COPYRIGHT © 2011 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 01, ISSUE 02, MANUSCRIPT CODE: 110135

Identification of Atrial Fibrillation (AFIB) of Heart Using Robust Statistical Tools and Approximate Entropy Method

A. H. M. Zadidul Karim *and* Md. Meganur Rhaman

Abstract— Atrial fibrillation (AFIB) is the most common arrhythmia encountered in clinical practice, affecting about 0.5-1% of the general population. Atrial fibrillation is not immediately life threatenting. However, of late it has become clear that AFIB often leads to severe complications such as heart failure and stroke. Atrial Fibrillation (AFIB) is an arrhythmia, an irregularity of the heart's rhythm. The irregularity of the impulses traveling down from the atria makes the ventricles beat irregularly. Sometimes AFIB can make the pulse fast and irregular or slow and irregular. AFIB alone is not a life-threatening arrhythmia, but it can be extremely bothersome and sometimes dangerous. A total of 16 sets of ECG recordings, 8 with normal rhythm and 8 of Atrial fibrillation (AFIB) patients are analyzed. By using some robust statistical tools decision making algorithm are designed in order to clarify normal and Atrial fibrillation (AFIB). This project work is based on time domain analysis and some non-linear methods. Heart rate variability (HRV) dataset are collected from MIT-BIH Arrhythmia data bank On this dataset of IHR, various time domain parameters like mean, variance, standard deviation (SD), the standard deviation of successive RR interval differences (SDSD), the root mean square of successive differences (RMSSD) and some non-linear parameters like Poincare plot parameters, approximate entropy (ApEn) are determined. The result obtained from the application of these techniques is analyzed to distinguish the ECG signals between the healthy person and that of the Atrial fibrillation (AFIB).

Index Terms— AFIB, ECG, Approximate Entropy, Time domain analysis.

1 INTRODUCTION

AFIB is the most common type of heart arrhythmia. People of all ages can get AFIB Young people with otherwise healthy normal hearts can develop AFIB [10]. It is most often found in older people with some other heart disease. It affects about 15% of people over age 85. Visual inspection of the ECG gives insights to cardiac arrhythmia disorders, but a proper analysis requires extracting hidden information from the ECG signal. The aim of our research is analyzing ECG signals, with the particular focus on Atrial Fibrillation arrhythmia condition. Variation in heart rate may be evaluated by a number of methods. Possibly the simplest to perform are the time domain measures. With these methods either the heart rate at any point in time or the intervals between successive normal complexes are determined. The statistical analysis of the calculated features indicate that they differ significantly between normal heart rhythm and the different arrhythmia types and hence, can be rather useful in ECG arrhythmia detection. The five statistical parameters considered for cardiac arrhythmia classification of the ECG signals are the mean of RR intervals, the variance of RR in tervals, the standard deviation of the RR intervals (SDNN), the standard deviation of differences between adjacent RR intervals (SDSD) and the root mean square successive difference of intervals which are extracted from heart rate signals (RMSSD).Time domain methods are easy to program in Matlab because of its simple mathematical expression. There is a wide range of variation found between normal and AFIB. We can easily separate normal and AFIB beat by using time domain method. For this reason I choose time domain analysis for separating normal and AFIB rhythms.

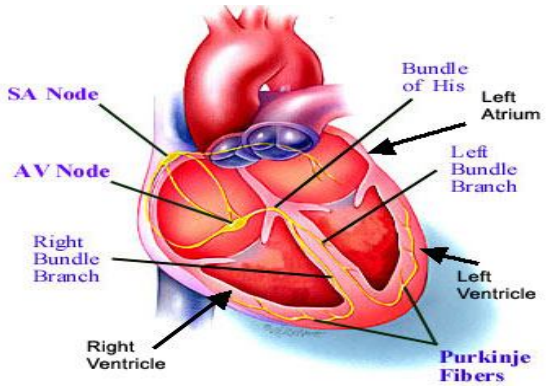


Fig. 1: Normal electrical conduction, arising from SA node and transmitted down pathway (outlined in yellow).

