

# ST Variability Analysis using Triangular Method, Linear Regression and SVM

S. Thulasi Prasad, Dr. S. Varadarajan

**Abstract**—The Cardio vascular diseases (CVDs) like Arrhythmia, Myocardial Ischemia and Myocardial Infarction (MI) may lead to sudden cardiac death if they are not identified in advance. With automated detecting system it is easy and faster for the analysts and doctors to diagnose these diseases from ECG rather than manually. In this paper an efficient and novel method, Triangular method, is proposed to extract ST segments and developed a Linear Regression model to detect ischemic beats for the analysis of ST-Segment Variability (STV). In the proposed method the ECG signal is preprocessed to remove powerline interference, motion artifacts and baseline wander. Later with simple QRS detection algorithm the QRS complex of each beat is detected. Next the RR intervals and the corresponding ST segments are extracted based on Triangular Method (TM) to form feature sets. From these feature sets a linear regression model is designed using Instantaneous heart rate (IHR) and ST segment level to set a reference threshold. Using the threshold set by the regression model the SVM classifier is used to identify the ischemic beats from the test feature sets of ECG signals. The Ischemic Intensity Factor and Ischemic Activity Factors were computed to evaluate the performance of the proposed method and found to be yielding better results when compared to Wavelet Transform based method.

**Index Terms**— IHR, Ischemia, Linear Regrsson, QRS Complex, RR interval, SA node, Scatter-plot, ST segments, SVM, Triangular method.

## 1 INTRODUCTION

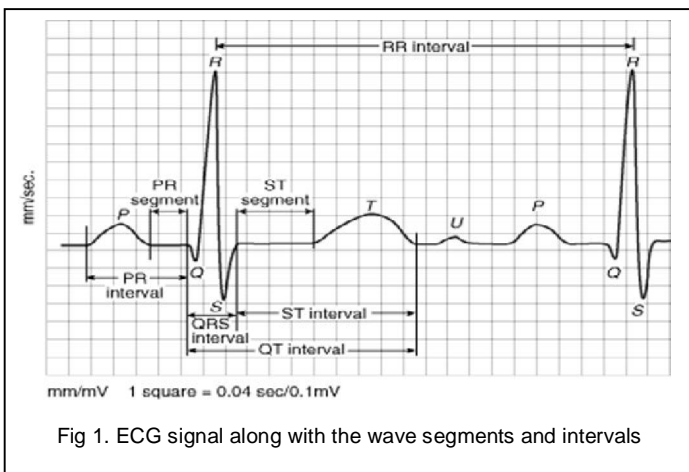
The ECG represents the electrical activity of human heart during the cardiac cycle [1]. Fig. 1 shows standard cardiac cycles of a normal ECG signal with the characteristics of different wave peak, wave segments and intervals. This consists of different positive and negative deflections corresponding to the electrical activity of the atrium and ventricles due to depolarization and repolarization. Such deflections are identified by tagging the letters P, Q, R, S and T repeatedly in every cardiac cycle. The Table 1 shows the various features and their characteristics of the standard ECG.

The electrical activity of human heart involves the generation of the electrical impulses from SA node and its transmission towards bundle branches via AV node [2]. These events are represented as a sequence of the P wave, PR interval, QRS complex, ST segment, T wave and QT interval as shown in the Fig. 1. The ST segment is the wave segment of an ECG that lies between the J-point and K-point [3]. This is usually not flat instead elevated or depressed.

The J-point is located as the first influx point after S-point and is identified as a point where the slope of the ECG wave is

TABLE 1  
 ECG CHARACTERISTIC FEATURES

Feature	Description
P Wave	A notched or biphasic wave with the duration of about 80 ms. This wave is generated due to sequential activation of the right and left atria
RR Interval	The normal duration of this wave is between 0.6 s to 1.2 s and represents the time required to transfer activation from atria to ventricles.
QRS Complex	It represents the depolarization of the ventricles and normally is around 0.06 to 0.10 sec (60 to 100 ms) in duration and larger than the P wave because the ventricles contain more muscle mass than the atria.
ST segment	A smooth waveform from the J-point (end of QRS) with gradual rising towards the peak of the T and terminates at K-point.



- S. Thulasi Prasad, currently working as Associate Professor in the Dept. of Electronics and Communication Engineering, Sree Vidyanikethan Engineering College, Tirupati, India, E-mail: stprasad123@yahoo.co.in
- Dr. S. Varadarajan, currently working as Professor in the Dept. of Electronics and Communication Engineering, Sri Venkateswara University College of Engineering, Tirupati, India, E-mail: varadasouri@gmail.com

flattest. Conventionally the K-point is located at a time interval of 40 ms from J-point in case of Tachycardia and at a time interval of 80 ms after J-point in case of bradycardia. Some times it is not possible to identify the exact beginning and end of ST segment [4], [5]. The automatic detection of cardiovascular diseases such as Myocardial Ischemia and Myocardial Infarction requires a simple and accurate detection of ST feature extraction from ECG signal. In this paper an efficient and a novel method, Triangular method, is proposed to extract ST segments and developed a Linear Regression model to detect ischemic beats for the ST-Segment Variability (STV) analysis [6]. With this method the ST segment abnormalities can be identified more efficiently and is possible to predict Myocardial Infarction.

