Statistical Distance of Quadtree Partitioned Attractors: Detecting Changes in the Behavior of ECG Signals

JCB de Figueiredo, SS Furuie

Heart Institute (InCor), University of São Paulo Medical School, Brazil

Abstract

In this work we describe a method to detect different types of behaviors in an ECG signal using quadtree-partitioned attractors. The method is based on the calculation of the statistical distance between two attractors in a phase space of a dynamical system where a quadtree decomposition was applied. A short portion of an ECG signal with normal behavior (of a patient) is used to reconstruct (using a time-delay technique) the trajectory of an attractor in a 2D phase space. The phase-space is then partitioned using a quadtree algorithm. The distributions in the quadtree mesh are interpreted as statistical distributions corresponding to the normal dynamical behavior of the ECG recording in the phase-space. An algorithm is then used to compute the correlation distance between this distribution and all other distributions that are built using reconstructed attractors from a sliding temporal window over the signal. It was noticed, in normal cases, that the distance was almost constant and below a threshold. For pathological cases (with abnormal transients), on the abnormal portion of ECG, the distance increased consistently with morphological changes.

1. Introduction

Many methods for ECG analysis are based in studies where it is assumed that the discrete time evolution of an ECG variable, obtained by monitoring, can be described by a differential dynamic system in a phase-space [1]. This time evolution generates a trajectory in the phase-space (named attractor) that can be interpreted as a statistical distribution [2]. In this paper, we present an analysis of ECG signals using an approach based on the calculation of a statistical distance (correlation distance) between two attractors in a phase-space. We start from a short portion of a discrete ECG signal, chosen to represent the supposed normal behavior of the ECG. A time embedding procedure is then used to generate from the signal an attractor in a low dimension phase-space. We called this attractor *reference attractor*. The points of the reference attractor are distributed in a uniform mesh in such a way that the value of each cell will represent the total of points that fall inside the cell. This distribution is then partitioned using a quadtree-decomposition algorithm. The threshold used as criterion for cell division corresponds to the medium cell value in the uniform mesh. We then compute the correlation distance between this quadtree *reference distribution* (assumed to be normal) and all other distributions that are built using attractors reconstructed from a sliding temporal window over the signal. The points of each one of these attractors are distributed over the quadtree mesh generated from the reference attractor. So we can measure how each of these distributions deviates from the distribution considered normal. This operation is performed using the correlation distance.

Figueiredo & Furuie, in a previous work [3], used this approach considering the uniform mesh as reference distribution. In the present work we will compare these previous results, showing that the use of quadtree partitions allows a more sensible detection of changes in the signal behavior. Besides, the use of quadtree meshes also permits the construction of more efficient (fast) algorithms to calculate these distances.

2. Method

Consider $E(t)$ as a normalized ECG signal $(0 \le E(t) \le 1)$. We can apply a time-delay embedding technique to reconstruct a phase-space representation of this signal that is topologically equivalent to the original phase-space [4]. The generation of this phase-space is achieved by choosing an integer m (the embedding dimension) and a time delay τ, thus forming the vector:

$$\vec{\xi}(t) = \{E(t), E(t+\tau), E(t+2\tau), ..., E(t+(m-1)\tau)\} \quad (1)$$

In principle, the dimension m of this phase-space is infinite, however some works showed that a low dimension phase- space (m=2) could be used to capture the important characteristics of an ECG signal [1]. A special care must be taken when choosing the value of τ. If τ is too small, $E(t)$ is close to $E(t+\tau)$ and the attractor is

0276-6547/01 $17.00 © 2001 IEEE

Computers in Cardiology 2001;28:521-524.

