

An Improved Method for Discriminating ECG Signals using Typical Nonlinear Dynamic Parameters and Recurrence Quantification Analysis in Cardiac Disease Therapy

M. Tang*, C.Q. Chang*, P.C.W. Fung[†]*, K.T. Chau* and F.H.Y. Chan*

*Department of Electrical and Electronic Engineering, The University of Hong Kong,
Pokfulam Road, Hong Kong SAR, PRC
Emails: {mantang, cqchang, ktchau, fhychan}@eee.hku.hk

[†]Division of Medical Physics, Department of Medicine, The University of Hong Kong,
Pokfulam Road, Hong Kong SAR, PRC
Email: hrspfcw@hkucc.hku.hk

Abstract—The discrimination of ECG signals using nonlinear dynamic parameters is of crucial importance in the cardiac disease therapy and chaos control for arrhythmia defibrillation in the cardiac system. However, the discrimination results of previous studies using features such as maximal Lyapunov exponent (λ_{\max}) and correlation dimension (D_2) alone are somewhat limited in recognition rate. In this paper, improved methods for computing λ_{\max} and D_2 are purposed. Another parameter from recurrence quantification analysis is incorporated to the new multi-feature Bayesian classifier with λ_{\max} and D_2 so as to improve the discrimination power. Experimental results have verified the prediction using Fisher discriminant that the maximal vertical line length (V_{\max}) from recurrence quantification analysis is the best to distinguish different ECG classes. Experimental results using the MIT-BIH Arrhythmia Database show improved and excellent overall accuracy (96.3%), average sensitivity (96.3%) and average specificity (98.15%) for discriminating sinus, premature ventricular contraction and ventricular flutter signals.

I. INTRODUCTION

There have been considerably studies of the deterministic chaotic behavior of heartbeat signals and it is highly desirable to control such cardiac chaos [1]. In particular, cardiac chaos control of arrhythmia for defibrillation using “anticontrol” algorithm, which maintains a desired level of chaoticity by achieving certain target values of nonlinear dynamic parameters, has shown to be capable of avoiding pathological behavior [2]. Therefore, the discrimination of ECG signals using nonlinear dynamic parameters is of crucial importance in the cardiac disease therapy and chaotic control for arrhythmia defibrillation in the cardiac system. Typical nonlinear dynamic parameters are the maximal Lyapunov exponent (λ_{\max}) and correlation dimension (D_2). Although there have been a number of studies in discriminating ECG signals using these nonlinear dynamic parameters [3], the results are not satisfactory. In our previous study [4], it was found that better discrimination results can be obtained if the Kantz algorithm is used to estimate λ_{\max} in ECG signals than the Wolf algorithm [3]. In this paper, we propose to improve the discrimination power by employing multiple features [5]. More precisely, another nonlinear dynamic

parameter will be incorporated to this classifier with the two typical parameters, λ_{\max} and D_2 , to form a multi-feature discriminator with improved discrimination power. Eight features from recurrence quantification analysis (RQA) [6], which is the extension of a graphical method called recurrence plot (RP) analysis [7], were computed. Fisher’s method predicts that V_{\max} , which is the maximum vertical line length, is the best feature to distinguish different ECG classes. The discrimination results by using each of the eight features match Fisher’s prediction. Thus, the additional parameter to the multi-feature discriminator is V_{\max} . Since D_2 is rather difficult to be estimated accurately in practice, the discrimination performance of classifiers based solely on this feature is usually limited [3]. To overcome this problem, an improved method to estimate D_2 is also proposed in this paper. To evaluate the performance of the proposed approach, multi-feature discrimination, which combines λ_{\max} , D_2 and V_{\max} , is applied to ECG datasets from the MIT-BIH Arrhythmia Database to discriminate normal and pathological ECG signals. Experimental results show that the overall accuracy, average sensitivity and average specificity obtained for discriminating sinus, premature ventricular contraction and ventricular flutter signals are respectively 96.3%, 96.3% and 98.15% which are significantly better than the conventional methods in [3].

II. METHODS

Phase space reconstruction [8] is the first step to compute the dynamical parameters λ_{\max} , D_2 and V_{\max} . Given the time series $x(t)$, an m -dimensional phase portrait Y_t is reconstructed such that $Y_t = \{x(t), x(t+\tau), \dots, x(t+(m-1)\tau)\}$, where τ is the time delay determined by finding the first minimum of mutual information. When $m > 2D + 1$, where D is the fractal dimension, the reconstructed phase space is topologically equivalent to the original system.