- *A.H.M. Zadidul Karim is with the Department of Electrical and Electronic Engineering, University of Asia Pacific (www.uap-bd.edu), Dhanmondi, Dhaka-1209, Bangladesh. E-mail: zadid_aust@yahoo.com*
- *Md. Meganur Rahman is with the Department Electrical and Electronic Engineering,Ahsanullah University of Science and Technology (www.aust.edu), Tejgaon Industrial Area, Dhaka-1208. E-mail: mizan_aust@yahoo.com*

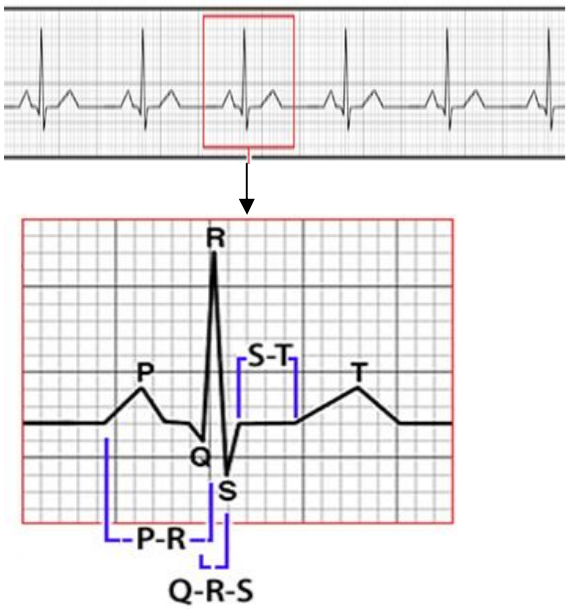


Fig. 2: A Typical Healthy ECG

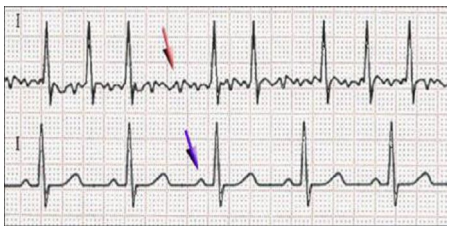


Fig. 3: AFIB Sustained ECG (top) vs. Healthy ECG (bottom)

Quantification of the unpredictability and complexity of the heart rate using approximate entropy [5] are being increasingly used because they can be computed from shorter ECG records. The aim of this study was to determine how and which of the variability and complexity parameters of the HRV derived from the Time domain analtsis and approximate entropy are different in patients with AFIB compared with subjects with normal rhythm. To evaluate the act of the proposed method, we used the MIT-BIH Arrhythmia Database directory. The ECG signals in this directory were sampled at 360 Hz and with a quantization resolution of 11bits/sample [11].

2 METHODS

An algorithm based on indexes of time domain analysis, approximate entropy, Time domain analysis are used to distinguish between ECG of normal and AEIB subjects. A total of fourteen ECG data sets are included in the study, seven with AFIB whilst the remaining seven with normal rhythm. The data sets of ECG are taken from MIT-BIH arrhythmia database.

2.2 Time domain analysis

Variation in heart rate may be evaluated by a number of methods. Possibly the simplest to perform are the time domain measures. With these methods either the heart rate at any point in time or the intervals between successive normal complexes are determined[1]. The statistical analysis of the calculated features indicate that they differ significantly between normal heart rhythm and the different arrhythmia types and hence, can be rather useful in ECG arrhythmia detection[3,7]. The five statistical parameters considered for cardiac arrhythmia classification of the ECG signals are the mean of RR intervals, the variance of RR intervals, the standard deviation of the RR intervals (SDNN), the standard deviation of differences between adjacent RR intervals (SDSD) and the root mean square successive difference of intervals which are extracted from heart rate signals (RMSSD).Time domain methods are easy to program in Matlab because of its simple mathematical expression. There is a wide range of variation found between normal and abnormal rhythms. We can easily separate normal and abnormal heart beat by using time domain method [8]. For this reason I choose time domain analysis for separating normal and abnormal rhythms.

Mean IHR

In this paper the mean is obtaining as

$$Mean = \frac{RR_1 + RR_2 + ... + RR_N}{N}$$

The sum of a list of numbers, divided by the total number of number in the list where $(RR_1, RR_2, \ldots\ldots, RR_N)$is the sample here total number of sample is N.

Variance: The variance is a measure of the dispersion of a set of values. The variance is the mean of the sum of the squares of the differences between the values and the Variance: The variance is a measure of the dispersion of a set of values. The variance is the mean of the sum of the squares of the differences between the values and the mean of the sample [7].