## 2 BACKGROUND

In 1988, Aldrich et al introduced an ECG database foundation to quantify the size of the total risk in Acute anterior and inferior myocardial infarction (AMI) based on ST-segment deviation. Later many scientists, researchers and engineers worked on the automated ST variability and proposed several analytical methods to detect Myocardial Ischemia and Infarction. The following paragraphs focus on some literature reviews.

Badilini et al (1992) performed analysis on the ST segment in the frequency domain and created a system to detect ischemia. In their analysis it was observed that the ischemic beats are rich in lower frequencies than the normal ones. Senhadji et al (1995) investigated ECG signals for beat classification using advanced signal analysis procedures known as Wavelet Transforms (WT). As the Wavelets are both oscillatory and localized in time, these are used to scrutinize both time and spectral information of the signal concurrently. Taddei et al (1995) used geometric algorithm to estimate the ST segment variations using a rule-based system for two lead ECG recordings. A graphical rule is used to separate Ischemic episodes from normal beats after representing estimated ST segment deviations in time series for each cardiac beat.

In 2007, Elif Derya Ubeylia used multiclass SVM with the error correcting output codes (ECOC) to classify ECG beats into four types of ECG beats such as normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat and atrial fibrillation beat. Wai Kei Lei (2008) developed an intelligent heart rhythm recognition system which functions based on integration of Hermite based orthogonal polynomial decomposition (OPD) and support vector machines (SVMs) classification. Simoliuniene et al (2008) proposed a method for detection and evaluation of T-wave alternance in ECG. Data collected from 24 subjects of possible myocardial infarction was preprocessed which includes baseline wander removal and T-wave duration adjustment using modified Bazett's formula.

Gu-Young Jeong (2010) developed an algorithm to detect the ST level changes, and then to classify them according to the ST shape type using the polynomial approximation. Ranganathan G et al (2010) proposed an autoregressive (AR) frequency analysis of the heart-rate variability (HRV) signal.

Spectral decomposition of the Heart Rate Variability (IHR) requires ECG recordings acquired over an entire night to evaluate the variances in the heart rate. Chetan G et al (2014) proposed an adaptable PLI canceller utilizing Least Mean Square (LMS) algorithm to remove 50/60 Hz powerline interference from electrocardiogram (ECG).

## 3 METHODOLOGY

In this paper an efficient and accurate method was proposed to extract ST segments of the ECG signals and to detect Ischemia (MI) beats from ECG signals using ST segment regression analysis between ST levels and Instantaneous Heart Rate (IHR). The Fig. 2 shows the block diagram of the proposed method. The methodology is described in the following sub sections

### 3.1 Preprocessing

Normally the ECG signal is corrupted by noise and artifacts such as baseline wander, power line interference and motion artifacts. As result of this, the J point and K points cannot be accurately detected and the ST analysis cannot be performed accurately [7].

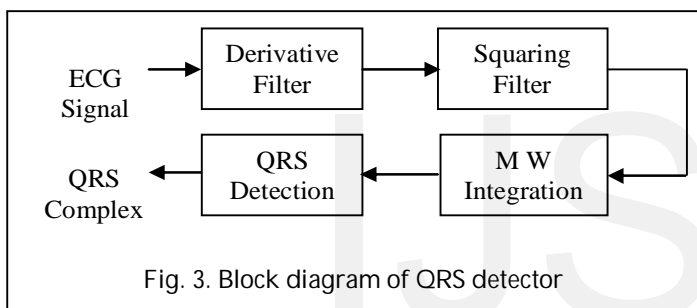
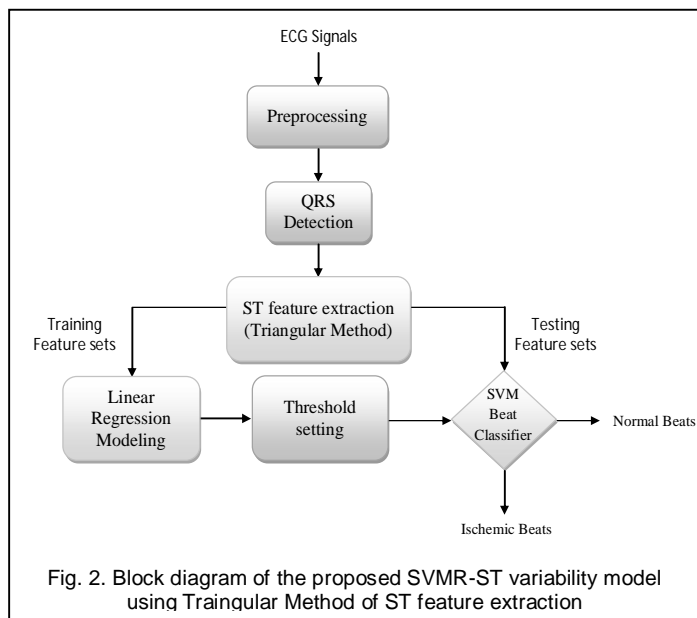
The preprocessing improves the signal to noise ratio (SNR) of ECG waves to such an extent that the ST segments can be measured more accurately. The preprocessing involves three steps: 1. High frequency noise motion and artifacts removal; 2. Baseline wanders removal and 3. Powerline interference removal [8].