In [4], the Kantz algorithm was proposed to estimate λ_{\max} of ECG signals. λ_{\max} is a quantitative measure to describe how small changes in the state of a system grow at an exponential rate in the phase space. To compute λ_{\max} , the

logarithm of the average evolved distance separation after a relative time interval Δn over all neighbours of the trajectory point Y_t within the region of a predefined small diameter is computed and averaged over a set of trajectory points as:

$$S(\Delta n) = \frac{1}{T} \sum_{t=1}^T \log_2 \sum_{i \in U(Y_t)} \frac{dist(Y_t, Y_i; \Delta n)}{|U(Y_t)|}, \quad (1)$$

where the distance is defined as $dist(Y_t, Y_i; \Delta n) = |x_{t+(m-1)\tau+\Delta n} - x_{i+(m-1)\tau+\Delta n}|$, $U(Y_t)$ is the set of neighbourhoods of Y_t , and $|U(Y_t)|$ is the number of neighbours of Y_t . Within some ranges of Δn , $S(\Delta n)$ exhibits a linear increase in Δn whereas its slope is an estimate of λ_{\max} .

On the other hand, the feature D_2 described in [3] represents the geometric structure of the underlying system and is defined by:

$$D_2 = \lim_{r \rightarrow 0} \log_2 C_m(r) / \log_2 r, \quad (2)$$

where m is the dimension of the phase space, and $C_m(r)$ is the correlation integral defined as:

$$C_m(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(r - |Y_i - Y_j|) \quad (3)$$

where Θ is the Heaviside step function satisfying $\Theta(x) = 1$ for $x > 0$ and $\Theta(x) = 0$ for $x \leq 0$. According to the algorithm in [3], D_2 is estimated by the local slope in the scaling region of the $\log_2 C_m(r) - \log_2 r$ plot. However, there is a practical difficulty in determining accurately where the proper scaling region is. There have been some methods to solve this problem in the literature [3], but they are not satisfactory at least in the application of ECG signal analysis. A new method is proposed in this paper. Despite the fact that D_2 is unaffected under different m as long as $m > 2D + 1$, practically its estimated value will not be exactly the same for different m . Therefore it makes sense to find a scaling region in the $\log - \log$ plot where the slope in this region is least affected by the change of m . For this purpose we make such plots for different m , and let the scaling region to be that around the following estimated point r_0 :

$$r_0 = \arg \max_r \frac{\text{mean}(d \log_2 C_m(r) / d \log_2 r)}{\text{std}(d \log_2 C_m(r) / d \log_2 r)}, \quad (4)$$

where $\text{std}(y)$ stands for the standard deviation of y . D_2 is then estimated by the average slope at this point:

$$D_2 = \text{mean}(d \log_2 C_m(r_0) / d \log_2 r_0). \quad (5)$$

The method of recurrence plot (RP) was first introduced to visualize the time dependent behavior of the dynamics of systems [7]. It represents the recurrence of the phase space trajectory to a certain state, which is a fundamental property of deterministic dynamical systems. The main step of this visualization is the calculation of the $N \times N$ matrix,

$$R_{i,j} = \Theta(\epsilon_i - \|\vec{Y}_i - \vec{Y}_j\|), \quad i, j = 1, \dots, N, \quad (6)$$

where ϵ_i is a cutoff distance defining a sphere centered at \vec{Y}_i , $\|\cdot\|$ is a norm (e.g. the Euclidean norm), and $\Theta(x)$ is the

Heaviside function. If \vec{Y}_j falls within the sphere, the state will be close to \vec{Y}_i and thus $R_{i,j} = 1$.

Recurrence quantification analysis (RQA) was then developed to quantify an RP [9] [10]. Some measures are extracted from the RP based on the recurrence point density and the diagonal structures in the recurrence plot, which are the recurrence rate (RR), the determinism (DET), the maximal length of diagonal structures (L_{\max}), the average length of the diagonal structures (L), and the entropy (ENT). Gao has recently introduced some measures from the RP based on the vertical structures in the RP, which are the laminarity (LAM), the trapping time (TT), and the maximal length of vertical structures (V_{\max}) [11].