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(RR_i - \overline{RR}\right)$$

The most obvious such measure is the mean value of RR intervals . In addition, several variables that measure the variability within the RR series exist. The standard deviation of RR intervals (SDNN) is defined as

$$SDNN = \sqrt{\frac{1}{N-1}\sum_{i=1}^{n}\left(RR_i - \overline{RR}\right)^2}$$

The standard deviation of successive RR interval differences (SDSD):
The standard deviation of successive interval differences (SDSD) given by

$$SDSD = \sqrt{\mathrm{var}(RR_N - RR_{N+1})}$$

SDSD can be used as a measure of the short-term variability.

The root mean square of successive differences (RMSSD)
The most commonly used measures derived from interval differences include the square root of the mean squared differences of successive NN intervals [7]. Calculation of root mean square is show in equation 5.

$$RMSSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} RR_i^2}$$

2.3 Approximate entropy analysis

Entropy is related to dynamical systems and is defined as the rate of information production. Two types of entropy, namely approximate entropy (ApEn) and sample entropy (SE) are calculated to determine the complexity of dynamical systems[2]. The algorithm for calculating AE and SE is almost similar. In our study, we used AE measure as it measures the system complexity more closely. This entropy measure is frequently applied to clinical cardiovascular and other time series analysis of different types of abnormalities [5, 8]. ApEn is a ``regularity statistic'' that quantifies the unpredictability of fluctuations in a time series. Intuitively, one may reason that the presence of repetitive patterns of fluctuation in a time series makes it more predictable than a time series in which such patterns are absent. ApEn reflects the likelihood that ``similar'' patterns of observations will not be followed by additional ``similar'' observations[4]. A time series containing many repetitive patterns has a relatively small ApEn; a less predictable (i.e., more complex) process has a higher ApEn. A brief summary of the calculations, as applied to a time series of heart rate measurements, $RR(i)$ Given a sequence N, consisting of N instantaneous heart rate measurements $RR(1), RR(2), ..., RR(N)$ we must choose values for two input parameters, m and r, to compute the approximate entropy, ApEn(N,m,r), of the sequence. The second of these parameters, m, specifies the pattern length, and the third, r, defines the criterion of similarity [9]. We denote a subsequence (or pattern) of heart rate measurements, beginning at measurement i within N, by the vector $p_m(i)$. Two patterns $p_m(i)$ and $p_m(j)$, are similar if the difference between any pair of corresponding measurements in the patterns is less than r, i.e., if

$|p_m(i) - p_m(j)| < r$ for $0 \le k < m$

Now consider the set $p_m(i)$ of all patterns of length m [i.e., pm (1), pm (2), pm (3),....., pm(N-m+1)], within N. For each subject and for each of the two experimental conditions, we considered a time series of N=2000 consecutive values; we calculated the standard deviation, SD, of and evaluated ApEn setting m=2 and with r increasing from 0.1 to 0.9. For a given r value, ApEn(r) was calculated.

We may now define: $C_i^m(r) = \dfrac{n_i^m(r)}{N-m+1}$

Where $n_i^m(r)$ is the number of patterns in Pm that are similar to $p_m(i)$ (given the similarity criterion r). The quantity $C_i^m(r)$ is the fraction of patterns of length that resemble the pattern of the same length that begins at interval i. We can calculate $C_i^m(r)$ for each pattern in Pm, and we define $C_i^m(r)$ as the mean of these $C_i^m(r)$ values. The quantity $C_i^m(r)$ expresses the prevalence of repetitive patterns of length m in N. Finally, we define the approximate entropy of N, for patterns of length and similarity criterion r, as

$$ApEn\ (N, m, r) = \ln\left[\frac{C_i^m(r)}{C_i^{m+1}(r)}\right]$$

Thus, if we find similar patterns in a heart rate time series, ApEn estimates the logarithmic likelihood that the next intervals after each of the patterns will differ[6]. Smaller values of ApEn imply a greater likelihood that similar patterns of measurements will be followed by additional similar measurements. If the time series is highly irregular, the occurrence of similar patterns will not be predictive for the following measurements, and Apen will be relatively large. In this project, we have estimated ApEn using the command of java program. For this program, we have taken instantaneous heart rate (IHR) values which are taken from MIT-BIH Arrhythmia Record. We use java vector class for intermediate data storage. However, after storing pertinent data, we constructed matrix of length 2 by taking two values from the data storage which are represented as

$$RR_1, RR_2, RR_3, RR_4, RR_5, RR_6, ..., RR_N$$

According to the algorithm of the ApEn we derived different values of approximate entropy. The flow chart for this algorithm is shown in fig (4.1).The comparison between two matrixes follows a threshold value which varies from 10% to 90% of the standard deviation for the data taken as input from IHR values. The comparison follows the following formula, $|vector(i+k) - vector(j+k)| <$ Per*standard deviation for $0 \le k < m$ Where per varies from 0.1 to 0.9 and vector is the object of Vector class in which we stored IHR values for our program[5]. The value of m represents the pattern length and the right side of the above inequality means the criterion of similarity.