The Butterworth lowpass digital filter having a cutoff frequency of 40 Hz is used to remove high-frequency noise from ECG signal [9]. The Butterworth highpass digital filter having a cutoff frequency of 0.05 Hz is used to remove the baseline wander from ECG signal [10]. A notch filter designed to 50 Hz is used to eliminate powerline interference.

### 3.2 QRS Detection

Detection of QRS complex is a vital step in the ECG signal analysis [11], [12]. A simple algorithm as illustrated in the Fig. 3 is used to detect QRS complex. In this algorithm first the R-peak is detected from the preprocessed ECG signal, later all other peaks are recognized with respect to the R-peak using QRS detection algorithm.

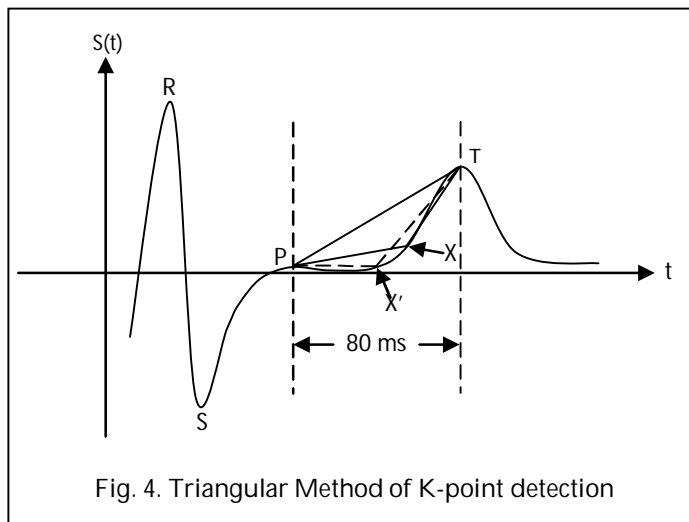
This algorithm involves the conversion of the filtered ECG signal into an ECG slope signal by the derivative filter to enhance the large magnitude samples compared to other sections samples. Later the squaring operation is applied to convert all the data values into positive values so that the higher frequency components of ECG signal are emphasized after the differentiation. Next the integrator is implemented with a moving window averaging filter and a threshold is set to the product of mean and maximum of the integrated signal output to detect the R-peak. Once the R-peak is identified, other peaks are recognized as the local negative maxima in each heart cycle on either side of R-peak. The negative maximum before R-peak is identified as Q-peak and the negative maximum after R-peak is identified S-peaks [13], [14].



### 3.3 Triangular method

The ST segment is the wave segment of ECG signal between the J-point and K-point in a time series of ECG wave [15], [16]. The J-point is located as the first inflex point after S-point and is identified as a point where the slope of the ECG wave is flattest. In this method the ECG signal is searched in the forward direction within a window of 20 ms to 100 ms from S-peak and J-point was recognized as one where the slope the ECG signal is either zero or changes its sign. Conventionally the K-point is located at a time interval of 40 ms from J-point in case of Tachycardia and at a time interval of 80 ms after J-point in case of bradycardia [17], [18]. But in the proposed method, the K-point is recognized with help of Triangular Method. The Triangular Method has been explained in the following sub section

ST segment is identified as the time separation between J-point and K-point. The J-point was located using slope of the ECG signal as explained in previous paragraph. But the K-point was identified using Triangular method instead of choosing the point located at 40 ms or 80 ms after J-point as in WT-ST modeling method. In the proposed method first a triangle is constructed with two fixed points as reference vertices and one variable vertex point at any point on the trace of the denoised ECG between the reference points of the triangle. One reference point is fixed at T-peak (vertex T), other refer-



ence point is fixed at 'P', that is 80 ms before T-peak and the variable point is 'X' as shown in the Fig. 4. The area of this triangle PTX is computed by moving the variable vertex point from one reference vertex (T-peak) to other reference vertex. Wherever the area of the triangle PTX gets the maximum value then time index corresponding to that point is considered as the K-point.

### 3.4 SVM Classification

Vapnik developed a new and powerful classification technique known as SVM for training and classification of learning system [19], [20]. This technique is extensively used for solving classification problems based on the various functions such as polynomial functions, radial basis functions, neural networks etc. SVM performs a non linear transforms of the original input space into a high dimensional feature space, creates Optimal separating hyper (OSH) plane and maximizes the margin between training data and the OSH. SVM formulates the OSH using quadratic optimization technique in a feature space. Then the support vectors are formed from the neighboring subset patterns of the OSH [21], [22].