Fisher's method, as described below, was used to determine which feature among the eight extracted from the RP is the best to distinguish the three ECG classes. Denote the overall mean for each feature in the three ECG classes as:

$$\bar{\mu}_Y = (\sum_{i=1}^g \mu_{iY}) / g, \quad g = 3, \quad (7)$$

where μ_{iY} is the mean of a feature in the i th class. Fisher's discriminant, which is the ratio of the data variability among the classes (the square of difference between each class mean and overall mean), to the common variability within classes (the variance of the total population), is calculated for each feature f_k from the RP:

$$a_k = \sum_{i=1}^g (\mu_{iY} - \bar{\mu}_Y)^2 / \sigma_Y^2, \quad (8)$$

where σ_Y^2 is the variance of population Y . Fisher's discriminant measures the discrimination power of each feature. The best feature is $f_{k_{opt}}$ where $k_{opt} = \arg \max_k (a_k)$.

With the estimated dynamic parameters, multi-feature based Bayesian and minimum distance discriminators can be constructed and used to classify normal and abnormal ECG signals.

III. EXPERIMENTAL DESIGN AND RESULTS

The proposed method is evaluated using the ECG signals from the MIT-BIH Arrhythmia Database [12]. The data sets used were sinus, premature ventricular contraction (PVC) and ventricular flutter (VF). Each type consists of 54 independent signals, each with 3-second long.

After reconstructing the phase space, we showed the associated maximal Lyapunov exponent curves $S(\Delta n)$ against Δn of the three ECG classes are shown in Fig. 1, where the dotted lines are the slope estimators obtained by least-square fit and the correlation integral curve $\log_2 C(r)$ against $\log_2 r$ and its instantaneous slope of a sinus signal in Fig. 2, where the curves from the bottom to top correspond to $m = 3, 4, \dots, 10$, respectively. From Fig. 1, the three slopes in the linear region are all positive and distinct, indicating the different chaotic natures of ECG signals in these three classes. In Fig. 2, D_2 is estimated to be 2.65.

The recurrence plots of the three ECG classes are shown in Fig. 3 (a), (b) and (c) where black dots mark a recurrence.

The ratio a_k in Fisher's method is computed for each feature from RP and it is shown in Table I. The Fisher

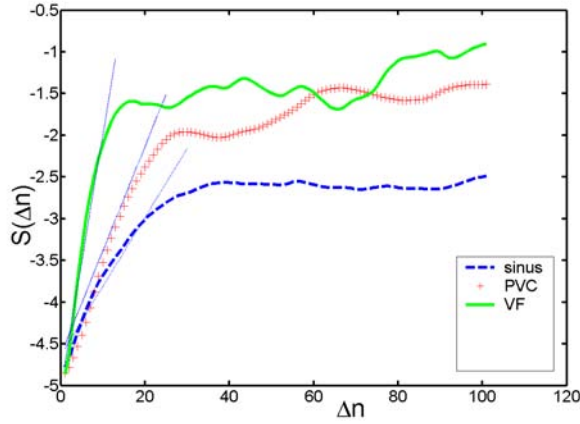


Fig. 1. $S(\Delta n)$ as function of Δn of sinus, PVC and VF

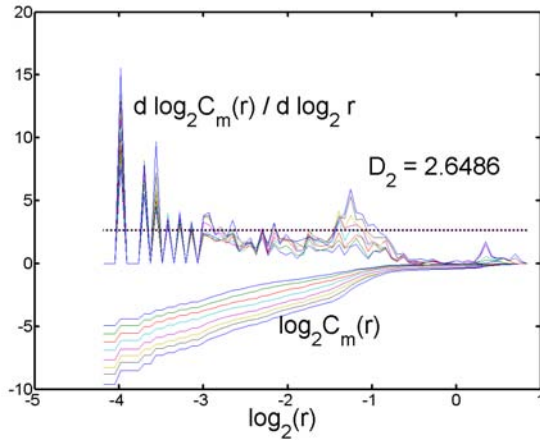
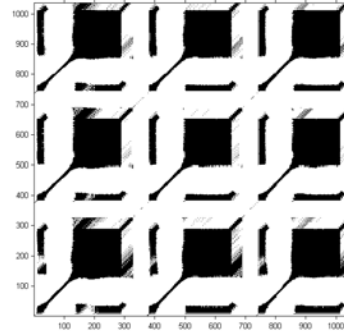