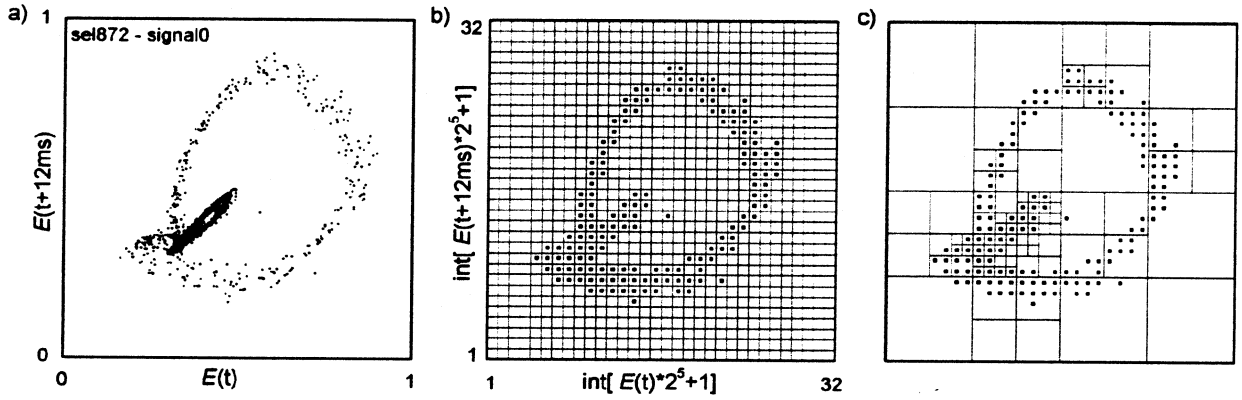

Figure 1. a) Attractor reconstructed from a normalized ECG b) The same attractor distributed over a uniform mesh with 5 bits of resolution (each cell represents the total of points that fall inside the cell). c) A quadtree mesh generated from the distribution in the uniform mesh (the threshold used was the mean value of the cells in the uniform mesh).

compressed to the vicinity of the phase-space diagonal. So it cannot be completely extended. If τ is too large, the attractor will fold and the phase graph deforms. A number of methods for finding a good value for τ have been proposed, but the practical method most used is visually inspect the attractors and look for a τ value that unfolds the reconstructed attractor but does not cause a folding back.

Figure 1a shows the normalized attractor obtained from the first 24s of the ECG signal *sel872* from the MIT Supraventricular Arrhythmia Database in Physionet [5]. This attractor was reconstructed using m=2 and τ=12ms. A discrete statistical distribution can be constructed from this attractor considering the vector

$$\vec{\Xi}(t) = \text{int}[\vec{\xi}(t)(2^b - 1) + 1] \qquad (2)$$

where b represents the resolution (in bits) for the space discretization. Figure 1b shows, for b=5, the statistical distribution obtained from the Figure 1a. In this figure, each mesh cell pi (i=1,2,3,...,2b) corresponds to the total of points of $\vec{\Xi}(t)$ that fall inside the cell i. Figueiredo & Furuie [3] used uniform meshes to construct statistical distributions to compare parts of an ECG signal. However, some problems can be pointed with the use of uniform meshes. The principal is that uniform meshes do not take in account the fact that some regions of the ECG attractor have low density and high spatial dispersion, as in the QRS complex for example, while other regions have high densities and low dispersion, as in the ST segment. Therefore, small deviations in the QRS complex are treated in the same way of small deviations in the ST complex, because all cells have the same weight. Other problem is related to the elevated time necessary to compare two distributions over uniform meshes with 2b, cells when b becomes large (high resolution analysis).

In order to solve these problems, we considered to use a quadtree-decomposition over the uniform mesh. In the quadtree-decomposition, the basic principle is to cover a planar region of interest by a square and then, recursively, partitionate squares into smaller squares until having (inside the squares) less points than an established threshold or the squares cannot be divided any more (resolution limit). Figure 1c shows a quadtree-decomposition of the distribution in Figure 1b. We adopted as threshold the medium value of the cells in the uniform mesh. That resulted in a new quadtree mesh, with 91 cells (933 less than the uniform mesh), and a new distribution, p={pi} (i=1,2,3,...,91), where the cell sizes are related to the attractor density points in the vicinity of the each cell.

To compare attractors, we will measure the correlation distance between the distributions p={pi} and g={gi} (i=1,2,3,...,k) generated by two attractors when its points are distributed over the reference mesh (that can be uniform or quadtree). This distance is given by

$$Dpg = 1 - \max[0, C(p,g)] \qquad (3)$$

where

$$C(p,g) = \frac{\sum_{i=1}^{k}(p_i - \bar{p})(g_i - \bar{g})}{\sqrt{\sum_{i=1}^{k}(p_i - \bar{p})^2 \sum_{i=1}^{k}(g_i - \bar{g})^2}} \qquad (4)$$

is the correlation function and $\bar{p}$ and $\bar{g}$ the mean values of the distributions. The equation (3) was written in such a way that the Dpg values were limited to the interval [0,1].

3. Results

To detect changes in the dynamics of an ECG signal, we will use the correlation distance (equation (3)) to calculate differences between a reference distribution $\{pi\}$ (obtained from a part of the signal supposed normal) and all others distributions $\{gi\}$ that are built using parts originated from a sliding temporal window over the signal. The points of each one of the reconstructed attractors (from the sliding windows) are distributed over the quadtree mesh generated with the reference distribution, so we can measure how each one of these distributions deviates from the reference distribution.

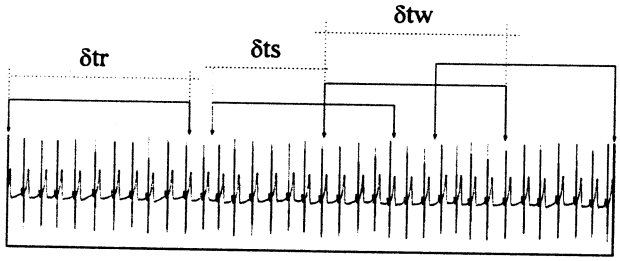

Figure 2. Scheme for the ECG signal analysis using the method described in this work.