If SVM N-dimensional input sample vector  $X_i \in R^d$  is transformed into K-dimensional feature space using  $\Phi(X)$  as shown in the Fig. 5. The equation of the hyperplane separating two different classes in the N-dimensional space is given by

$$C(X) = W \cdot \Phi(X) + b \quad (1)$$

$$\sum_{i=1}^K C(X) = W \cdot \Phi(X) + b \quad (2)$$

Where  $W \cdot X$  is the dot product of weighting vector 'W' and input sample space 'X', and 'b' is the bias or threshold. The decision function is evaluated based on the condition given by (3)

$$\text{sign} [C(x)] \geq 1 \quad (3)$$

The Non-linear transformation of original samples' space into high-dimension space is performed with help of proper

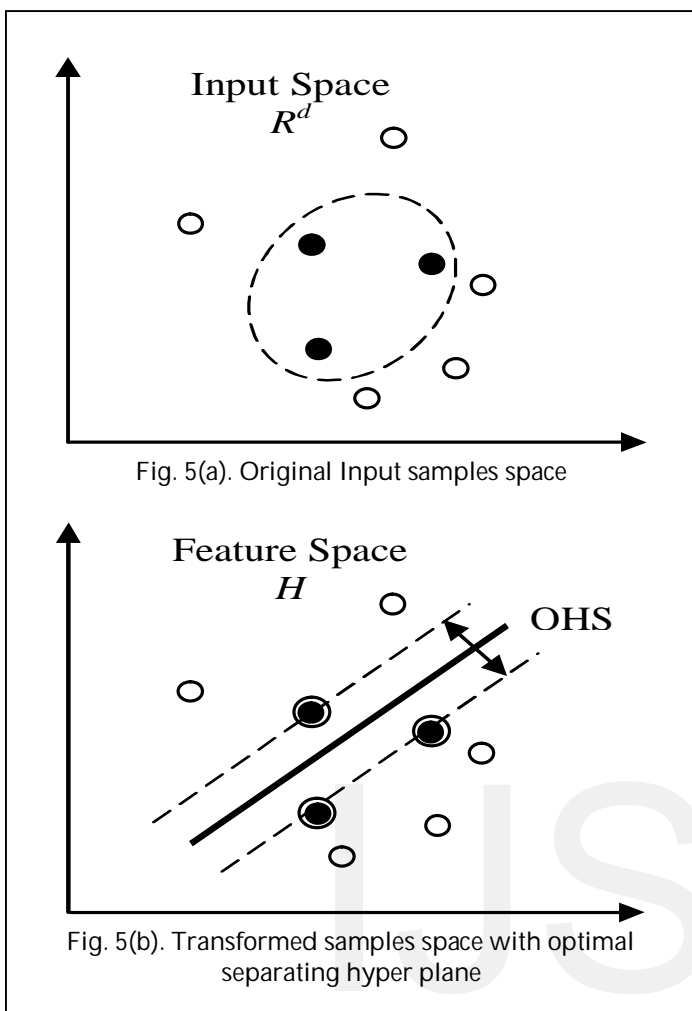


Fig. 5(a). Original Input samples space

Fig. 5(b). Transformed samples space with optimal separating hyper plane

kernel mapping and the best linear classification surface of the samples. The kernel function,  $K(X_i, X)$  is defined as the inner product of the vector as given by the equation

$$\Phi^T(X_i) \Phi(X) \tag{4}$$

Some of the best known kernels functions are listed below.

1. Linear Kernel function:  $K(X, Z) = X.Z$ , where  $X.Z$  is the dot product of  $X$  and  $Z$ .
2. Polynomial kernel function:  $K(X, Z) = (X.Z + 1)^d$ , where ' $d$ ' is the degree of polynomial.
3. Radial basis function:

$$K(X, Z) = \exp\left[\frac{-|X - Z|^2}{2\sigma^2}\right] \tag{5}$$

where ' $\sigma$ ' is width of the function

In SVM the task of separating vectors  $X_i$  into two classes is formulated by describing the destination values either  $d_i = 1$  or  $d_i = -1$ , as the maximal separation margin. The best linear classification surface function is obtained by maximizing the quadratic function  $Q(x)$  in terms of Lagrange multipliers ' $a_i$ ' and the destination values ' $d_i$ ' associated with the input vectors  $X_i$ , defined as;

$$\sum_{i=1}^n a_i - 0.5 \sum_{i=1}^n \sum_{j=1}^n a_i a_j d_i d_j K(X_i, X_j) \tag{6}$$

Subjected to the conditions

$$\sum_{i=1}^n a_i d_i = 0 \quad 0 \leq a_i \leq C \tag{7}$$

Where ' $C$ ' is the user-defined regularization constant and ' $n$ ' is the number of learning data pairs  $(X_i, d_i)$ . Finally the expression of the output signal  $C(X)$  is given as;

$$\sum_{i=1}^{N_{sv}} a_i d_i K(X_i, X) + b \tag{8}$$

And the best classification function of SVM given by;

$$\text{sign} \left\{ \sum_{i=1}^{N_{sv}} a_i d_i K(X_i, X) + b \right\} \tag{9}$$

To apply SVM on ECG signals for beat classification it necessary to know the proper kernel function. The kernel function that is used in this research work is Radial basis function which expressed as;

$$K(X, Z) = \exp\left[\frac{-|X - Z|^2}{2\sigma^2}\right] \tag{10}$$

Where ' $\sigma$ ' is width of the function.