Fig. 2. Correlation and its instantaneous slope curves of a sinus signal

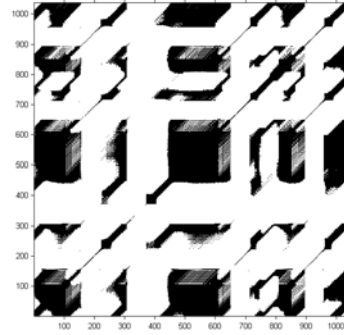
prediction matches the classification result that V_{\max} gives the highest overall accuracy if each of the ten features is used for distinguishing the three ECG classes (Table II). V_{\max} reveals information about the time duration of the laminar states, i.e. chaos-chaos transitions, thus making the investigation of intermittency possible. V_{\max} is analogous to the standard RQA measure, the maximal length of diagonal structures L_{\max} [10].

Since V_{\max} from RP has attained highest overall accuracy in the classification, it is incorporated into the multi-feature classifier with λ_{\max} and D_2 to improve the classification performance.

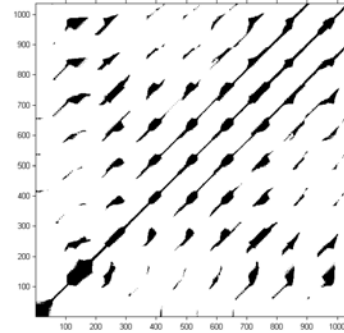
Table IV shows λ_{\max} , D_2 and V_{\max} of the three types of ECG signals. The discrimination results for the Bayesian and minimum distance discriminators using respectively the λ_{\max} , D_2 , V_{\max} and their combination as the feature vectors are shown in Table III. The overall accuracy, average sensitivity and average specificity for discriminating the three ECG classes using the proposed multi-feature Bayesian discriminator are 96.3%, 96.3% and 98.15%, respectively. It can be seen that, for the same set of feature vectors, the use of



(a) Recurrence plot of a sinus



(b) Recurrence plot of a PVC signal



(c) Recurrence plot of a VF signal

Fig. 3. Recurrence plot of sinus, PVC, and VF signals

the Bayesian discriminator gives better discrimination results than minimum distance discriminator because it makes use of the prior information in the sample statistics. The improved discriminating result suggests that the combination of these three chaotic parameters could effectively describe different levels of chaoticity of the three types of ECG signals. Hence, these three dynamic parameters have great potential in controlling cardiac chaos for arrhythmia defibrillation by using the anticontrol algorithm for cardiac disease therapy.

TABLE I
VARIABILITY BETWEEN THE GROUPS OF Y-VALUES RELATIVE TO THE COMMON VARIABILITY WITHIN CLASSES

| Feature | RR | DET | L | L_{\max} | ENT | LAM | TT | V_{\max} |
|---------|------|------|------|------------|------|------|------|------------|
| a_k | 2.41 | 1.39 | 1.85 | 1.55 | 2.02 | 1.59 | 2.49 | 2.56 |

TABLE II
CLASSIFICATION RESULTS USING EACH FEATURE FROM RP BY BAYESIAN CLASSIFIER

| Feature | V_{\max} | TT | RR | ENT | L | LAM | L_{\max} | DET |
|----------------------|------------|------|------|------|------|------|------------|------|
| Overall Accuracy (%) | 88.9 | 81.5 | 78.4 | 71.6 | 66.7 | 63.0 | 60.5 | 62.7 |

TABLE III
DISCRIMINATION RESULTS IN % FOR BAYESIAN AND MINIMUM DISTANCE DISCRIMINATORS USING DIFFERENT CHAOTIC DYNAMIC FEATURES (BAY.: BAYESIAN, MD: MINIMUM DISTANCE)

| Discriminator | Bay. | MD | Bay. | MD | Bay. | MD | Bay. | MD | Bay. | MD |
|---------------------|------------------|------------------|-------|-------|------------|------------|-----------------------|-----------------------|---------------------------------|---------------------------------|
| Feature | λ_{\max} | λ_{\max} | D_2 | D_2 | V_{\max} | V_{\max} | λ_{\max}, D_2 | λ_{\max}, D_2 | $\lambda_{\max}, D_2, V_{\max}$ | $\lambda_{\max}, D_2, V_{\max}$ |
| Average sensitivity | 88.02 | 77.78 | 40.95 | 44.31 | 88.89 | 85.8 | 91.16 | 86.33 | 96.3 | 93.21 |
| Average specificity | 91.17 | 88.89 | 82.43 | 73.18 | 94.44 | 92.9 | 93.30 | 92.07 | 98.15 | 96.6 |
| Overall accuracy | 85.19 | 77.78 | 49.38 | 45.06 | 88.89 | 85.8 | 88.89 | 85.26 | 96.3 | 93.21 |

