

Pattern Classification of ECG signals for Cardiac Disease Prediction using Probabilistic Models

S. Elango^{a, *}, J. Sundararajan^b

^a Research Scholar, Anna University of Technology, Coimbatore, TN, India

^b Principal, Pawai College of Technology, Namakkal, TN, India

Article Info

Article history:

Received 12 July 2013

Accepted 14 July 2013

Available online 23 July 2013

Keywords:

Wavelets transform

Probabilistic Model

ECG

Pattern mining

Signal Processing

Abstract

We propose a new methodology for the prediction of cardiac disease with the help of wavelet transform and Hidden markov model. In most of the occasions the variation in the human heart rate is unidentifiable due to the small voltage variation could not recordable with graph generated by ECG. The small difference in the ECG may also notify the presence of cardiac problem which could not be identified with human intervenes. The Electro cardiogram signals are passed through wavelet filter for the identification of small variation and then classified with probabilistic models. The wavelet transform boost the signals so that the small variation also become accountable and then we used hidden markov model for the identification of pattern which is relevant to identify cardiac diseases.

© 2013 TUJEST. All rights reserved.

1. Introduction

The Electro Cardiogram represent how the human heart is responding for the electrical signal and it shows the activity of heart that is what we call heart beat. The ECG records the electrical activity of the heart, where each heart beat is displayed as a series of electrical waves characterized by peaks and valleys. Any ECG gives two kinds of information. One, the duration of the electrical wave crossing the heart which in turn decides whether the electrical activity is normal or slow or irregular and the second is the amount of electrical activity passing through the heart muscle which enables to find whether the parts of the heart are too large or overworked.

Wavelets transform a signal processing technique used in various applications to decompose, filter, feature extraction etc.... Wavelet transform has huge impact in biomedical systems for signal processing. For many signals, the low-frequency content is the most important part. It is what gives the signal its identity. The high-

frequency content, on the other hand, imparts flavor or nuance. To gain a better appreciation of this process, it is performed a one-stage discrete wavelet transform of a signal. The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components.

In wavelet analysis, a signal is split into an approximation and a detail. The approximation is then itself split into a second-level approximation and detail, and the process is repeated. The transformed signal provides information about the time and the frequency. Using this approximated information low frequency data could be identified, which is more important in cardiac disease prediction.

Pattern mining is a process of similar valued data points from large set of data sets. In digital signal processing the same can be applied to identify and select similar frequency signals with time and frequency. Hidden Markov model is a statistical Markov model in

*Corresponding Author:

S. Elango, E-mail: nandhu_prema@yahoo.co.in

which the system being modeled is assumed to be a Markov process with unobserved states. A HMM can be considered the simplest dynamic Bayesian network. In normal markov model the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; even if the model parameters are known exactly, the model is still 'hidden'.

Here we discuss a new probabilistic model, it uses the data which are sequentially generated according to time or index. For example from the following example we classify the human heart into three states.

State1: Diseased
State2: Symptoms
State3: Normal.

Table1. The pattern of heart rate and transition states.

	Diseased	Symptom	Normal
Heart Rate	190	170	95
	195	172	89
	189	165	112

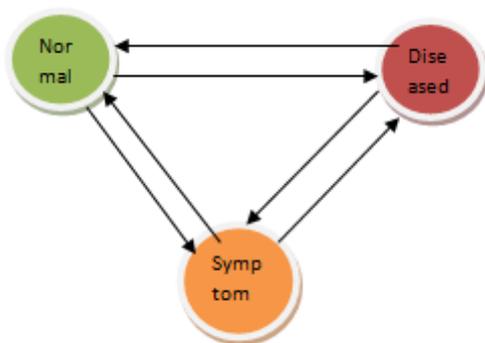


Figure1: How the state changes.

The problem is identifying the state and transition in cardiac diseases. The state of heart will be varying with slight change in heart rate which is a hidden one and will not be identified with normal methods and change of state to another state is a probabilistic one. So that we use hidden markov model which is a probabilistic one and used for pattern classification.

2. Background

There are various researches ongoing for the problem of ECG signal analysis. We discuss few of them here related to the problem.

Classification of cardiac signal using time domain method [1], is discussed which uses statistical parameters. The four statistical parameters considered for cardiac arrhythmia classification of the ECG signals are the standard deviation of the NN intervals (SDNN), the standard deviation of differences between adjacent NN intervals (SDSD), the root mean square successive difference of intervals which are extracted from heart rate signals (RMSSD) and the proportion derived by dividing NN50 by the total number of NN intervals (pNN50).

Recently wavelets have been used in a large number of biomedical applications. The wavelet packet method is a generalization of wavelet decomposition that offers a rich range of possibilities for signal analysis. The multi-resolution framework makes wavelets into a very powerful compression [5] and filter tool [6], and the time and frequency localization of wavelets makes it into a powerful tool for feature extraction [7].