As example we analyzed the discrete signal *sel872* of the MIT public database in Physionet [5]. The maximum value of this signal was normalized to unity. The first initial δtr=24s was chosen to reconstruct, using a time delay embedding with m=2 and τ=12ms (equation (1)), an attractor from which the reference distribution $\{pi\}$ was generated. A value of b=9 (see equation (2)) was adopted. It resulted in a uniform mesh of 262144 cells for the phase-space discretization. Note that we were not concerned if this initial part of the signal is normal from the pathological point of view. The concern was to choose a part of the signal that approximately represented the common behavior in the whole signal. In this case, the initial 24s satisfied the requirement. Using windows of δtw=8s separated by δts=3s, we ran the signal reconstructing, for each window, the attractors corresponding to the distributions $\{gi\}$ and calculating the distances *Dpg* (equation (3)) between these distributions and the reference one. The calculated values of *Dpg* were attributed to the time instant corresponding to the center of each sliding window. Figure 2 illustrates this scheme and Figure 3 shows the quadtree mesh obtained with the above parameters.

In figure 4, we can see the results obtained using the uniform mesh and the quadtree mesh as reference distribution. The low values of *Dpg* correspond to parts of the signal where the dynamic behavior is similar to the reference interval. These values oscillate around *Dpg*=0.2 for both meshes. The regions where the correlation distance presents pronounced peaks indicate an accentuated change in the dynamic behavior of the ECG. In these cases, the quadtree mesh presented more sensibility to detect the changes as we can see comparing the peak amplitudes of the two curves. In addition, the time necessary to compute the curve using the quadtree mesh was of the order of 13 times faster than using the uniform mesh. It is due to the fact the number of cells in the quadtree mesh (1516) was about 173 times smaller than the number of cells in the uniform mesh (262144).

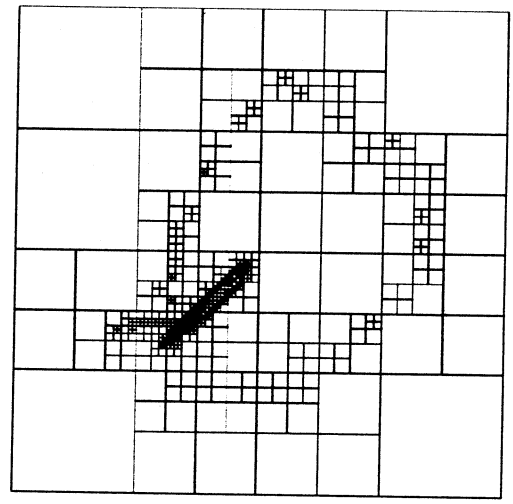

Figure 3. Quadtree mesh generated from the first 24s of sel872 for: m=2,τ=12ms and b=9. (The threshold used was the mean value of the cells in the uniform mesh).

For identifying the kind of change occurred in the peaks, we need to reconstruct the attractors and compare with the reference attractor. In figure 4 we selected three time instants, two corresponding to the highest peaks in the *Dpg* graph (indicating the greatest changes in the dynamical behavior). These instants are t=05:32 and t=08.13. We also chose another instant (t=09:54) where the distance *Dpg* presented a low value. For each one of these instants, we took the 8s of the window used for obtaining the *Dpg* value and used it to reconstruct the corresponding 2D attractor. The results are showed in figure 5. We can notice that the dynamics in the instants t=05:32 (Figure 5b) and t=08:13 (Figure 5c) present large differences in relation to the dynamical behavior used as reference for the signal (attractor in Figure 5a). The attractor of Figure 5b presents significant morphological changes in some QRS segments and the attractor of the Figure 5c seemingly presents fluctuations in the segments P and ST segments (that could be caused by noise). In the Figure 5 we can also see that, in concordance with the *Dpg* graph, the attractor of the Figure 5d have a dynamical behavior close to the reference.