The SVM classifier is trained with training ST feature set obtained after setting the threshold against the extracted ST features by the regression model. Once the SVM classifier is trained it is ready to classify the beats into normal and Ischemic.

#### 4 RESULTS

The proposed method was tested on the ECG datasets acquired from a local hospital. First the ECG of each patient is preprocessed with a bandpass filter to remove powerline interference, motion artifacts and baselinewanders. For instance the original and preprocessed ECG signals of the patient ID: 077 are shown in the Fig. 6 and Fig. 7 respectively for few cycles.

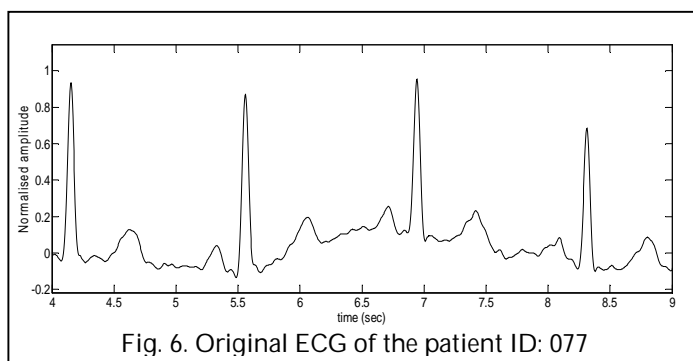
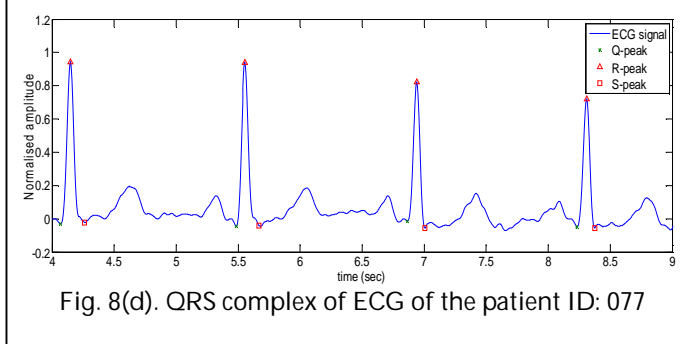
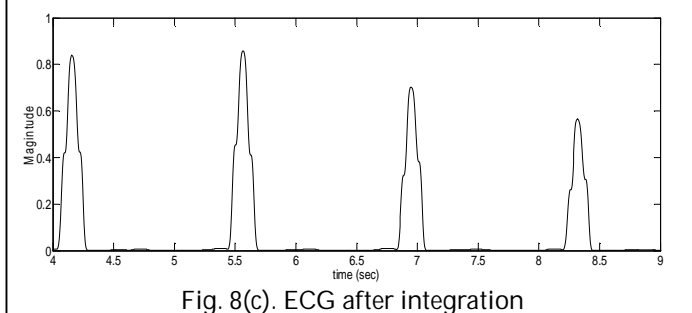
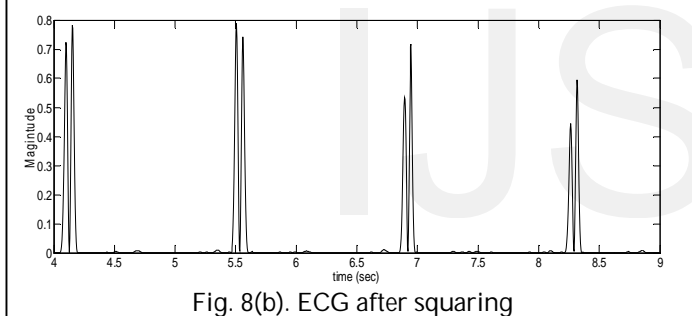
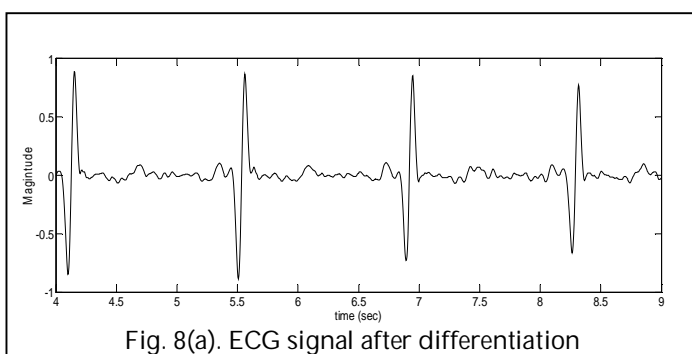
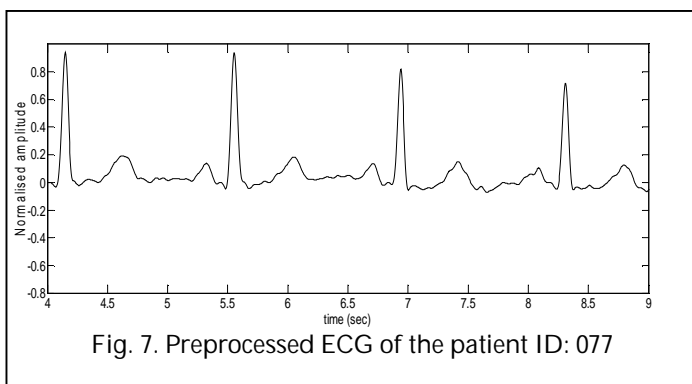