There is some works on precise detection of ECG using FFT and wavelet, Karel et al. proposed the performance criteria to measure the quality of a wavelet, based on the principle of maximization of variance [8].

In [9] Mahmoodabadi developed and evaluated an electrocardiogram (ECG) feature extraction system based on the multi-resolution wavelet transform. David at [10] presented a method to reduce the baseline wandering of an electrocardiogram signal. Shantha at [11] discussed the design of good wavelet for cardiac signal from the perspective of orthogonal filter banks.

Detection of ECG characteristic points using wavelets transforms [2] shows that with multi scale information in wavelets it is easy to characterize the ECG waves and the QRS complex. In [2] the difference from high P and T waves, noise, baseline drift and interference were recognized.

Comparing Wavelet Transforms for Recognizing Cardiac Patterns [3], proposed to recognize wavelet pattern by comparing cardiac patterns. The choice of the wavelet family as well as the selection of the analyzing function into these families have been discussed to the Daubechies decompositions provided by the spline wavelet (6 levels) and the complex wavelet (10 levels) .

In [4], Amara Graps proposed a method that though D6 algorithm is more complex and has a slightly higher computational overhead but it picks up detail that is missed by the Harr wavelet algorithm, which is simpler than the former. D6 of Debauchees is similar in shape to QRS complex and their energy spectrum is concentrated around low frequencies.

All the methods discussed have the problem of identifying hidden layer ie the hidden signals(low signal values) are not identified and no pattern matching is performed for the detection of cardiac disease. We

propose a new vector quantization method to predict the cardiac disease using hidden markov models.

3. Proposed System

The propose method uses hidden markov model for the quantization of vectors in order to identify the possibility of cardiac disease. The echo cardiogram contains various portions like P, Q, R signals. Each having some space and time in the graph, we use these time and spatial information's for identification purpose.

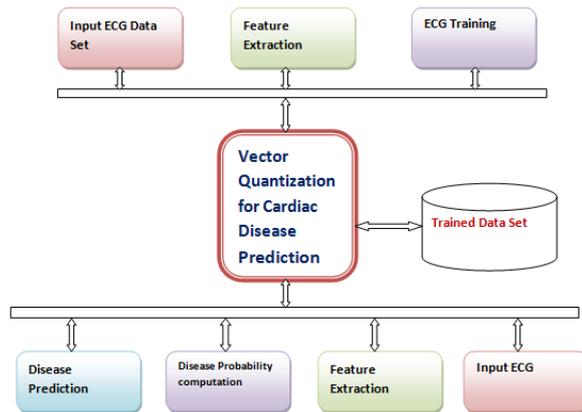


Fig 2. Proposed System Architecture.

3.1. Feature Extraction

The electrocardiogram contains various time domain and space domain values; they are amplitudes and intervals of various sectors. We extract P-R interval, R-R interval, Q-T interval, S-T interval, P-wave interval, QRS interval and the amplitude values of P, R, Q, T waves. Each features extracted is stored in the data base for further manipulation.

Algorithm:

Step1: read data set D_s .

Step2: for each ecg E_i from D_s

Extract the following features.

P_r = P-R interval

R_r = R-R interval.

Q_t = Q-T interval.

S_t = S-T interval

P = P wave interval.

Q_{rs} = QRS interval.

PA - P wave Amplitude.

QA - Q wave Amplitude.

RA - R wave Amplitude.

TA - T wave Amplitude.

Construct vector V_i and add to vector

set V_s .

$$V_s = \sum V_i(P_r, R_r, Q_t, S_t, P, Q_{rs}, PA, QA, RA, TA).$$

End.

Step3: stop.

3.2. ECG Training

At this stage the extracted features are used for training and we assign class labels for each feature vector v_i from the feature set V_s . Initially each class is assigned with set of vectors and with the input vector we compute the probability of vector to be assigned with each class label. Which class label has more probability will be assigned to the vector. For a single feature vector there are ten features and for each feature in the feature vector, a cumulative probability is calculated and based on the presence and absence of waves and the values of the various features. The computed probability value will be used for indexing.

Algorithm:

Step1: read feature vectors V_s

Step2: for each vector V_i in V_s

Identify

If($P_{w_i} \neq \text{Null}$) // presence of inverted P-wave-

P_{w_i}

Assign Dextrocardia as label to E_i .

If($R-R$ interval < 0.6 s)

Assign Tachycardia as label to E_i .

If($R-R$ interval > 1 s)

Assign Bradycardia as label to E_i .

If (Tall T-wave && P-wave = Null)

Assign Hyperkalcaemia as label to E_i .

If (Complete drop out of cardiac cycle)

Sinoatrial block as label to E_i .

If (Inverted T-wave)

Assign Myocardialischaemia as a label

to E_i .