In our program, to calculate ApEn we evaluate the mean of the fraction of patterns of length m that resemble the pattern of the same length whose value depends on the distance between two vector. We attribute seven data set of healthy people as input and seven data set of AFIB and after analyzing by our program, we come across the discernible differences in output between healthy people and the people with AFIB for the value of the criterion of similarity which is greater than 0.2.

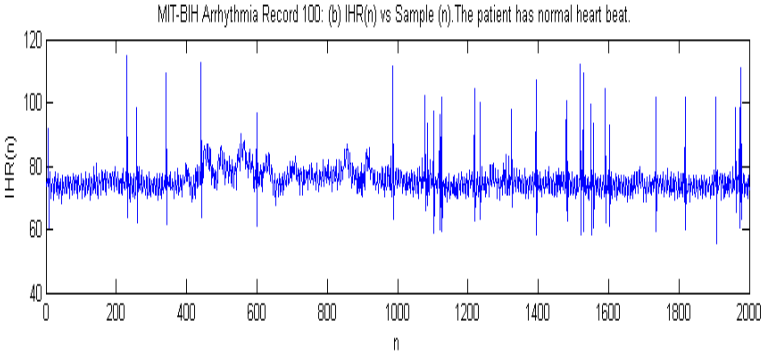
We attribute seven data sets of healthy people as input and seven data set of AFIB patient and after analyzing by our study, we come across the discernible differences in output between healthy people and the people with AFIB for the value of the criterion of similarity which is greater than 0.2.

4 RESULT

Fourteen ECG recordings were analyzed. Fig. 4 and 5 depict the IHR time series of typical normal and AFIB data sets. The time domain(TD) parameters (Mean, SD, Variance, SDSD and RMSSD) were measured for each data set and the average values were calculated. Table 1 summarizes the results from TD indexes of the two groups. Here, the average values of 7 data sets (MIT-BIH data set # 100, 105, 111, 112, 116, 118 and 121) of normal ECG are computed. Same is done for rest 7 data sets (201, 202, 210, 217, 219,221 and 222) of AFIB patients. The ApEn are analyzed to see if any significant difference is found between normal and AFIB data series.

From Table 1, it is obvious that there is a clear reduction of time domain parameters in the healthy group. Table 2 shows the time domain parameters of AFIB. From table 3 we found a significant difference between the two groups exists. The average value of variance for healthy group is 57.13 while that for AFIB is 560.25. The average of SD is 6.8 and 22.54 for healthy and AFIB groups, respectively. A significant difference was found between two groups as the SDSD (8.74) for healthy group is much lower than that of AFIB group (33.51). Also, a significant difference exists in the RMSSD between the two groups, the value being 8.66 and 26.19 for healthy and AFIB groups, respectively.