TABLE IV
MAXIMAL LYAPUNOV EXPONENTS, CORRELATION DIMENSIONS AND MAXIMUM VERTICAL LINE LENGTH OF THE 3 ECG CLASSES

| | Sinus | PVC | VF |
|------------------|--------------------|--------------------|-------------------|
| λ_{\max} | 15.48 ± 2.23 | 21.54 ± 3.00 | 31.97 ± 8.37 |
| D_2 | 2.78 ± 1.07 | 1.92 ± 0.70 | 2.14 ± 0.48 |
| V_{\max} | 162.91 ± 13.92 | 211.41 ± 42.91 | 34.59 ± 27.03 |

IV. CONCLUSION

We have applied nonlinear dynamic analysis to quantitatively differentiate sinus, PVC and VF by three nonlinear dynamic parameters λ_{\max} , D_2 and V_{\max} . Kantz algorithm, instead of Wolf algorithm in previous studies, is used to estimate λ_{\max} with improved discriminating power. Great improvement in D_2 estimation is also obtained by introducing a new method to determine the scaling region of the correlation integral plot. Experimental results have verified Fisher's prediction that the maximal vertical line length (V_{\max}) among the other seven features from RQA is the best to distinguish different ECG classes. Experimental results using the MIT-BIH Arrhythmia Database show that significant improvement can be obtained by using λ_{\max} , D_2 and V_{\max} as a combination in multi-feature discrimination. It also reveals the possibility of accurate cardiac chaos control using the anticontrol algorithm and better cardiac disease therapy, as these three parameters define distinct levels of chaoticity among different ECG classes.

ACKNOWLEDGMENTS

This work is supported in part by Hong Kong RGC grants under HKU7180/03E and N_HKU703/03.

REFERENCES

- [1] Qu Z, "Nonlinear Dynamic Control of Irregular Cardiac Rhythms," *J Cardiovasc Electrophysiol*, vol.15, pp.1186-1187, 2004.
- [2] R.Ramaswamy, S. Sinha and N. Gupte, "Targeting chaos through adaptive control," *Physical Review E*, vol.57, pp.R2507-2510, 1998.
- [3] M. I. Owis, A. H. Abou-Zied, A.B. M. Youssef, and Y. M. Kadah, "Study of Features Based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification," *IEEE Transaction on Biomedical Engineering*, vol.49, pp.733-736, 2002.
- [4] M. Tang, C.Q. Chang, P.C.W. Fung, K.T. Chau, F.H.Y. Chan, "An Effective Way of Distinguishing Premature Ventricular Contraction from Sinus based on Analysis of Chaotic Feature," *Proc. MEDSIP2004*, Malta, pp.85-90, 2004.
- [5] W. C. Xu, H. F. Tse, F. H. Y. Chan, P. C. W. Fung, K. L. F. Lee, and C. P. Lau, "A New Bayesian Discriminator for Detection of Atrial Tachyarrhythmias," *Circulation*, vol.105, pp.1472-1479, 2002.
- [6] L. L. Trulla, A. Giuliani, J. P. Zbilut, and C. L. Webber, Jr., "Recurrence quantification analysis of the logistic equation with transients," *Physics Letters A*, vol.223, no.4, 255-260, 1996.
- [7] J. P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence Plots of Dynamical Systems," *Europhysics Letter*, vol.5, pp.973-977, 1987.
- [8] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed., Cambridge University Press, Cambridge, 2004, pp.35-36.
- [9] C.L. Webber, Jr., and J.P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of Applied Physiology*, vol.76, no.2, pp.965-973, 1994.
- [10] J. P. Zbilut, V. L. Webber, Jr., "Embeddings and delays as derived from quantification of recurrence plots," *Physics Letters A*, vol.171, no.3-4, pp.199-203, 1992.
- [11] J. Gao, H. Cai, "On the structures and quantification of recurrence plots," *Physics Letters A*, vol.270, pp.75-87, 2000.
- [12] The MIT-BIH Arrhythmia Database, Harvard-MIT Div. Health Sci. Technol., Cambridge, MA., 1997, 3rd ed.