If (QRS interval $< 0.1s$)

Assign Hypercalcaemia as label to E_i

If (Irregular)

Sudden cardiac death

End if

Step3: stop.

3.3. Disease Probability computation

Here we compute the disease probability like state transition models. There are three states in our probabilistic model like normal, diseased, symptoms. There are eight cases in disease, we compute probability for each case of the diseases and if the input vector doesn't match with any of the case and have some abnormal values then it will be assigned to the state of symptom. The probability value must be more than 0.6 to be get assigned to a case of disease. If the probability is less than 0.5 and greater than 0.1 then we can say that there are symptoms.

Algorithm:

Step1: read input feature vector I_v .
 Step2: initialize probability matrix P_b .
 Step2: for each case of disease D_i
 For each feature vector $V_i(D_i)$
 Cumulative distance $Cd =$
 $cd + \sum \text{Dist}(I_v(a_{i-n}) - V_i(a_{i-n}))$.
 End
 $Pb_i = Cd \times \log N$.
 N-total number of feature vectors in
 particular case.
 End
 Step 3: End.

3.4. Disease Prediction

This procedure predicts the disease affected by the person based on computed probability value and conditions for the diseases and so on. If the input vector have good probability towards a disease then the disease will be easily identified, otherwise the symptoms for each disease will be analyzed and prompted.

Step1: start
 Step2: read probability matrix pb .
 Step3: sort P_b according to probability value.
 Step4: select top few disease probability and check for the presence of attribute
 Condition. (R-R interval > 1s).
 If(condition == true)
 Assign disease label to the feature
 vector.
 Else
 Assign symptom as state.
 End.
 Step5: End.

4. Results and Discussion

We have used 2000 Electrocardiogram records for the evaluation of our algorithm. From 2000 records, 60 percent of the records are used for training and 40 percent has used as testing samples. The proposed method have produced good results and have taken little time for the training and testing process. The quality of disease prediction is higher and produces good results compare to other works in this era.

5. Conclusion

We have proposed a vector quantization of ECG signals for the prediction of cardiac diseases using a probabilistic model. The proposed system produces good results compare to other existing works. The overall time complexity has been reduced with the help of this model.

References

- [1] B. Anuradha, K. Suresh Kumar and V. C. Veera Reddy , classification of cardiac signals using time domain methods , arpn journal of engineering and applied sciences, vol. 3, no. 3, 2008.
- [2] Cuiwei Li, Chongxun Zheng and Changfeng Tai. 1995. Detection of ECG characteristic points using wavelets transforms. IEEE Transactions on Biomedical Engineering. 42(1): 21-28.
- [3] Senhadi, L., Carrault, G., Bellanger, J.J., and Passariello, G. 1995. Comparing Wavelet Transforms for Recognizing Cardiac Patterns. IEEE Engineering in Medicine and Biology Magazine, Publication. 14(2): 167-173 .
- [4] Amara Graps. 1995. An Introduction to Wavelets. IEEE Comp.Sc. and Eng. 2(2).
- [5] Ivo Provazník and Jiří Kozumplík. 1997. Wavelet transform in electrocardiography data compression. International Journal of Medical Informatics. 45(1-2):111-128, June.
- [6] M.P. Wachowiak, G.S. Rash, P.M. Quesada, and A.H. Desoky. 2000. Wavelet based noise removal for biomechanical signals: a comparative study. IEEE Trans. Biomed. Eng. 47(3):360-368, March.
- [7] Cuiwei Li, Chongxun Zheng, and Changfeng Tai. 1995. Detection of ECG characteristic points using wavelet transforms. IEEE Trans. Biomed. Eng. 42(1): 21-28, January.
- [8] Karel J.M.H., Peeters R.L.M., Westra R.L., Moermans K.M.S., Haddad S.A.P. and Serdijn W.A. 2005. Optimal discrete wavelet design for cardiac signal processing, Proceedings of the IEEE, Engineering in Medicine and Biology 27th Annual Conference.
- [9] Mahamoodabadi S.Z., Ahmedian A., Abolhasani M.D. 2005. ECG feature extraction using daubechies wavelet, Proceedings of the fifth IASTED International Conference VISUALIZATION, IMAGING and IMAGE PROCESSING, September 7-9, Benidorm, Spain.
- [10] David Cuesta-Frau, Daniel Novak, Vladimir Eck, Juan C. Pérez-Cortés. 2002. Electrocardiogram Baseline Removal Using Wavelet Approximations, Computer-Based Medical Systems, (CBMS 2002), Proceedings of the 15th IEEE Symposium.
- [11] R. Shanthe Selva Kumari, S. Bharathi, V. Sadasivam. 2007. Design of Optimal Discrete Wavelet for ECG Signal using orthogonal filter bank, International Conference on Computational Intelligence and Multimedia Applications, IEEE.