Fig. 6. Original ECG of the patient ID: 077



The QRS complex is detected by performing differentiation, squaring, integration sequentially on the filtered ECG signal and setting a threshold using (11).

$$\text{Threshold} = \text{mean}[X(n)] * \text{max}[X(n)] \quad (11)$$

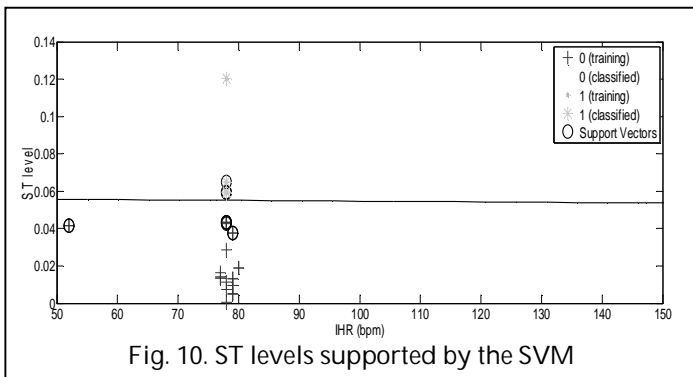
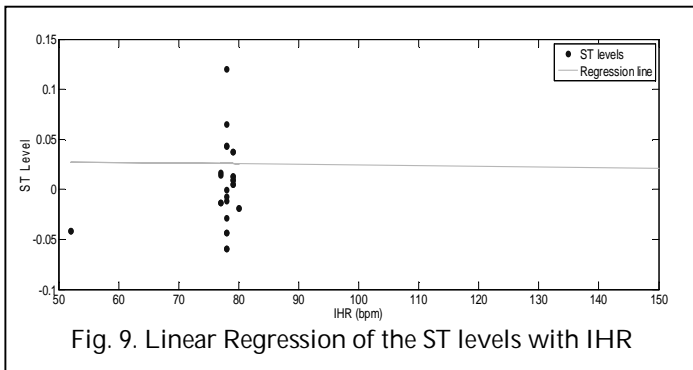
Fig. 8 shows the result of operations in QRS detection. Next the J-point is identified from the slope signal. From the filtered time series ECG the K-point is detected for each beat by using the Triangular Method and ST segments are extracted. Next the mean amplitude of the PQ-segment is determined from the wave segment between the offset of P-wave and onset of QRS wave and is considered as the isoelectric reference level. By using this isoelectric level as the reference level, the ST segment level of each beat is determined as the difference between the amplitude of the ST segment and isoelectric reference level. Instantaneous heart rate (IHR) from the RR interval using (12) is computed for each beat.

$$\text{Instantaneous heart rate (IHR)} = 60/RR \text{ bpm} \quad (12)$$

Both ST levels and Instantaneous Heart Rates (IHRs) are saved as feature sets for the ST analysis. For instance the IHR and ST levels of ECG signal of the patient ID: 077 for the first 30 beats are tabulated as in the Table 2.

TABLE 2  
IHRs AND ST LEVELS OF THE ECG OF THE PATIENT ID: 077

Beat No	IHR (bpm)	ST Level	Beat No	IHR (bpm)	ST Level
1	54	0.0417	16	52	0.0416
2	45	0.0410	17	78	0.1206
3	78	0.0649	18	78	0.0288
4	78	0.0113	19	78	0.0004
5	78	0.0074	20	77	0.0143
6	77	0.0165	21	78	0.0435
7	78	0.0428	22	79	0.0374
8	79	0.0377	23	78	0.0593
9	78	0.0594	24	78	0.0434
10	78	0.0433	25	77	0.0135
11	77	0.0135	26	79	0.0096
12	79	0.0096	27	79	0.0133
13	79	0.0133	28	80	0.0189
14	80	0.0188	29	79	0.0051
15	79	0.0050	30	52	0.0415



