A Two-stage Solution Algorithm for Paroxysmal Atrial Fibrillation Prediction

KS Lynn, HD Chiang

School of ECE, Cornell University, Ithaca, USA

Abstract

We propose a two-stage solution algorithm to predict the onset of paroxysmal atrial fibrillation (PAF) based on half-hour heart rate variability (HRV) signals. Nonlinear feature based on vectors calculated from return map and difference map constructed by HRV signal were developed. The extracted features were fed into their corresponding k-nearest-neighbor classifiers for parameter adjustment and classification.

According to the official scoring results, our algorithm scored 34 points in the screening stage and 40 points in the prediction stage. In addition, the developed algorithm appears to be very robust against measuring noises. For example, with different QRS detectors, the classification results only change slightly (within 5%).

1. Introduction

Human's heart-rate variation (HRV) signals have been shown to exhibit fractal and chaotic properties [1-3]. There are important evidences of significant nonlinear dynamics in HRV signals. Many papers published during the past few years indicated the capability of nonlinear tools in distinguishing several heart-diseased signals that were considered non-separable from normal heart signals using linear tools [4-6].

There is no closed-form mathematical expression for the dynamics of human's heart. In addition, due to measuring noise and chaotic behavior in the HRV signal, long-term prediction of HRV signal is almost impossible. However, there are still ways to perform PAF prediction without the knowledge of exact future HRV signal. In this paper, we develop a two-stage solution algorithm to accomplish this task.

There are three major steps in the proposed algorithm: data preprocessing, feature extraction, and feature classification. At the step of data preprocessing, the original input data is examined and modified into an appropriate format. This includes data transformation, noise removal, bad data ejection, etc. However, since officially issued data was used, it is assumed that all the data are in proper format with reasonable noise and therefore, only transformations of R-R intervals to HRV were performed at this step. At the feature extraction step, useful information that best represent characteristics of data is extracted and rearranged into proper format. The extracted information, or features, will then be fed into a classifier for classification during the feature classification step. In this paper, we develop a two-stage solution algorithm for predicting the onset of PAF. The objective of the first stage, which is a screening stage, is to distinguish HRV signals of group N (subject with no documented PAF) from those of group A (subject with PAF but no PAF episodes in the record). The goal of the second stage, a prediction stage, is to further separate group-A signals into records that are far away (at least 45 minutes) from PAF episodes and records that are immediately followed by PAF episodes. We will focus on feature extraction and classification. We will be talking about how features are selected and extracted for both screening stage and prediction stage in section 2. In section 3, we will explain the architecture of the system. Experimental results will be provided in section 4 while discussions about our proposed algorithm will be covered in section 5. A final concluding remark of our algorithm will be presented in section 6.

2. Feature extraction

Features, according to the way they are extracted, can be classified into three categories: time domain, frequency domain, and state-space domain. Features extracted directly from sequential measurements are called time domain features. Features extracted from signal that is processed by Fourier transform or other similar tools are called frequency domain features. In addition, Features extracted from system's trajectory are called state-space domain features. Recently, state-space domain features have been widely discussed due to their outstanding capability of capturing features from the reconstructed trajectories that are proven topologically equivalent to the trajectories of the original system. The trajectory reconstruction technique was developed based on Takens' embedding theorem [7] in which he proved that the trajectory of a dynamical system can be topologically reconstructed from one of the system's outputs as long as the system is on an attractor.

The simplest feature, also called an index, consists merely a single number. Researchers have been trying to apply existing tools or develop new tools to compute indices from a series of measurements. These indices attempted to represent certain properties of the data [8-10]. When applied to HRV signals, some of them were able to reflect the health condition of human's heart (healthy or heart-diseased) [3-6, 11]. However, up to now, there are no reliable indices that can be used to distinguish HRV signals of PAF patients and those of non-PAF subjects.

In our work, we abandoned the idea of deriving such an index. Instead, we chose to compute a vector from the original HRV signal that carries more useful information. According to our experiments, the return plot (or Poincaré

0276-6547/01 $17.00 © 2001 IEEE

plot) and the difference plot, which are state-space domain features, are promising when used for both stages.

Let an HRV signal be denoted as $X=[x_1, x_2, \ldots, x_p]$. The return map is generated by plotting x_i versus x_{i+1}, $i=1, 2, \ldots, p-1$ while the difference map is constructed by drawing $x_{i+1}-x_i$ versus $x_{i+2}-x_{i+1}$, $i=1, 2, \ldots, p-2$. They are state-space domain features because the return map is actually a 2D projection of trajectory points while the difference map is a 2D projection of trajectory vectors. Figure 1 shows return map and difference map of a normal subject.

