry diagnosis for coronary heart disease in traditional Chinese medicine by using multi-label learning. BMC Complementary Alternative Med., Vol. 10. 10.1186/1472-6882-10-37

Proakis, J.G. and D.G. Manolakis, 2007. Digital Signal Processing: Principles Algorithm and Applications. 4th Ed., Pearson Education, Inc., India, ISBN-13: 9788131710005, Pages: 1156.

Suppappola, S. and Y. Sun, 1994. Nonlinear transforms of ECG signals for digital QRS detection: A quantitative analysis. IEEE Trans. Biomed. Eng., 41: 397-400.

Swarnalatha, R. and D.V. Prasad, 2010. Maternal ECG cancellation in abdominal signal using ANFIS and wavelets. J. Applied Sci., 10: 868-877.

Vijila, C.K.S., M.E.P. Kanagasabapathy and S. Johnson, 2006. Fetal ECG extraction using softcomputing technique. J. Applied Sci., 6: 251-256.

Wagner, G.S., 2001. Practical Electrocardiography. 10th Edn., Lippincott, Williams and Wilkins, Philadelphia.

Webster, J., 2009. Medical Instrumentation Application and Design. 3rd Edn., John Wiley and Sons, New York.