The Linear Regression algorithm is employed between IHR and ST levels from the Table 2 to derive design coefficients. The regression modeling can be visualized by constructing a regression curve from the regression model coefficients as shown in the Fig. 9. This regression curve is used as a threshold to compare the ST levels of the testing feature sets. All the beats whose ST levels are above the regression threshold are identified as probable Ischemic beats. Then SVM is employed on the detected probable Ischemic beats to separate the final and true Ischemic beats. Fig. 10 shows the beats supported by the SVM. Thus the Ischemic beats are populated for further ST variability analysis. The ST variability is analysed with the parameters like Ischemic Intensity Factor, Ischemic Activity Factor (per 30 beats) and Peak to Average Value (PAV) by using (13), (14) and (15).

$$Ischemic Intensity Factor (IIF) = \frac{N_{Isc}}{N_T} \quad (13)$$

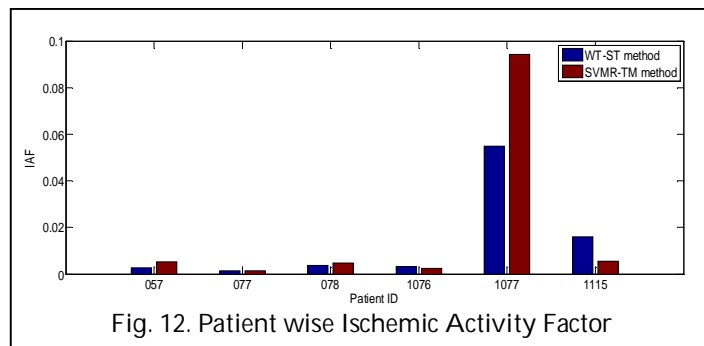
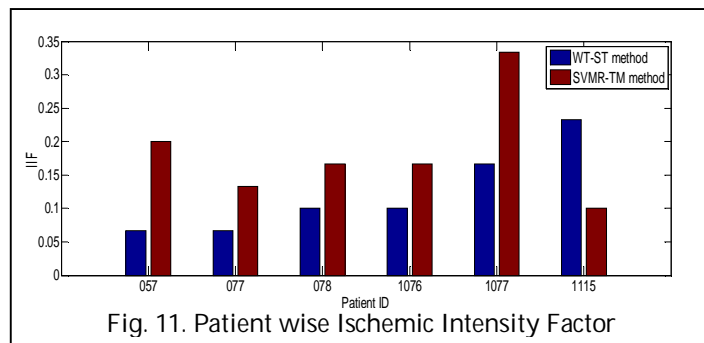
$$Ischemic Activity Factor (IAF) = N_{Isc} * ST_{avg} \quad (14)$$

$$Peak to Average Value (PAV) = \frac{ST_{peak}}{ST_{avg}} \quad (15)$$

Where  $N_{Isc}$  is the number of detected Ischemic beats,  $N_T$  is the total number of beats,  $ST_{avg}$  is the average ST level and  $ST_{peak}$  is the peak value of the respective ECG signal ST segments. For example the IIF, IAF and PAV of patient ID: 077 are evaluated as 0.133, 0.002 and 5.14 respectively for SVMR-TM method. The same procedure is repeated on ECGs of six patients using WT-ST method and SVMR-TM methods. From the count of Ischemic beats detected in both methods, all other performance evaluation parameters are computed and tabulated as in the Table 3.

**TABLE 3**  
**PERFORMANCE EVALUATION AND COMPARISON**

Patient ID	057	077	078	1076	1077	1115	
WT-ST method	No of Ischemic beats	2	2	3	3	5	7
	Ischemic Intensity Factor (IIF)	0.067	0.067	0.100	0.100	0.167	0.233
	Ischemic Activity Factor (IAF)	0.003	0.001	0.004	0.003	0.055	0.016
	Peak to Average Value (PAV)	5.18	9.24	5.46	6.40	2.569	6.47
SVMR - TM method	No of Ischemic beats	6	4	5	5	10	3
	Ischemic Intensity Factor (IIF)	0.200	0.133	0.167	0.167	0.333	0.100
	Ischemic Activity Factor (IAF)	0.005	0.002	0.005	0.003	0.094	0.006
	Peak to Average Value (PAV)	2.21	5.14	4.33	7.91	2.99	7.53



Patient wise IIF and IAF are shown in the Fig. 11 and Fig. 12 respectively for all six patients according the Table 3 data. From the results in the Table 3 as well as from the Fig. 11 and Fig. 12 it is clear that the SVMR (Support Vector Machines Regression) - ST variability modeling with Triangular Method of ST feature extraction is yielding higher Ischemic Intensity Factor (IIF), Ischemic Activity Factor (IAF) and Peak to Average Values than the Wavelet Transform (WT-ST) based method. The results show that the performance of SVMR-TM method is better than WT-ST method.