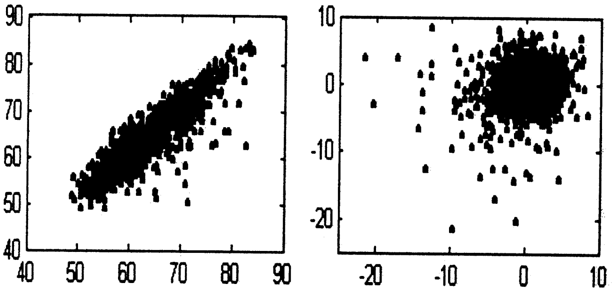

Figure 1. The return map (left) and the difference map (right) of a normal subject

In order to extract feature vectors from the return map and the difference map, we first divided these maps into m-by-m and n-by-n equal-sized lattices. An m-by-m matrix and an n-by-n matrix are associated with the latticed maps in the way that each matrix element is set to the fraction of points lying within the corresponding lattice. The feature vector of the return map and that of the difference map are constructed by re-shape the m-by-m matrix and the n-by-n matrix into m^2-by-1 and n^2-by-1 vectors, respectively.

3. Feature classification

In our work, we employed the k-nearest neighbors to classify the computed feature vectors. Each feature vector computed from a testing data set is compared against all feature vectors computed from training data sets. The k-closest vectors in the training sets are used to determine which class the testing vector belongs to and calculate the corresponding confidence level. To compute the confidence level of each class, we used a distance-based scoring algorithm instead of directly counting the number of class member in the k-nearest neighbors. Assume there are totally f classes and let the distance between the testing vector and its k-nearest neighboring training vectors be denoted as $d_{i,j}$ where $i \in [1, f]$ and $j = 1, 2, \ldots, k$. A weighting factor associated with each $d_{i,j}$ is defined as $w_{i,j} = d_{\max} / d_{i,j}$ where $d_{\max} = \max(d_{i,j})$. The confidence level associated with each class is defined by

$$C_p = \sum_{i=p,\, j} w_{i,j} \Big/ \sum_{j} w_{i,j} \qquad (1)$$

and the testing feature vector is said to belong to class p if C_p is the largest among C_1 to C_f.

4. Experimental results

We used the data sets issued for CinC Challenge 2001 as our training data and testing data. The training data set consists of 50 subjects, half of which experienced PAF (group N) while the other half do not have PAF (group A). Each of the record in the data set is a half-hour ECG recording randomly chosen from a 24-hour ECG record. Two half-hour recordings were extracted from each subject's 24-hour ECG record. For subjects with no documented PAF, these two half-hour recordings were extracted randomly. For subjects experienced PAF, one half-hour recording was chosen at least 45 minutes away from PAF episode and the other half-hour recording was selected right next to a PAF episode. The testing set was obtained from another 50 subjects. The details of the data set are similar to those of the training set except that the number of PAF subjects is unknown.

The feature vector was constructed by a 3600-by-1 vector computed from 60x60 latticed return plot (ranging from 30~230) and a 4900-by-1 vector from 70x70 latticed difference plot (ranging from −150 ~150). A 5-nearest-neighbor algorithm is used as classifier. With the choices of feature vector and classifier, we successfully distinguished 34 out of 50 subjects in the screening part of the competition and 39 out of 50 subjects in the prediction part of the competition.

5. Discussions

Conventional approaches attempt to derive a single index according to certain property of the signal. In addition, most of the approaches are developed based on the noise-free assumption. Some algorithms are capable of maintaining their accuracy when low-percentage, normally or uniformly distributed noises are added. However, real world signals such as human's HRV signal are usually contaminated with noise of unknown distribution and magnitude. Nonlinear indices such as detrended fluctuation analysis (DFA), Fano factor (time domain indices), interval-based periodogram (frequency domain index), capacity dimension, or correlation dimension (state-space domain indices) calculating certain power-law relationship of the signal usually suffer from determining the slope of a curve caused by noise, as shown in Figure 2.

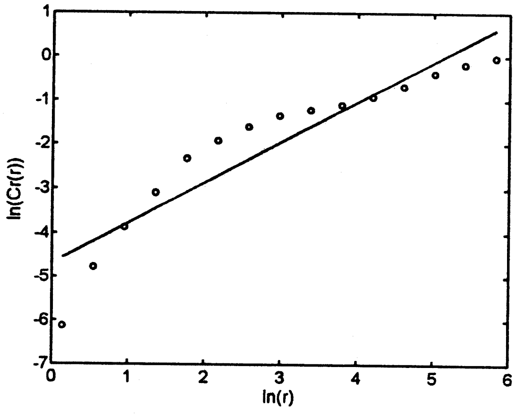

Figure 2. The correlation integral of an HRV signal. As it can be seen, to calculate correlation dimension from the slope of the regression line can be inappropriate.

In addition, most of the above-mentioned algorithms assume data points are measured from a steady system where no transient behavior is observed. However, human's heart is in fact a non-autonomous system, which means the parameters of the system change with time and so do the properties of the system. Under this condition, algorithms measuring such properties that require long duration of data may generate inaccurate results. This can be verified by chopping long HRV signal into short sections and apply the above-mentioned algorithms to compute their corresponding properties in each section. Sometimes the values computed from these sequential sections exhibit large variations, as shown in Figure 4.