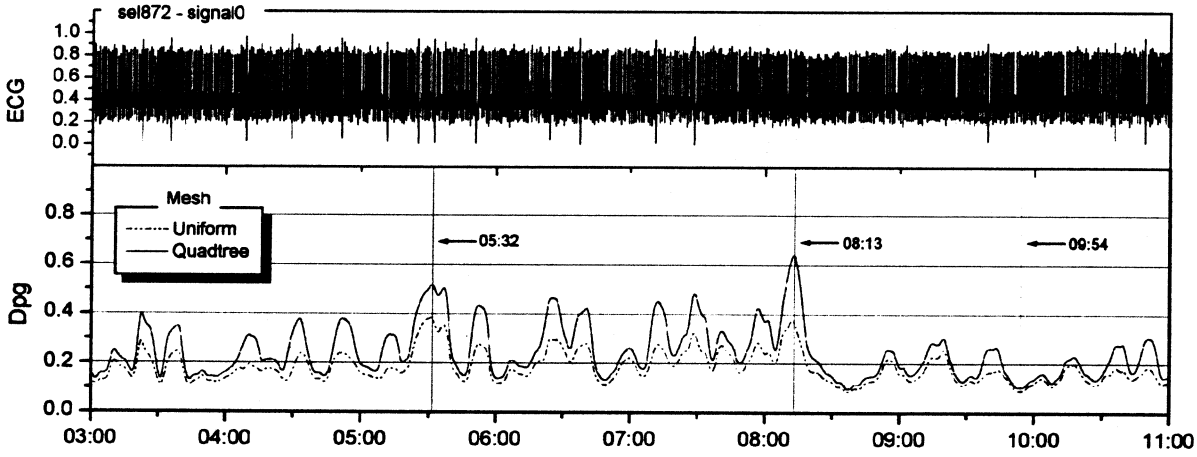

Figure 4. *Dpg* analysis of Signal *sel872* from MIT database for: m=2,τ=12ms and b=9.

4. Conclusions

In this work, we showed that the correlation distance can be used for detection of changes in the behavior of an ECG signal when the normal (or standard) behavior of the signal is known. The simple form of the equation (3) can be easily implemented and produces results more quickly than other numerical methods used to obtain measures for dynamic comparison, such as the Integral of Correlation [2]. The use of quadtree-decomposition for phase-space blocking makes the processing faster and more accurate due to the fact the phase-space blocking is adjustable with the local density of points of the attractor.

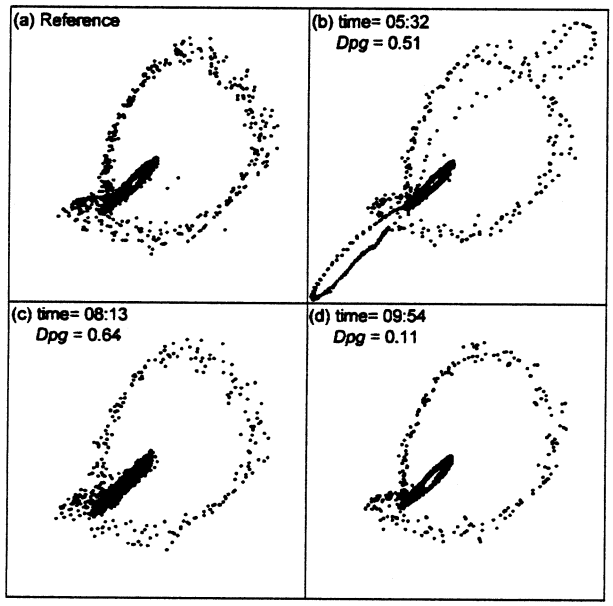

Figure 5. Attractors corresponding to the time instants indicated in the figure 4.

In spite of this work has presented the analysis of only one ECG signal, other signals were analyzed and the results obtained were also good, though further studies are necessary to solve problems such as the appropriate choice of the parameters and the use of the distances when comparing different signals. The next stage will be the comparison of normal and pathological signals in the attempt of establishing the existence or not of defined distances between different reference attractors. Finally, the robustness of this method should also be investigated regarding to the existence of noise in the signals or when the signals have low temporal resolution.

References

[1] Wang Z, Ning X, Zhang YD, G. Nonlinear dynamic characteristics analysis of synchronous 12-Lead ECG signals. IEEE Engineering in Medicine and Biology 2000; September/October: 110-5.

[2] Diambra L. Divergence measure between chaotic attractors. Physical Review E 2001; 64:0352XX

[3] de Figueiredo JCB, Furuie S. Statistical distances to detect changes in the normal behavior of ECG-Holter signals. SPIE - Medical Imaging 2001; San Diego: SPIE PRESS; 2001.

[4] Takens F. Detecting strange attractors in turbulence. Lecture Notes in Mathematics vol 898 1981;366-81.

[5] Goldberger AL, Amaral LA, Glass L , et al. PhysioBank, PhysioToolkit, and PhysioNet - Components of a new research resource for complex physiologic signals. Circulation 2000; 101:E215-e220.

Address for correspondence.

Júlio César Bastos de Figueiredo
Instituto do Coração (InCor) - HCFMUSP
Av. Dr. Enéas Carvalho de Aguiar, 44
São Paulo - SP - 05403-000
Email: Julio.Figueiredo@incor.usp.br