Fig. 4: (a) IHR of MIT-BIH Record_100: The patient has normal heart



beat

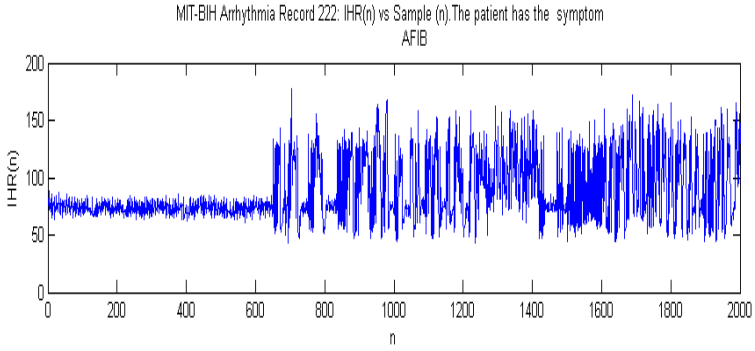


Fig. 5: IHR of MIT-BIH Record_222. The patient has AFIB beats

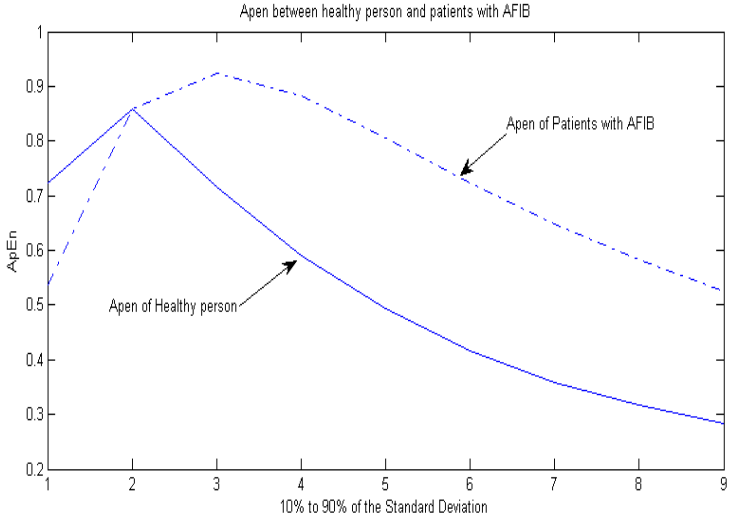


Fig. 6: ApEn between the patients has normal heart beat vs. AFIB

TABLE 1 HRV PARAMETERS IN THE TIME DOMAIN FOR THE PATIENT HAS THE SYMPTOM OF AFIB.

Record	Mean	Variance	SDNN	SDSD	RMSSD
201	73.70	855.65	29.24	31.82	29.20
202	80.16	986.35	31.39	20.47	23.56
210	91.40	357.23	18.89	26.55	26.55
217	74.66	106.33	10.30	65.30	13.58
219	75.53	259.66	16.11	21.28	21.28
221	87.69	556.88	23.59	36.82	36.84
222	87.46	799.66	28.27	32.32	32.30

TABLE 2 HRV PARAMETERS IN THE TIME DOMAIN FOR HEALTHY GROUPS

Record	Mean	Variance	SD	SDSD	RMSSD
100	75.54	24.72	4.97	6.71	6.70
105	85.86	122.89	11.08	15.70	15.12
111	70.51	14.23	3.77	3.97	3.97
112	84.99	6.49	2.55	2.52	2.52
116	80.59	113.13	10.63	17.40	17.40
118	75.64	82.21	9.07	11.29	11.29
121	62.02	36.21	6.02	3.58	3.59

TABLE 3 AVERAGE HRV PARAMETERS IN THE TIME DOMAIN.

Average	Healthy groups	AFIB
Mean	76.45	81.51
Variance	57.13	560.25
Standard Devision	6.87	22.54
SDSD	8.74	33.51
RMSSD	8.66	26.19

TABLE 4 APPROXIMATE ENTROPY VALUES OF HEALTHY GROUPS AND THE GROUPS WITH AFIB

Approximate Entropy	Healthy groups	AFIB
Apen (m=2,r=0.1SD)	0.7230	0.5378
Apen (m=2,r=0.2SD)	0.8575	0.8581
Apen (m=2,r=0.3SD)	0.7158	0.9248
Apen (m=2,r=0.4SD)	0.5896	0.8826
Apen (m=2,r=0.5SD)	0.4944	0.8065
Apen (m=2,r=0.6SD)	0.4163	0.7236
Apen (m=2,r=0.7SD)	0.3576	0.6492
Apen (m=2,r=0.8SD)	0.3165	0.5834
Apen (m=2,r=0.9SD)	0.2822	0.5260

To explore the complexity of the heart rate variability, the ApEn of the IHR signals was calculated. Figure 3 demonstrates the change of ApEn with m=2 and r=0.1*SD to 0.9*SD of IHR data for normal and AFIB subjects. The mean values of ApEn of the healthy group were found to be lower than that of AFIB at all r values except at 0.1*SD and 0.2*SD. Statistically, the ApEn of healthy group is found to be significantly different from that of AFIB group at r > 0.2*SD. The average AFIB values at different r are summarized in Table 4.

5 CONCLUSION

This work describes the application of approximate entropy (ApEn), time domain analysis to differentiate the normal rhythm from the AFIB. It should be noticed that approximate entropy (AE) statistics produce incompatible results. The most important thing is that AE counts each sequence as matching itself. We can attribute this self bias as the source of inconsistent result of AE. In contrast, sample entropy (SE) abates this self matching [5]. The SE statistics is free of the bias grounded by self-matching. The name refers to the applicability to time series data which is sampled from a continuous process. In addition, the algorithm advocates ways to make use of sample statistics to appraise the information existing in the data.