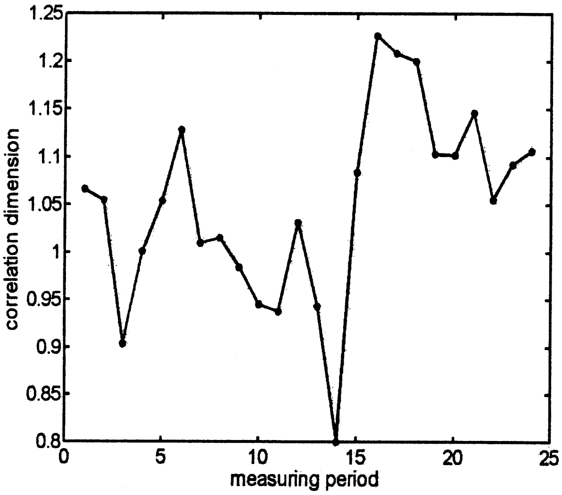

Figure 4. Correlation dimension calculated at different measuring periods showing large variations

In our work, the selected feature vectors are robust to measuring noise in the sense that, under the present of noise, the distribution of data points is similar to that of the original (non-contaminated) points. By "distribution", we mean the values in quantized return plot and difference plot. In addition, the selected feature vectors are also suitable for capturing properties of data generated by non-autonomous system. This is because data measured at different timing of the system are simply added to the pictures of return map and difference map. Therefore the system behavior (i.e. the location of system trajectory in state space) at different period of time can be captured. However, to use either the return map or the difference map alone is inadequate to capture all the properties of the original trajectory. The return map loses the connectivity information while the difference map lacks of the location information. The combination of both tools provides a better way to represent features in the original trajectory.

In the proposed algorithm, we used k-nearest neighbors as classifier in both stages for its capability of providing confidence level on its classification. In our work, we modified the conventional way of computing the confidence level by multiplying weighting (penalty) factors on each of the k neighboring feature vectors according to the distance between them and the testing feature vector. This significantly improved the classification accuracy comparing with that used conventional approach. Moreover, with the confidence level, one can access the reliability of the classification result and decide whether to accept the result or to proceed to another test.

6. Concluding remarks

In this paper, we proposed a two-stage solution algorithm for paroxysmal atrial fibrillation prediction. The algorithm is shown to outperform the conventional single index approach in the sense that it is not only robust to the measuring noise but also suitable for non-autonomous system such as human's heart. In addition, the algorithm provides confidence level for each of its classification result so one can decide whether to accept the result or to proceed to another test. As a whole, the algorithm is simple, easy-to-implement, easy-to-modify, and, most of all, it is reliable.

References

[1] Goldberger AL. Applications of chaos to physiology and medicine. In: Kim JH, Stringer J, editors. Applied Chaos. John Wiley & Sons, Inc, 1992:321-31.

[2] Poon CS, Barahona M, Merrill CK. Method and apparatus for detecting nonlinearity in an electrocardiographic signal. US pattern No. 5792062, issued at Aug. 11, 1998.

[3] Peng CK, Hausdorff JM, Goldberger AL. Fractal mechanisms in neural control: Human heartbeat and gait dynamics in health and disease. In: Walleczek J, editor. Nonlinear Dynamics, Self-Organization, and Biomedicine. Cambridge. Cambridge University Press, 1999.

[4] Kurths J, Voss A, Saparin P, Witt A, Kleiner HJ, Wessel N. Quantitative analysis of heart rate variability. Chaos 1995;5(1):88-94.

[5] Barczynski RJ, Kubasik A. Nonlinear methods in heart rate variability. T.A.S.K. quarterly 1997;1(2):195-207.

[6] Otakar F, Jiri H. Applying nonlinear dynamics to ECG signal processing. IEEE Engineering in Medicine and Biology 1998; 17(2):96-101.

[7] Takens F. in Dynamical Systems and Turbulence. In: Rand DA, Young LS, editors. Lecture Notes in Mathematics. Berlin : Springer- Verlag, 1981; 898:366-381.

[8] Moon FC. Chaotic and fractal dynamics: an introduction for applied scientists and engineers. New York: Wiley, 1992.

[9] Abarbanel HDI. Analysis of observed chaotic data. New York: Springer-Verlag, 1996.

[10] Judd K. Control and chaos. Boston: Birkhäuser, 1997.

[11] Turcott RG, Teich MC. Fractal character of the electrocardiogram: distinguishing heart- failure and normal patients. Annals of Biomedical Engineering 1996; 24:269-93.

Address for correspondence

Ke-Shiuan Lynn
418 Rhodes Hall
School of Electrical and Computer Engineering
Cornell University
Ithaca, NY, 14853
U.S.A.