## 5 CONCLUSION

In this paper a new method of feature extraction based on Triangular Method and SVM linear regression is proposed. Experimental results show that the feature selection using regression method greatly improves the classification quality. This is because some ST features may be missed when the hard threshold is used as in conventional methods. In the proposed method the effect of raising slope of the T-wave is also included in feature selection with the help of the Triangular Method of feature extraction. This also improves the accuracy of Ischemic detection and the classification quality. The proposed method was tested on the ECG signals, which include the normal beats and ischemia beats, obtained from a local hospital. The Ischemic Intensity Factor (IIF) and Ischemic Activity Factor (IAF) are evaluated for both the proposed method and WT-ST method for all the six patients. The proposed method is yielding better results compared to the WT-ST method for all the patients. Thus the Support Vector Machines linear Regression modeling using Triangular Method of ST feature extraction (SVMR-TM) will be very useful and helpful for the doctors and ECG analysts to recognize Ischemia and analyze the ECG signals more accurately.

## REFERENCES

- [1] S. Chugh, "Textbook of Clinical Electrocardiography", Tunbridge Wells, UK, Anshan Ltd., 2006
- [2] T.-C. Chou, "Electrocardiography in Clinical Practice," *The Journal of the American Medical Association*, vol. 256, no. 23, pp. 3291-3292, 1986.
- [3] Stamkopoulos, T., Maglaveras, N., Bamidis, P. D., and Pappas, C., "Wave segmentation using nonstationary properties of ECG", *In Computers in Cardiology*, 529-532. IEEE Press, 2000.
- [4] G.-Y. Jeong, K.-H Yu, "Design of Ambulatory ECG Monitoring System to detect ST pattern change," *SICE-ICASE International Joint Conference*, Oct 2006.
- [5] Ruch TC, Patton HD (eds.) (1982), "Physiology and Biophysics", 20th ed., 1242 pp. W. B. Saunders, Philadelphia.
- [6] Pettersson, J., Pahlm, O., and Carro, E., "Changes in High-Frequency QRS Components Are More Sensitive Than ST-Segment Deviation for Detecting Acute Coronary Artery Occlusion", *Journal of the American College of Cardiology*, Vol. 36, pp. 1827-1834, 2000.
- [7] Thakor, N.V., Yi-Sheng, Z. "Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection", *IEEE Trans. Biomed. Eng.* 1991; 38: 785-794.
- [8] P. S. Hamilton, "A comparison of adaptive and nonadaptive filters for the reduction of powerline interference in the ECG", *IEEE Trans. Biomed. Eng.* 1996; 43:105-109.
- [9] Dr. Christian Cuadrado, "A DSP Practical Application: Working on ECG Signal, Applications of Digital Signal Processing", ISBN: 978-953-307-406-1
- [10] Belle A. Sheno, "Introduction to digital signal processing and filter design", John Wiley and Sons, p.120.ISBN 978-0-471-46482-2 :2006.
- [11] Friesen GM, Jannette TC, Jadallah M A, Yates SL, Quint SR, Nagle HT, "A comparison of the noise sensitivity of nine QRS detection algorithms", *IEEE Trans Biomed Eng*, 37: 85-98, 1990.
- [12] Schlegel, T.T., Kulecz, W.B., Depalma, J.L., Feiveson, A.H., Wilson, J.S., Rahman, M.A., and Bungo, M.W., "Real-Time 12-Lead High-Frequency QRS Electrocardiography for Enhanced Detection of Myocardial Ischemia and Coronary Artery Disease", *Mayo Clinic Proceedings*, Vol. 79, pp. 339-350, 2004
- [13] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm", *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 230-236, March 1985.
- [14] I. I. Christov, "Real time electrocardiogram QRS detection using combined adaptive threshold", *BioMedical Engineering OnLine*, vol. 3, no. 1, p. 28, 2004.
- [15] di Bernardo, D., Langley, P., and Murray, A., "Effect of Changes in Heart Rate and in Action Potential Duration on the Electrocardiogram T Wave Shape", *Physiological Measurement*, Vol. 23, No. 2, pp. 355-364, 2002.
- [16] Y. Suzuki, K. Ono, "Personal computer system for ECG ST-segment recognition based on neural networks," *Medical and Biological Engineering and Computing*, vol 30, pp. 2-8, Jan 1992.
- [17] X. Li, et al., "ST-T complex automatic analysis of the electrocardiogram signal based on wavelet transform," *Proc. 29th IEEE Bioengineering*, pp. 144-145, 2003.
- [18] P. de Chazal, M. O'Dwyer and R. B. Reilly, "Automatic Classification of Heart beats using ECG Morphology and Heart beat Interval Features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196-1206, July 2004.
- [19] Vapnik, V., "The Nature of Statistical Learning Theory", New York, Springer-Verlag, pp.187-189. 1995.
- [20] S. R. Gunn, "Support Vector Machines for Classification and Regression," Technical Report, Dept. of Electronics and Computer Science, University of Southampton, 1998.
- [21] T. Joachims, "Making large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning," MIT-Press, B. Schölkopf, C. Burges and A. Smola (ed.), 1999.
- [22] Takeuchi, K., Collier, N., "Bio-Medical Entity Extraction using Support Vector Machines", *Proceedings of the ACL 2003 Workshop on Natural Language Processing in Biomedicine*, 57-64, 2003.