Time domain domain analysis of the RR interval variability of AFIB and normal subjects shows that there is significant difference in these measures for AFIB patients with respect to normal subjects. Lower values of ApEn reflect more regular time series while higher values are associated with less predictable (more complex) time series [6]. The lower ApEn values for the healthy group indicates an increase in regularity and a decrease in complexity in the IHR.

The major finding of this study shows that Time domain parameters and ApEn of HRV based on the complexity information of heart rate is able to better distinguish normal heart beat from the AFIB. A lower ApEn is found for normal rhythm data sets and higher in AFIB data sets. All assessed conventional HRV parameters in the time domain (Variance, SDNN, SDSD, RMSSD) were reduced in Healthy groups. The results show that there is a significant difference between the Time domain parameters of normal rhythm data sets and that of AFIB data sets. With respect to the found results mentioned above, nonlinear measures can be investigated further in future studies of heart rate variability in different cardiac diseases. The successful treatment of these measures may result in a better understanding of the nature of HRV.

REFERENCES

[1] Thakor, N.V.; Pan, K , "Tachycardia and fibrillation detection by automatic implantable cardioverter-defibrillators: sequential test-

ing in time domain," IEEE Trans. Biomed. Engg. Vol. 09, pp. 21-24, March.1990.

[2] Richman JS, Moorman JR, "Physiological time-series analysis using approximate entropy and sample entropy" Am J Physiol 2000, Vol. 278: H2039–H2049.

[3] Thakor, N.V.; Pan, K , "Tachycardia and fibrillation detection by automatic implantable cardioverter-defibrillators: sequential testing in time domain," IEEE Trans. Biomed. Engg. Vol. 09, pp. 21-24, March.1990.

[4] Volkan Tuzcu, Selman Nas "Sample Entropy Analysis of Heart Rhythm Following Cardiac Transplantation" 2005 IEEE International Conference on Systems, Man and Cybernetics Waikoloa, Hawaii October 10-12, 2005.

[5] Ahsan H. Khandoker, Herbert F. Jelinek, Marimuthu Palaniswami, "Heart rate variability and complexity in people with diabetes associated cardiac autonomic neuropathy" 30th Annual International IEEE EMBS Conference Vancouver, British Columbia, Canada, August 20-24, 2008.

[6] P Castiglioni, M Di Rienzo, "How the Threshold "R" Influences Approximate Entropy Analysis of Heart-Rate Variability" Computers in Cardiology 2008;Vol. 35: pp. 561–564.

[7] B. Anuradha, K. Suresh Kumar and V. C. Veera Reddy " Classification of Cardiac Signals Uing Time Domain Methods" ARPN Journal of Engineering and Applied Sciences, VOL. 3, NO. 3, JUNE 2008.

[8] D.A. Litvack, T.F. Oberlander, L.H. Carney, and J.P. Saul. "Time and frequency domain methods for heart rate variability analysis" a methodological comparison. Psychophys-Vol. 32,492–504, 1995.

[9] Steven M. Pincus "Approximate entropy as a measure of system complexity" Proc. Nati. Acad. Sci. USA, Vol. 88, pp. 2297-2301, March 1991.

[10] R Alcaraz, JJ Rieta, "Robust Prediction of Atrial Fibrillation Termination Using Wavelet Bidomain Entropy Analysis" Computers in Cardiology 2007, Vol. 34, pp. 577–580.

[11] MIT-BIH Arrhythmia Database CD-ROM, 3rd ed. Cambridge, MA: Harvard–MIT Div. Health Sci. Technol., 1997.

A.H.M Zadidul Karim has been serving as an Assistant professor in the Department of Electrical and Electronic Engineering (EEE), University of Asia Pacific (UAP). He joined at UAP in April 2007 after completion his BSc. Engg. (EEE). He has completed his MEngg. from BUET. Right now he teaches courses on Digital Electronics, Digital signal processing (DSP), Electrical Machines, Electrical and Electronics Circuits. Besides these, he is a convener of EEE Sports Club.

MD. MEGANUR RHAMAN has been serving as an Assistant professor in the Department of Electrical and Electronic Engineering (EEE) Department, Ahsanullah University of Science and Technology (AUST). He joined AUST in Octobor 2005 after the completion of his BSc. Engg. (EEE). He has completed his MEngg. from BUET. Right now he teaches courses on Measurement antd instrumentation, Digital Electronics, Digital signal processing (DSP), Electrical Machines, Electrical and Electronics Circuits.